Young suns and infant planets: probing the origins of solar systems

Bohn, A.J.

Chapter 1

Introduction

Are we alone in the Universe? This question has fascinated humans for thousands of years, and it is still one of the most intriguing yet unsolved problems to date. For many centuries and millenia, this question has been of mostly philosophical nature, driven by the human desire to understand their place and purpose within the cosmos. In this thesis introduction I will briefly review past discourses on humanity's belief in extraterrestrial life (see Section 1.1). From a theoretical concept that was already discussed more than 2'000 years ago, an active search for signposts of intelligent life outside Earth emerged within the last century. Merely three decades ago, humankind started to discover planets around other stars than our Sun, of which some even might exhibit favorable conditions to host *life*. In Section 1.2 I will introduce this recent *exoplanet revolution*. I will discuss planet detection methods, challenges, and future prospects in this rapidly evolving field. Several vital questions remain unanswered, as for instance regarding the formation and evolution of planetary systems. In Section 1.3 I will describe how current research aims to understand these phenomena. Special focus will be given to the contributions that originated from my work during the past four years (see Section 1.4). Driven by this enormous progress and supported by major technological developments, humankind might even succeed in obtaining first signs of life outside Earth within the next century.

1.1 The search for life in the Universe

Already in ancient Greece, philosophers of various schools were arguing whether our Earth was unique, or there exist several other worlds that might harbor equally intelligent life. For instance, Metrodorus of Chios (4$^{\text{th}}$ century BCE) was a keen supporter of the theory on the plurality of worlds (Pseudo-Plutarch, Placita Philosophorum 879b-c; Greek version published by Bernardakis 1893; translation provided by Goodwin 1874):

> Μητρόδωρος δέ φησιν ἄτοπον εἶναι ἐν μεγάλῳ πεδίῳ ἕνα στάχυν γενηθῆναι καὶ ἕνα κόσμον ἐν τῷ ἀπείρῳ.
>
> ὅτι δ᾽ ἄπειρος κατὰ τὸ πλῆθος, δῆλον ἐκ τοῦ ἄπειρα τὰ αἴτια εἶναι:

> To Metrodorus it seems absurd, that in a large field one only stalk should grow, and in an infinite space one only world exist;
>
> and that this universe is infinite is manifest by this, that there are causes infinite.

εἰ γὰρ ὁ μὲν κόσμος πεπερασμένος, ὅτι δ᾽ αἴτια πάντα ἄπειρα, ἐξ ὧν ὅδε ὁ κόσμος γέγονεν, ἀνάγκη ἀπείρους εἶναι.	Now if this world were finite and the causes which produced it infinite, it is necessary that the worlds likewise be infinite;
ὅπου γὰρ τὰ πάντα αἴτια, ἐκεῖ καὶ τὰ ἀποτελέσματα αἴτια δ᾽ ἤτοι αἱ ἄτομοι ἢ τὰ στοιχεῖα.	for where all causes do concur, there the effects also must appear, let the causes be what they will, either atoms or elements.

Opposing to this *principle of plenitude* (Lovejoy 1936), several of the most renowned natural scientists from this time did not believe in the existence of further Earth analogs. Among them was Plato ($5^{th} - 4^{th}$ century BCE), who argued (Plato, Timaeus 31a-b; Greek version published by Burnet 1903; translation provided by Lamb 1925):

πότερον οὖν ὀρθῶς ἕνα οὐρανὸν προσειρήκαμεν, ἢ πολλοὺς καὶ ἀπείρους λέγειν ἦν ὀρθότερον;	Are we right, then, in describing the Heaven as one, or would it be more correct to speak of heavens as many or infinite in number?
ἕνα, εἴπερ κατὰ τὸ παράδειγμα δεδημιουργημένος ἔσται.	One it must be termed, if it is to be framed after its Pattern.
τὸ γὰρ περιέχον πάντα ὁπόσα νοητὰ ζῷα μεθ᾽ ἑτέρου δεύτερον οὐκ ἄν ποτ᾽ εἴη:	For that which embraces all intelligible Living Creatures could never be second, with another beside it;
πάλιν γὰρ ἂν ἕτερον εἶναι τὸ περὶ ἐκείνω δέοι ζῷον, οὗ μέρος ἂν εἴτην ἐκείνω, καὶ οὐκ ἂν ἔτι ἐκείνοιν ἀλλ᾽ ἐκείνῳ τῷ περιέχοντι τόδ᾽ ἂν ἀφωμοιωμένον λέγοιτο ὀρθότερον.	for if so, there must needs exist yet another Living Creature, which should embrace them both, and of which they two would each be a part; in which case this Universe could no longer be rightly described as modeled on these two, but rather on that third Creature which contains them both.
ἵνα οὖν τόδε κατὰ τὴν μόνωσιν ὅμοιον ᾖ τῷ παντελεῖ ζῴῳ, διὰ ταῦτα οὔτε δύο οὔτ᾽ ἀπείρους ἐποίησεν ὁ ποιῶν κόσμους, ἀλλ᾽ εἷς ὅδε μονογενὴς οὐρανὸς γεγονὼς ἔστιν καὶ ἔτ᾽ ἔσται.	Wherefore, in order that this Creature might resemble the all perfect Living Creature in respect of its uniqueness, for this reason its Maker made neither two Universes nor an infinite number, but there is and will continue to be this one generated Heaven, unique of its kind.

Also Aristotle (4^{th} century BCE), one of Plato's most famous apprentices, was heavily disputing the theory of life outside Earth, as this hypothesis was in stark contrast with his physical model of the cosmos. Even though some ancient scholars did not agree with this geocentric viewpoint of Aristotle, the broad majority accepted this cosmology in the years to come. Among the small opposition was the Roman philosopher Lucretius (1^{st} century BCE), who said (Lucretius, De Rerum Natura, Liber II, 1067-1076; translation provided by Rouse & Smith 1924):

Praeterea cum materies est multa parata,/ cum locus est praesto nec res nec causa moratur/ ulla, geri debent ni mirum et confieri res./	Besides, when abundant matter is ready, when space is to hand, and no thing and no cause hinders, things must assuredly be done and completed.
nunc et seminibus si tanta est copia, quantam/ enumerare aetas animantum non queat omnis,/	And if there is at this moment both so great store of seeds as all the time of living existence could not suffice to tell,

quis eadem natura manet, quae semina rerum/ conicere in loca quaeque queat simili ratione/ atque huc sunt coniecta, necesse est confiteare/ esse alios aliis terrarum in partibus orbis/ et varias hominum gentis et saecla ferarum.	and if the same power and the same nature abides, able to throw the seeds of things together in any place in the same way as they have been thrown together into this place, then you are bound to confess that there are other worlds in other regions and different races of men and generations of wild beasts.

Despite these doubts, the Aristotelian model – that postulates the uniqueness of Earth and its unparalleled ability to harbor life in the Universe – was allowed to endure, and dominated humanity's view of the world for several centuries until the late Middle Ages. Accordingly, this model was also promoted by the Church, as it was well in line with the uniqueness of Jesus Christ and the clerical, geocentric view of the cosmos. Nevertheless, there remained some inconsistencies between the Christian and Aristotelian philosophy. Albert Magnus (13th century CE) pointed out that if there was an omnipotent God, why should he create just the one Earth and not all possible realizations of it (Grant 1936). In the end, he neglected this flaw in his reasoning, because a plurality of worlds could just not be congruent with the uniqueness of Jesus Christ. As a pious bishop, there was no way this fundamental principle of the unique incarnation – the Christian concept that God became human (flesh),[1] embodied in his son Jesus Christ – could not be valid.

This dogma began to change with the advent of the Copernican revolution that began to emerge in the early 16th century. Based on his astronomical observations, Nicolaus Copernicus (1473–1543 CE) introduced the heliocentric model of our Solar System, which was disputing to the geocentric model of Ptolemy (2nd century CE) that had been widely accepted before. These developments were carefully observed by the Church, which continued propagating Earth as the center of the Universe. Scholars that were opposing this theory were oppressed and silenced. The most prominent victims of this rigorous policy of clear denial of scientific facts were Giordano Bruno (1548–1600 CE) and Galileo di Vincenzo Bonaiuti de' Galilei (1564–1642 CE). Bruno was an Italian monk who was inspired by the recent discoveries of Copernicus. The refutation of Earth as the center of the Solar System let alone the origin of the surrounding Universe motivated Bruno to revive the historic belief in plurality. For the first time this concept was partly supported by scientific facts, which made Bruno postulate (Bruno, De l'Infinito, Universo e Mondi, Dialogo Terzo; translation provided by Ponnamperuma 1964):

Uno dunque è il cielo, il spacio immenso, il seno, il continente universale, l'eterea regione per la quale il tutto discorre e si muove.	Sky, Universe, all-embracing ether, and immeasurable space alive with movement – all these are of one nature.
Ivi innumerabili stelle, astri, globi, soli e terre sensibilmente si veggono [...]. [...]	In space there are countless constellations, suns, and planets [...]. [...]
Sono dunque soli innumerabili, sono terre infinite, che similmente circuiscono quei soli; come veggiamo questi sette circuire questo sole a noi vicino. [...]	There are also numberless earths circling around their suns, no worse and no less inhabited than this globe of ours. [...]

[1] Incarnation is derived from the Latin word *caro*, which means flesh.

> La raggione è, perché noi veggiamo gli soli, che son gli più grandi, anzi grandissimi corpi, ma non veggiamo le terre, le quali, per esserno corpi molto minori, sono invisibili; [...].
>
> We see only the suns because they give light; the planets remain invisible, for they are small and dark; [...].

As this theory was openly disputing the uniqueness of the incarnation, the Church did not endorse this concept which challenged one of the main pillars of Christianity. Bruno was tried and burnt at the stake in 1600, making him one of the earliest martyrs of science (Paterson 1971). His beliefs, however, were not purely scientific but arose from his theologically motivated criticism of the Christian concept of a unique incarnation (Yates 1964). In fact, Bruno's work was a huge setback for scholars that were promoting the Copernican view of the cosmos for scientific reasons. By condemning Bruno and his theories, the Church also had to dispute the heliocentric model and the underlying scientific facts. People including Galileo Galilei, who insisted on promoting this theory, were denounced as heretics and subjected to the clerical inquisition.

Even though Galilei himself did not believe that other planets might be inhabited by forms of life (Drake 1957), his technological developments supported questioning the uniqueness of Earth and its position in the cosmos. The advent of optical telescopes revealed craters and elevations on the lunar surface, satellites of Jupiter, and a plenitude of stars that has not been visible to the naked eye before. Paired with the recent discovery of the Americas in the 15th century, these new astronomical insights fostered an open mindset towards further undiscovered phenomena far beyond the imagination of previous generations (e.g., Fontenelle 1686). The Dutch astronomer Christiaan Huygens (1629–1695 CE) strongly believed in the existence of a multitude of worlds similar to ours (Huygens, Cosmotheoros Huygens & Huygens 1698):[2]

> What a wonderful and amazing Scheme have we here of the magnificent Vastness of the Universe! So many Suns, so many Earths, and every one of them stock'd with so many Herbs, Trees and Animals, and adorn'd with so many Seas and Mountains! And how must our wonder and admiration be encreased when we consider the prodigious distance and multitude of the Stars?
>
> [...] how vast those Orbs must be, and how inconsiderable this Earth, the Theatre upon which all our mighty Designs, all our Navigations, and all our Wars are transacted, is when compared to them. A very fit consideration, and matter of Reflection, for those Kings and Princes who sacrifice the Lives of so many People, only to flatter their Ambition in being Masters of some pitiful corner of this small Spot.

These words reflect the so-called *principle of mediocrity* which postulates that the likelihood of an item, if randomly drawn from a sample that contains several subsets, is proportional to the sizes of the individual subsets. Transferred to the question regarding the uniqueness of a habitable Earth, this principle implies that life should also exist on Earth-like planets throughout the Universe, as it does exist on Earth (e.g., Rauchfuss 2008).

Based on these technical developments and new discoveries, the belief in the plurality of worlds was widely spread by the end of the 18th century. The idea that there could exist unknown and strange civilizations in the Universe was not dis-

[2] As I could not find the original Latin version of Huygens' Cosmotheoros, only the English translation is provided.

regarded as an impossible scenario any longer. Accordingly, this theme began to appear in literature of the 19th century, for instance in the novels of Jules Gabriel Verne (1828 − 1905; e.g., Autour de la Lune) or Herbert George Wells (1866 − 1946; e.g., The War of the Worlds). Driven by these fictional influence, searches for signs of such an extraterrestrial intelligence emerged. These efforts predominately focused on Solar System objects, owing to the technical limitations of this time. Inspired by the previous observations of Giovanni Schiaparelli (1835–1910),[3] Percival Lowell (1855–1916) was convinced that he had detected canals on Mars reminiscent of intelligent inhabitants of the red planet (Lowell 1895, 1906). Even though Evans & Maunder (1903) argued that the supposed canals are most likely an optical illusion, and higher-resolution images collected in 1909 showed that these features were in fact geological structures, the belief of intelligent Martian inhabitants remained.

Up to the present day, no unambiguous evidence for the existence of extraterrestrial life in our Solar System has been found. In addition to this local search for extraterrestrial intelligence in our immediate neighborhood, larger surveys emerged that were hunting for signs of such civilizations beyond the borders of our Solar System. Cocconi & Morrison (1959) suggested to search for such technosignatures at radio wavelengths that might also be used for communication among other intelligent forms of life. This approach was pursued by pioneers such as Frank Donald Drake, who searched for radio signals from nearby stars such as ϵ Eridani or τ Ceti, yet without any success (Drake 1961, 1979). Nevertheless, the search for extraterrestrial intelligence (SETI) emerged from these pioneering efforts and began to gain recognition among the astronomical community (Sagan 1982). This search still continues, driven by humanity's desire to know whether there are Earth analogs that harbor similar forms of intelligent life. Privately funded projects like the Breakthrough Listen Search for Intelligent Life continue to monitor interesting stellar host stars to find evidence for such civilizations in our galactic neighborhood (e.g., Worden et al. 2017). For a more detailed overview of the past, present, and future of SETI, the reader is referred to Tipler (1981), Papagiannis (1985), or Shuch (2011).

1.2 The exoplanet revolution

Despite all these mostly theoretical considerations, planets outside our Solar System – which are the necessary requirement for extraterrestrial life – remained a theoretical construct without any physical evidence until the end of the 20th century. In 1992 Wolszczan & Frail discovered two planetary-mass objects around the millisecond pulsar PSR B1257+12 with the Arecibo radio observatory in Puerto Rico. This discovery came as a big surprise, as planets were not expected to survive the supernova that must have preceded the current post-main sequence state the system was observed in. It was hypothesized that these planets actually formed after the main sequence period of the host, within a debris disk that was composed of the remains of the pulsar's stellar binary companion (Wolszczan 1994). Recent simulations by Fagginger Auer & Portegies Zwart (2021), however, indicate that planets might even stay bound during and after a supernova if the host system is a stellar binary.

[3]The Italian astronomer Giovanni Schiaparelli observed Mars in 1877 and reported the detection of several *canali* on the planet's surface. Whereas the original Italian word means as much as *channels* (which can be geological structures), it was mistranslated as *canals* (e.g., Sheehan 1988). As canals are by definition human-made, this unfortunate translation supported the belief in life on Mars.

Only three years after the discovery of these pulsar planets, Mayor & Queloz (1995) succeeded in detecting a gas giant planet around the solar-type star 51 Peg. This remarkable discovery of the fist planetary-mass object orbiting a main sequence star was awarded the Nobel prize in physics in 2019. Most peculiar though was the short orbit that this planet exhibits. An orbital period of merely 4.2 days indicated a semi-major axis of $\sim$0.05 au, which is twenty times smaller than the Earth-Sun distance in our Solar System and about 13 % of Mercury's semi-major axis.

With these intriguing discoveries it became clear that indeed planets do exist around celestial bodies outside our Solar System; the system architectures and formation scenarios of these extrasolar environments, however, can be vastly different from what we have observed in our immediate neighborhood.

The initial discoveries of planets outside our Solar System by Wolszczan & Frail (1992) and Mayor & Queloz (1995) launched a new era of extraterrestrial planetology, leading to 4375 confirmed exoplanets that are listed in the NASA exoplanet archive[4] as of April 10, 2021. The semi-major axes and masses of all these exoplanets are shown in Figure 1.1. Together with the Solar System planets, a parameter space of almost 6 orders of magnitude is covered by both weight and orbital separations. As such, TRAPPIST-1 b is one of the lowest-mass exoplanets with a mass equivalent to that of the Earth and a semi-major axis of merely 0.01 au (Gillon et al. 2016, 2017; Grimm et al. 2018), whilst GU Psc b is more than 2000 au away from its stellar host and about 11 times as massive as Jupiter (Naud et al. 2014). The colors used in Figure 1.1 are indicative of the different detection methods that were used to discover these exoplanets. The techniques that are commonly used to detect and characterize extrasolar planets can be divided into two broad categories: direct and indirect methods. Whereas indirect methods infer the presence of a planetary companion by monitoring of stellar observables (such as its reflex motion or flux intensity), direct imaging aims for obtaining spatially resolved images of extrasolar planetary systems.

The clustering of planets that are detected with the same technique indicates that each method favors a specific region of the parameter space for the detection of extrasolar planets. Whereas the transit method succeeds in detecting planets at small semi-major axes that are usually smaller than 1 au, direct imaging excels for separations that are significantly larger than 10 au. The underlying causes that are shaping the individual discovery space for each method are briefly discussed in the following Sections 1.2.1 and 1.2.2. A substantial part of the parameter space exhibits no significant number of exoplanet detections. For instance, we do not know any extrasolar planet that is less massive than Jupiter and more than 20 au away from its host star. This void, however, must not be interpreted as the absence of such planets around stars other than our Sun; it is rather the manifestation of the limitations that each of the discovery methods exhibits. Even though there might be a plethora of exoplanets that are residing within these unprobed regimes in Figure 1.1, our current instrumentation is not sensitive enough to reveal this as of yet undiscovered population.

Our knowledge of extrasolar planetary systems is limited by the available data, which is also obvious from a comparison to the Solar System planets in Figure 1.1. Jupiter is the only planet that has extrasolar analogs of similar mass and semi-major axis. Again, this is not a confirmation that our Solar System planets are unique and that similar objects do not exist around other stars; it is rather a strong indication that

[4] https://exoplanetarchive.ipac.caltech.edu/

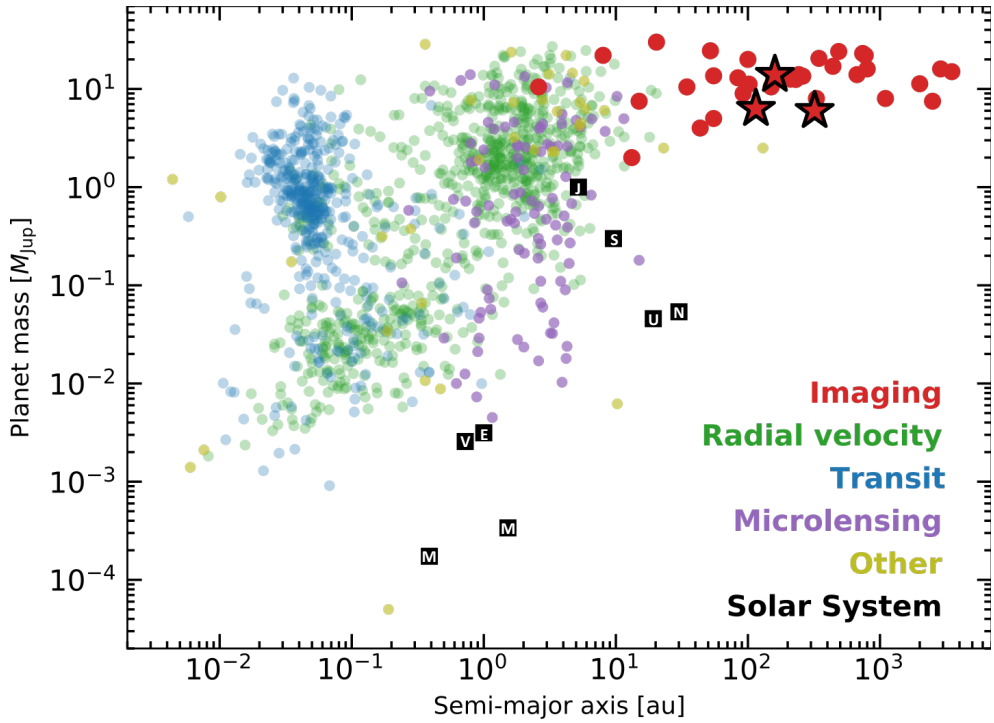

Figure 1.1: Confirmed extrasolar and Solar System planets as of April 10, 2021. The data were collected from the NASA exoplanet archive.[4] We present the planet mass as a function of its semi-major axis. The exoplanets are categorized with respect to their primary detection method as indicated by the colored labels in the bottom right. The red stars highlight the planets that were discovered as part of this thesis. As the orbital parameters for many directly imaged companions are poorly constrained, the semi-major axes are usually approximated by the projected separations; the masses of the directly imaged planets are inferred from theoretical models. For the majority of the planets that are detected by RV measurements, the presented mass is just the minimum mass $M \sin(i)$, as the true mass is degenerate with the orbit inclination.

we are missing a large fraction of extrasolar planets due to limitations of our current technologies. Especially the next generation of large ground and space-based observatories will play a crucial role in exploring this uncharted territory of exoplanet parameter space. The direct detection of Earth-like planets might actually be in reach of first-generation instruments at the Extremely Large Telescope (ELT) of the European Southern Observatory (ESO). Quanz et al. (2015) simulate that the Mid-infrared ELT Imager and Spectrograph (METIS; Brandl et al. 2014) will be sensitive enough to detect ∼10 terrestrial planets with equilibrium temperatures between 200 and 500 K around the nearest stars. METIS is expected to see first light in the late 2020s; accordingly, the first image of an Earth analog outside our Solar System might even be collected within this decade. Future space-based observatories such as the Large Ultraviolet Optical Infrared Surveyor (LUVOIR; The LUVOIR Team 2019) and the Habitable Exoplanet Imaging Mission (HabEx; Mennesson et al. 2016; Gaudi et al. 2020) might accompany these ground-based efforts within the next decades. Even bigger mission such as the Large Interferometer For Exoplanets (LIFE; Quanz et al.

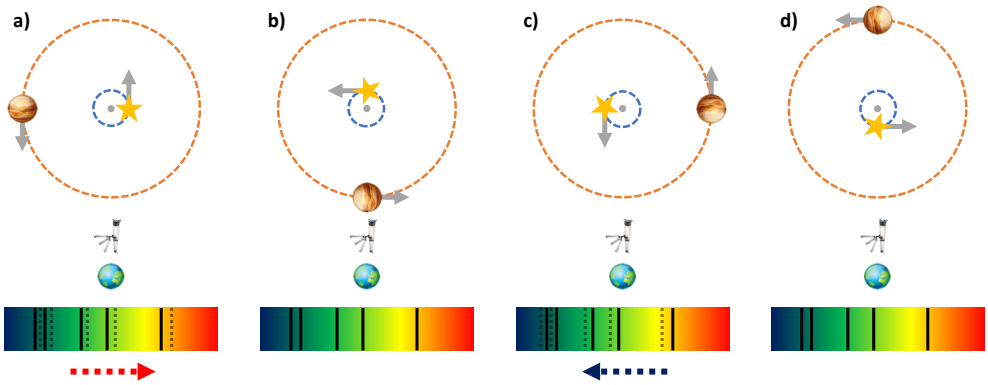

Figure 1.2: The radial velocity method for the detection of exoplanets. In a two-body system both planet and star orbit around the common center of mass (grey dot). Therefore, the presence of a planet induces a periodic change of the stellar velocity vector. The radial component of this velocity change can be measured by monitoring the stellar spectrum. If the star moves away from Earth (panel a) spectral lines are red-shifted and if the star moves towards us the same lines appear blue-shifted (panel c). The solid black lines in the spectral inlays represent the rest-frame spectrum of the star; the dotted lines indicate the redshift (panel a) and blueshift (panel c) of absorption lines due to the orbital motion of the planet. Measuring this periodic change in the stellar radial velocity allows us to infer the orbital period and the minimum mass of the exoplanet.

2019, 2021) are currently designed. LIFE is supposed to characterize the atmospheres of terrestrial exoplanets, and might be able to reveal the first biosignatures outside Earth. Despite the enormous technical challenges that still need to be overcome to conduct such a measurement, humankind might be able to answer the initial question as to whether we are alone in the Universe before the end of the 21st century.

1.2.1 Indirect detections of exoplanets

The majority of known exoplanets have been detected by indirect detection methods (see Figure 1.1): the NASA Exoplanet Archive lists 4324 out of 4375 (∼99 %) to have been discovered indirectly as of April 10, 2021. These indirect techniques infer the presence of an extrasolar planet from the analysis of stellar observables such as its reflex motion (as for instance radial velocity or astrometric measurements) or intensity variations of the primary star that are caused by planetary transit or microlensing events. In this section I will give a brief overview of the main indirect detection techniques, their strengths and limitations. For a more complete review of this topic the reader is referred to Wright & Gaudi (2013).

The radial velocity method (837 discovered planets as of April 10, 2021)

The first exoplanet around a main sequence star was discovered by Mayor & Queloz (1995) who evaluated RV measurements of the G2 star 51 Peg as part of a larger survey for potential extrasolar planets. As visualized in Figure 1.2, a planet that is in orbit around a star introduces a periodic motion of the host, as both star and planet are revolving around their common center of mass. Of course, the reflex motion of the star is several orders of magnitude smaller than the orbital motion of the planet.

Jupiter for instance has an orbital velocity of approximately $13\,\mathrm{km\,s^{-1}}$, but it induces an RV amplitude of only $12.4\,\mathrm{m\,s^{-1}}$ to the motion of the Sun, when viewed from a location within the orbital plane (e.g., Cochran & Hatzes 1996). For Earth in orbit around the Sun the amplitude of velocity change is $0.1\,\mathrm{m\,s^{-1}}$. Nevertheless, it is possible to detect perturbations to stellar velocities in the order of $1\,\mathrm{m\,s^{-1}}$ that can be attributed to orbiting exoplanets (e.g., Pepe et al. 2021). The radial component of this movement can be identified by spectroscopic analysis since absorption lines in the stellar spectrum are blue-shifted when the star moves towards Earth and red-shifted when it moves away from Earth. Observations that cover a full orbit of the exoplanet allow us to identify its orbital period. A minimum mass of the planet can be derived as its true mass is degenerate with the system inclination: a heavy planet on a close-to face-on orbit can cause the same RV signal as a low-mass planet that is orbiting almost edge-on. This degeneracy can be broken by inferring the planet's inclination by other means, for instance if the planet is transiting (e.g., Barbieri et al. 2007), the companion can be imaged (e.g., Maire et al. 2020), precise astrometric data is available (e.g., Benedict et al. 2006), or circumstellar material provides hints towards the most likely system geometry (e.g., Trilling & Brown 1998).

The sensitivity of RV measurements scales as

$$(\mathrm{S/N})_{\mathrm{RV}} \propto M_{\mathrm{p}} P^{-\frac{1}{3}} M_{\star}^{-\frac{2}{3}} \propto M_{\mathrm{p}} a^{-\frac{1}{2}} M_{\star}^{-\frac{1}{2}} \,, \tag{1.1}$$

where $M_{\star}$ is the mass of the primary star and M_{p} denotes the planet mass, P its orbital period and a its semi-major axis (Wright & Gaudi 2013). This method, therefore, favors the detection of massive planets in close orbits around their hosts. This bias is clearly visible in Figure 1.1, as the majority of RV planets exhibit masses above $1\,M_{\mathrm{Jup}}$ and semi-major axes that are smaller than 1 au. Although the signal-to-noise ratio of RV measurements theoretically increases with decreasing host star mass, this does not necessarily mean that M dwarfs are the best targets for large surveys to search for RV planets. Other factors such as the decreasing luminosity and the activity of these low-mass stars have a significant impact on the feasibility to detect planets via this method (e.g., Saar et al. 1998). Yet scientists were able to overcome these challenges in some cases, and detected Earth-sized planets around the two closest M dwarfs to Earth: Ribas et al. (2018) discovered a planet with a minimum mass of $3.2\,M_{\oplus}$ around Barnard's star; and Proxima Centauri, the closest star to Earth at a distance of 1.3 pc, was found to harbor two super-Earths, of which the inner one might be able to support liquid water on its surface (Anglada-Escudé et al. 2016; Damasso et al. 2020).

Transiting exoplanets (3325 discovered planets as of April 10, 2021)

The first planet that was detected by the transit method was HD 209458 b (Henry et al. 2000; Charbonneau et al. 2000), yet the existence of the planet was already confirmed with RV measurements by the time the transit observations were carried out. Indeed, OGLE-TR-56 b is the first extrasolar planet whose discovery can be attributed to the transit technique (Konacki et al. 2003). This method exploits the fact that some exoplanets must move across the line of sight between Earth and their host star during their orbit. As visualized in Figure 1.3 this obscuration results in a periodic dimming of the stellar brightness, similar to a solar eclipse observed here on Earth, yet of much smaller magnitude. The amplitude of the dimming event is

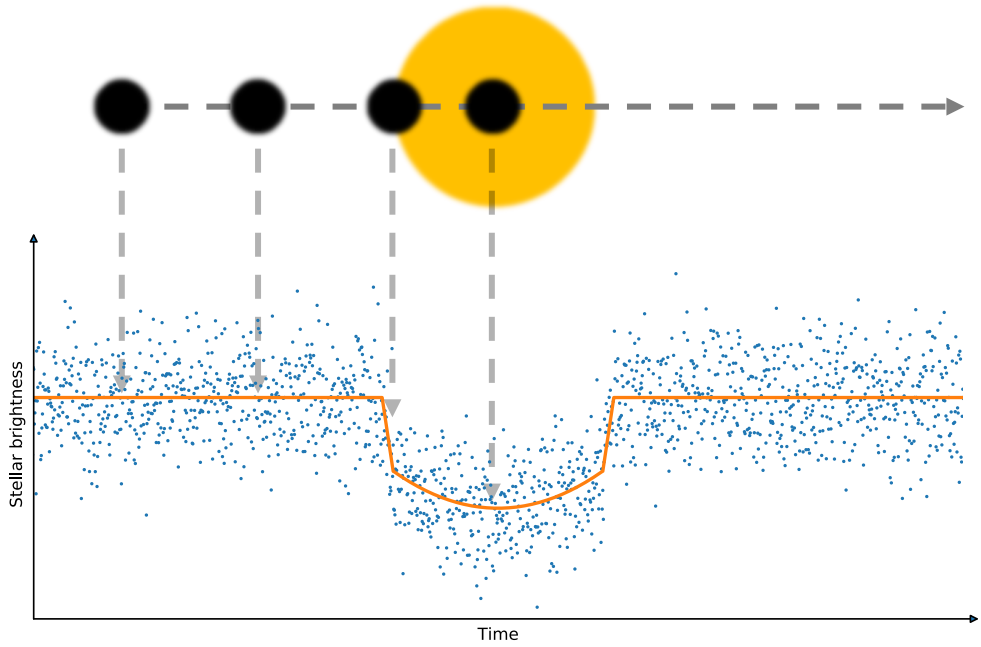

Figure 1.3: The transit method for the detection of exoplanets. Planets can move across the line of sight between observer and primary star. This results in a periodic dimming event that can be measured via continuous monitoring of the stellar luminosity.

directly related to the planetary and stellar radii involved. Based on purely geometrical considerations the stellar fractional intensity reduction, the so-called transit depth, is modeled as

$$\frac{\Delta I}{I} = \frac{R_\mathrm{p}^2}{R_\star^2}\,, \tag{1.2}$$

where R_p and $R_\star$ denote the radii of the planet and the star, respectively. If the stellar radius is well known, the planetary radius can be inferred from a transiting event. For solar-type stars the mangitudes of these eclipsing events are in the order of 1 % for Jovian companions and 0.01 % for terrestrial planets. Due to the smaller stellar radii, M dwarfs are particularly favorable to search for Earth-sized planets with the transit method. The transit depth is increased compared to solar-type primaries and therefore easier to detect. Instruments like TRAPPIST (Gillon et al. 2011; Jehin et al. 2011) or SPECULOOS (Delrez et al. 2018) exploit this advantage in the search for transiting, terrestrial exoplanets. One of the most intriguing results from these surveys was the discovery of seven temperate Earth-sized planets around the M dwarf TRAPPIST-1 (Gillon et al. 2017) that launched a cascade of follow-up observations. An even larger transit depth of ∼57 % is exhibited by the planetary system around the white dwarf WD 1856+534, which is hosting a Jupiter-sized companion that is eclipsing its post-main sequence host in a grazing transit (Vanderburg et al. 2020).

A single stellar dimming event, however, is not a solid indicator for a transiting planet. To confirm a companion, the transit has to occur periodically, which allows us to determine the planet's orbital period. If the stellar mass is known, the orbital

semi-major axis can be derived from the period. But even periodic stellar dimming events might originate from other sources than transiting exoplanets (e.g. Brown 2003; Charbonneau et al. 2004). There is a variety of non-planetary sources that could mimic transit-like events, such as stellar variability and other unfortunate stellar constellations. For instance, grazing eclipses of stellar binaries can look like planetary transit signatures. Another frequent source of false-positive detections are eclipsing stellar binaries that are close to the scientific target on sky, yet not resolved by the recording telescope. When extracting the stellar flux, the contributions of all stars in the aperture are combined, which can mimic the signal of a transiting exoplanet. To test for the two latter scenarios, high-contrast imaging at high spatial resolution as introduced in Section 1.2.2 is a powerful technique. The detection of astronomical objects that are hidden below the angular resolution limit of transit survey telescopes is important to correct the derived planetary parameters for the flux contribution of the identified contaminant (e.g. Evans et al. 2016b; Southworth et al. 2020). This can lead to revised planetary radii that differ by up to 10% from the previously derived values, unaware of the close contaminant (for the exoplanet WASP-20 b; Evans et al. 2016b). Such follow-up observations of transiting exoplanet host stars at higher spatial resolution can also reveal gravitationally bound stellar companions (e.g. Daemgen et al. 2009; Faedi et al. 2013a; Bergfors et al. 2013; Adams et al. 2013; Ngo et al. 2015). Studying the multiplicity of exoplanet host systems is important to understand the formation and evolution of planetary environments. This topic will be discussed extensively in Chapter 2 of this thesis.

The majority of exoplanets has been discovered by space-based transit surveys, which fundamentally contributed to our knowledge of planetary occurrence rates and the diversity of planetary system architectures. The Kepler space mission has contributed more than 2'500 objects to our census of extrasolar planets (Borucki et al. 2010), and the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) is currently expanding the sample of short-period planets around bright and nearby stars. But also ground-based surveys successfully detected planets via the transit method: especially the Wide Angle Search for Planets (WASP; Pollacco et al. 2006) discovered more than 150 transiting exoplanets.

Despite this plethora of discoveries, there are a few shortcomings associated with exoplanet detections via the transit method. First, there is no intrinsic possibility to derive a planetary mass from the observed dimming event. These parameters have to be obtained via follow-up measurements using other techniques such as the RV method. Second, the probability of a planet to eclipsing its host star, p_{transit}, decreases with the semi-major axis a of the planet as per

$$p_{\mathrm{transit}} \propto \frac{R_\star}{a} \,, \tag{1.3}$$

again purely based on geometric considerations (Borucki & Summers 1984). Wright & Gaudi (2013) derive a sensitivity of

$$(\mathrm{S/N})_{\mathrm{transit}} \propto p_{\mathrm{transit}}^{-\frac{1}{3}} R_\mathrm{p}^2 M_\star^{-\frac{5}{3}} \propto a^{-\frac{1}{2}} R_\mathrm{p}^2 M_\star^{-\frac{3}{2}} \tag{1.4}$$

to detect transiting planets, when assuming that the stellar radius scales linearly with the stellar mass. This assumption is a decent approximation for stars with masses up to $1\,M_\odot$. Accordingly, the population of transiting exoplanets is heavily biased towards short-period giant planets with typical planet masses between $10\,M_\oplus$ and $10\,M_\mathrm{Jup}$ at separations of 10^{-2} au to 1 au (see Figure 1.1).

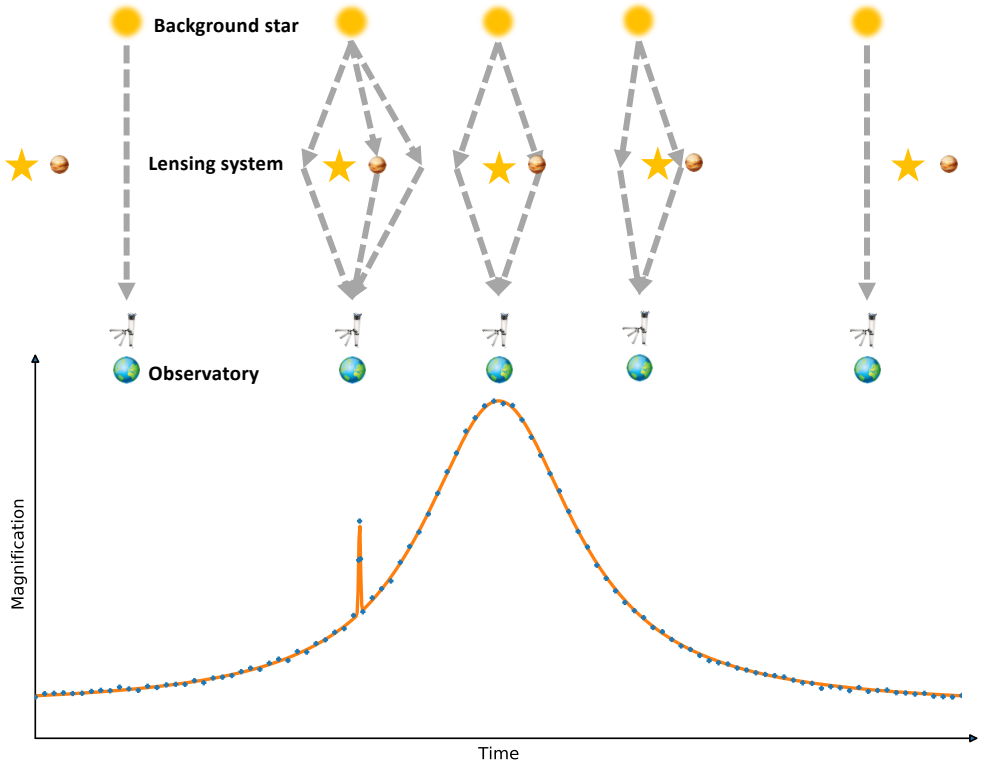

Figure 1.4: Microlensing events for the detection of exoplanets. According to Einstein's general theory of relativity light curves are bent in the presence of matter. Hence, a foreground star that is passing across the line of sight from Earth, can artificially magnify the recorded intensity of a more distant background object, similar to a lens in optical physics. A planet in the lensed system can be identified as perturbation of the magnification lightcurve that is resembling the gravitational lensing of a single star.

Microlensing events (108 discovered planets as of April 10, 2021)

The microlensing method to detect extrasolar planets makes use of Einstein's general theory of relativity (Einstein 1916). This framework postulates that matter curves space-time, and hence massive objects can bend the trajectories of light. Dyson et al. (1920) were able to experimentally confirm this theoretical concept for the first time, by measuring the angular displacement of stars due to the gravitational influence of the Sun during a solar eclipse in 1919. As visualized in Figure 1.4 this concept is also applicable to other stellar constellations that can be observed here on Earth. If a foreground star is moving across the line of sight, the light emitted by a background star is deflected due to general relativity. Accordingly, this foreground star can act as a lens that is magnifying the background source's intensity measured on a detector. If the lensing system harbors a planetary companion, the gravitational influence of this planet can create an additional magnification on top of the intensity variations caused by its stellar host. This second-order magnification is of much smaller magnitude and acts on much shorter time-scales than the primary magnification. Yet with sufficient

temporal sampling of the lightcurve of the background star, this variation can be identified and attributed to a planet.

This method was successfully employed by Bond et al. (2004) for the first time to detect OGLE 2003-BLG-235L b (MOA 2003-BLG-53L b). The observed microlensing event was attributed to a binary system with a mass ratio of $q = 0.0039$ between the primary star and its companion. As this observational technique just provides this mass ratio and not the individual masses of both objects, follow-up measurements were necessary to break this degeneracy. Bennett et al. (2006) identified the lensing star in Hubble Space Telescope data, determined its spectral type, and concluded that OGLE 2003-BLG-235L b is likely a gas giant planet with a mass of $2.6\,M_{\mathrm{Jup}}$ with an orbital semi-major axis of 4.3 au. Indeed, microlensing detections of exoplanets are not possible for arbitrary system geometries. The method heavily favors the detection of planets that are close to the Einstein radius (Einstein 1936) of the lensing foreground star. The optimal separation of a planet to be detected via these means is

$$a_{\mathrm{opt}} \approx 2.85\,\mathrm{au}\sqrt{\frac{2M_\star}{M_\odot}} \tag{1.5}$$

(see Wright & Gaudi 2013). Due to this intrinsic bias, the majority of microlensing planets have been detected in a range of semi-major axes from approximately 1 au to 10 au (see Figure 1.1). Future missions such as the Nancy Grace Roman Space Telescope (Spergel et al. 2015) are supposed to discover more than one thousand bound exoplanets by microlensing events (Penny et al. 2019). Furthermore, this mission might reveal a significant population of unbound planetary companions that are freely floating through our Galaxy (Barclay et al. 2017). Such isolated planets were observed in the past (Sumi et al. 2011); they are are postulated to be ejected from their natal circum-stellar systems (e.g., Rasio & Ford 1996).

Other indirect detection methods (54 discovered planets as of April 10, 2021)

There are several other indirect methods that can detect extrasolar planets. The number of planets detected via these means, however, is small compared to the three major indirect methods discussed above. Future instrumentation might enable a plethora of new detections – a planet population that is just out of reach of current observatories.

Astrometry The reflex motion of a planetary host star that is induced by an orbiting companion does not only have a radial component as probed by RV measurements. The star also moves perpendicular to the line of sight in the plane of the sky. This motion can be measured and associated with a planetary-mass companion. The combination of RV observations with astrometric measurements breaks the mass degeneracy by revealing the orbital inclination (Tuomi et al. 2009). However, extremely high astrometric precision is required to reveal the presence of an exoplanet: whereas a Jupiter analog orbiting a Sun-like star creates an astrometric signal in the order of $\sim$0.1 mas, the astrometric amplitude of an Earth analog is about 0.1 µas (Wright & Gaudi 2013). Hence only a handful of planets were claimed to be discovered by this technique (e.g., Muterspaugh et al. 2010; Sahlmann et al. 2013), and most of the results need to be confirmed by longer observational baselines or complementary

detection methods. The *Gaia* space observatory of the European Space Agency (Gaia Collaboration et al. 2016) is the most promising mission to significantly increase the small number of astrometrically confirmed exoplanets. *Gaia* is monitoring the motions of more than 1 billion stars with unprecedented precision during a nominal mission duration of 5 years. Perryman et al. (2014) simulated that these data should reveal about 21'000 new planets via the astrometric wobbles of their respective host stars.

Transit timing When a system with a single planet is aligned such that this companion is transiting, this transit occurs strictly periodically. If further planets are present, their gravitational influence will affect the orbit of the initial planet. Even if the additional planets do not transit themselves, their presence manifests itself in a deviation from the strict periodicity of the initial transit signal. Using this technique, non-transiting planets can thus be detected in known transiting planet-hosting systems (e.g., Nesvorný et al. 2012).

Pulsar timing Even though the first exoplanets were discovered by this method (Wolszczan & Frail 1992), the absolute number of pulsar planets is small: Only seven planets are associated with this category according to the NASA exoplanet archive. The method relies on modulations of the pulsar period by an orbiting plant. Due to the induced motion the frequency of the pulsar is altered, which allows us to derive the mass and the orbital period of the planetary companion.

Pulsation timing Similar to the modulation of the pulsar frequency, orbiting planets can affect the arrival time of lightcurve modulations due to stellar variability. Silvotti et al. (2007) used this technique to reveal a Jovian gas giant orbiting the extreme horizontal branch star V391 Pegasi. The second planet discovered by the pulsation timing method was found around KIC 7917485 in the Kepler field of view (Murphy et al. 2016).

Disk kinematics One of the most recently developed techniques to indirectly detect exoplanets is the analysis of kinematic structures in protoplanetary disks. The advent of the Atacama Large Millimeter/submillimeter Array (ALMA) facilitated studies at radio wavelengths with unprecedented spectral and spatial resolution (e.g., Ansdell et al. 2016; Andrews et al. 2018). Forming planets will impact the gas dynamics in the protoplanetary disk, and therefore these signatures can be searched in the spectroscopic data. Teague et al. (2018) and Pinte et al. (2018) exploited these novel opportunities and announced the detection of several protoplanets around the young star HD 163296; yet other methods could not confirm these proposed protoplanets so far (e.g., Mesa et al. 2019a).

1.2.2 Direct detections of exoplanets

Only $\sim$1 % of all known exoplanets have been discovered by taking a direct image of the companion (51 out of 4375 as of April 10, 2021). This small number is due to the challenging nature of the problem to image a faint planet right next to its host star that is many orders of magnitude brighter. Hence, the nature of the underlying problem for our imaging system is twofold: not only the huge contrast between star

and planet has to be overcome, but this sensitivity has to be provided at very small angular separations from the star.

Huge contrast and small angular separation: the main challenges of direct imaging

The typical angular separations when trying to image Solar System analogs can be assessed geometrically and scaled with the distance of the targeted system. Whereas a Jupiter analog to a star at 5 pc is separated 1″ from its primary, the same system geometry at 100 pc corresponds to an angular separation of 0.″05. An Earth twin would have an angular separation of 0.″2 and 0.″01 for a system at 5 pc and 100 pc, respectively. For comparison, a 10 m-class telescope (which is the largest category of currently operating optical telescopes) has a diffraction limit of about 0.″015 at optical wavelengths (600 nm) and 0.″05 in the near infrared (2 µm). But this diffraction limit based on the Rayleigh criterion (see equation 1.11) is not equivalent to the true spatial resolving power of a ground-based observatory, as discussed later. From these geometrical considerations, it seems plausible that the direct imaging search for exoplanets should focus on the stars that are closest to the Sun. These provide more favorable angular separations for Solar System equivalents, and planets should be easier to detect.

The second variable, however, that needs to be considered in this problem is the contrast between the star and the planet. To quantify the magnitude of this challenge, one has to investigate the different types of electromagnetic radiation that are emitted by a planetary-mass companion. Planets cool down after their formation, as their masses and hence the internal temperatures are too low to fuse hydrogen or deuterium in the core (e.g., Burrows et al. 1997). The major sources of electromagnetic radiation that can be detected from such a planetary-mass object are

1. reflected light of the host star
2. and thermal radiation
   a) from reprocessed starlight
   b) or due to release of gravitational energy.

To first order, sources 1 and 2a do not depend on the age of the system, as they just rely on the stellar radiation that is either reflected or absorbed by the planet. The spectral energy distribution (SED) of reflected light from an exoplanet is shaped by the emission spectrum of its primary star, altered by atmospheric molecular absorption and scattering processes (e.g., Selsis et al. 2008). The reflected light contrast between star and planet $c_{\mathrm{ref}}(\lambda)$ can be approximated by

$$c_{\mathrm{ref}}(\lambda) = A_{\mathrm{g}}(\lambda) \left(\frac{R_{\mathrm{p}}}{a} \right)^2 f_{\mathrm{ref}}(\lambda, \alpha) \,, \tag{1.6}$$

where $A_{\mathrm{g}}(\lambda)$ and $f_{\mathrm{ref}}(\lambda, \alpha)$ denote the wavelength-dependent geometric albedo and phase function of the planet (e.g., Seager 2010). The latter quantity also depends on the phase angle α of the planet, which determines the amount of flux reflected off the planet towards the direction of Earth. For a fixed planet radius, geometric albedo, and phase angle, the reflected light contrast at a certain wavelength becomes more favorable the higher the stellar irradiation, i.e. the smaller the planet's semi-major axis.

Of course, this effect is exactly the opposite in our imaging system, whose contrast degrades the closer a source is to the primary star (when assuming a regular Airy-shaped point spread function for both sources; see, e.g., Figure 1.6). For an Earth analog with a geometric albedo of 0.4 and a phase function of 0.5, equation (1.6) provides a reflected light contrast in the order of 4×10^{-10} at visible wavelengths. The detection of Jupiter analogs is not much more favorable in reflected light, as the larger semi-major axes reduce the amount of stellar irradiation, providing contrasts in the order of 2×10^{-9}. These requirements are significantly lower than the sensitivities of current instrumentation. Ground-based optical studies can achieve a total intensity contrast of approximately 1×10^{-6}, which is more than three orders of magnitude higher than the required specifications (e.g., Hunziker et al. 2020). Future space-based observatories such as the Nancy Grace Roman Space Telescope might be able to achieve these challenging requirements (e.g., Robinson et al. 2016; Trauger et al. 2016; Zimmerman et al. 2016; Girard et al. 2020).

As the reflected light spectrum of an exoplanet roughly follows the SED of the stellar primary, leading to the extremely challenging contrasts $10^{-9} - 10^{-10}$, the thermal emission of a planet is usually better suited for its direct detection. The SED maximum of solar or earlier type stars is located at visible wavelengths, whereas the peak emission of much cooler planets is in the near to mid infrared wavelength regime. In this region, the stellar luminosity decreases as a function of λ^{-4} according to the Rayleigh–Jeans limit of its approximated blackbody emission. This leads to more favorable star-to-planet contrasts at near to mid infrared wavelengths. When approximating both star and planet as a blackbody $B_\lambda(T)$, the thermal planet contrast is

$$c_{\text{therm}}(\lambda) = \left(\frac{R_\text{p}}{R_\star}\right)^2 \frac{B_\lambda(T_\text{p})}{B_\lambda(T_\star)}\,, \tag{1.7}$$

with T_p and $T_\star$ denoting the planetary and stellar temperatures, respectively (Wright & Gaudi 2013). Assuming that we observe both blackbodies in the Rayleigh–Jeans limit, equation (1.7) simplifies to

$$c_{\text{therm}}(\lambda) = \left(\frac{R_\text{p}}{R_\star}\right)^2 \frac{T_\text{p}}{T_\star}\,, \tag{1.8}$$

which can be written as

$$c_{\text{therm}}(\lambda) = \left(\frac{R_\text{p}}{R_\star}\right)^2 \frac{T_\text{eq}}{T_\star}\,, \tag{1.9}$$

when requiring the planet to be in thermal equilibrium ($T_\text{p} = T_\text{eq}$). T_eq denotes the theoretical equilibrium temperature of a planet, which is derived under the assumption that the energy from the absorbed stellar irradiation is completely re-emitted as thermal blackbody radiation of the planet: in this framework it is assumed that the planet is isothermal (i.e. there are no temperature gradients across its surface) and that it has no intrinsic source of energy. Utilizing our example of an Earth and Jupiter analog with equilibrium temperatures of 250 K (Earth) and 110 K (Jupiter) and an effective solar temperature of 5'800 K, provides contrast ratios of 4×10^{-6} and 2×10^{-4}, respectively. This is an improvement of 4–5 orders of magnitude compared to the previously discussed reflected light contrasts.

However, no such Solar System planet analog with a similar age and equilibrium temperature could be directly observed so far. The reason for this shortcoming is simple: for equation (1.9) to hold, we assumed that λ is in the Rayleigh–Jeans limit of the Planck function, which requires that

$$\lambda \gg \frac{hc}{k_\mathrm{B} T} \,. \tag{1.10}$$

The right-hand side of this inequality is in the order of 50 µm for a terrestrial equilibrium temperature of 250 K. Observations at this wavelength range are impossible from the ground as our own atmosphere is not transparent at these frequencies. Another issue that comes along with observations in the infrared is closely related to the reason that favors the detection of terrestrial planets in this frequency regime. As everything on Earth has a temperature similar to this equilibrium temperature of 250 K, the amount of noise originating from the atmosphere, the telescope, and other surfaces is exponentially increasing for wavelengths that are longer than 1 µm. Observations at 10 µm are feasible, but the background noise is usually larger than the signal from any astrophysical source (e.g., Absil et al. 2004). When going to shorter wavelengths, condition (1.10) is violated and the star-to-planet contrast degrades drastically, due to the exponential dependence of $B_\lambda(T)$ in the opposing Wien regime. Evaluation of the initial equation (1.7) provides a thermal contrast of 5×10^{-10} when observing an Earth analog at 5 µm and even 7×10^{-17} at 2 µm. So, space-based observatories are strictly mandatory for the direct detection of old, evolved exoplanets with equilibrium temperatures below 300 K. But even then, these observatories must be actively cooled to minimize the thermal noise of the observations. Besides, the ability to resolve star and planet scales inversely proportional with wavelength leading to enormous aperture sizes that are required to provide the additionally required spatial resolution. A solution might be a space-based nulling interferometer such as the proposed LIFE mission (Quanz et al. 2021).

The only class of exoplanets that could be directly probed with current instrumentation are those of category 2b that emit additional energy released by the gravitational contraction following their recent formation. In other words, these are young exoplanets that exhibit higher temperatures and that are therefore more luminous than their older counterparts. Young Jovian giants with ages that are smaller than 20 Myr and masses in the range of $1\,M_\mathrm{Jup} - 13\,M_\mathrm{Jup}$ usually have temperatures of 1'000 K – 2'000 K. For solar-type host stars with $T_\star = 5'800\,\mathrm{K}$ and $R_\star = 1\,R_\odot$ and Jupiter-sized planets ($R_\mathrm{p} = 1\,R_\mathrm{Jup}$), this corresponds to contrasts in the order of $2 \times 10^{-5} - 7 \times 10^{-4}$ and $4 \times 10^{-4} - 2 \times 10^{-3}$ at 2 µm and 5 µm, respectively (see equation 1.7). This regime is well probable with current telescopes and instruments as discussed below.

In Figure 1.5 we visualize the thermal contrast of gas giant companions as a function of age for several wavelengths. We used AMES-Cond evolutionary models (Allard et al. 2001; Chabrier et al. 2000) to obtain the effective temperatures for planets of various masses at ages in the range 1 Myr – 10 Gyr. We simulated a solar-like primary whose effective temperature as a function of age was determined in a similar fashion using MIST models (Dotter 2016; Choi et al. 2016). For host star and planet we applied radii of $1\,R_\odot$ and $1\,R_\mathrm{Jup}$, respectively, which we assumed to be constant in time. It is clearly visible that the star to planet contrast, c_therm, for a fixed object mass decreases as a function of time, mainly driven by the cooling of

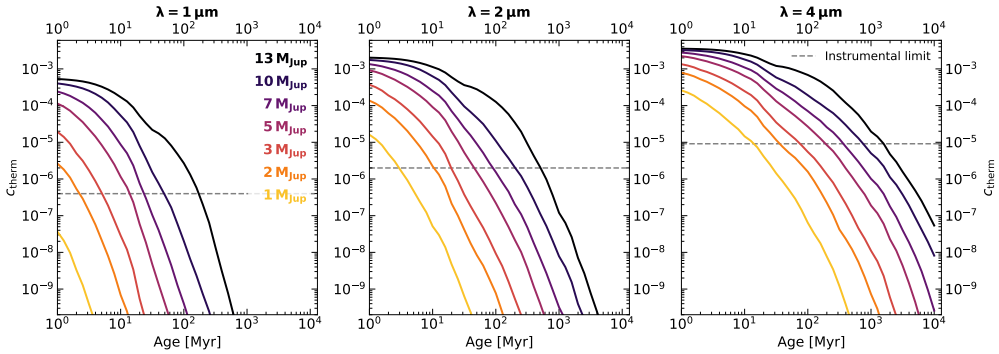

Figure 1.5: Thermal star-to-planet contrast as a function of system age. The contrasts are derived using equation (1.7). Planetary and stellar temperatures are derived from AMES-Cond and MIST evolutionary models, respectively. A range of planet masses are assessed as indicated in the top right of the left panel. We further assume a constant planetary and stellar radius of 1 R_{Jup} and 1 $R_{\odot}$, respectively. The contrast is evaluated for infrared wavelengths of 1 µm (left panel), 2 µm (middle panel), and 4 µm (right panel). The dashed lines represent instrumental contrast limits for SPHERE (1 µm and 2 µm; simulated with the SPHERE exposure time calculator[6]) and NACO (4 µm; Launhardt et al. 2020).

the planetary object. Independent of age, the contrast is more favorable for higher-mass objects, which also agrees with the higher temperatures that are associated with heavier objects.[5] When evaluating c_{therm} for different wavelengths, it becomes clear that the planet contrast becomes more favorable at longer wavelengths. As noted before, the thermal background noise on our detectors increases at longer wavelengths as well, deteriorating the achievable contrast. To obtain reasonable estimates for actual instrumental sensitivities at the wavelengths of this comparison (1 µm, 2 µm, and 4 µm), we relied on the performance of advanced high-contrast imaging instruments for each corresponding wavelength regime: for 1 µm and 2 µm this was the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE; Beuzit et al. 2019) and for 4 µm the NAOS adaptive optics system combined with the CONICA camera (NACO; Lenzen et al. 2003; Rousset et al. 2003). The NACO contrast at 4 µm was adopted from data on ϵ Eridani ($m_{4\mu\mathrm{m}} \approx 1.6$ mag) published by Launhardt et al. (2020). For SPHERE we utilized the dedicated exposure time calculator[6] and simulated the contrast for 1 h on-target integration on a bright, solar-type star ($m = 0$ mag). For both instruments, we assessed the contrast performance at 1″, which is visualized by the dashed lines in Figure 1.5. As expected, the instrumental contrast performance decreases when going to longer wavelengths. A wavelength of 4 µm seems to be most favorable for the detection of young planets: for system ages that are smaller than 10 Myr planets with masses down to 1 M_{Jup} can be detected, and for older systems planets with masses close to the deuterium burning limit of $\sim$13 M_{Jup} seem to be within reach of NACO. But this must not be interpreted as a general conclusion. First, the angular resolution is proportional to wavelength, which means that at shorter wavelengths smaller angular scales can be probed. Second, these simulations rely on synthetic evolutionary models for sub-stellar objects

[5] Note that the energy budget to calculate the planetary luminosity does not include any reprocessed stellar light in this scenario.

[6] Online available at `https://www.eso.org/observing/etc/`

which have their own uncertainties. And third, observations at shorter wavelengths are usually more efficient as longer integration times can be applied and less time is spent off-target to monitor and model the thermal sky background. It is true though that wavelengths from H to L band are best suited for the detection of gas giant exoplanets as all undisputed planet discoveries by direct imaging were exclusively made in this wavelength regime (e.g., Marois et al. 2008, 2010; Lagrange et al. 2010; Rameau et al. 2013; Macintosh et al. 2015; Chauvin et al. 2017b; Keppler et al. 2018). Besides, younger planets are brighter and therefore easier to detect. Due to this relation, the majority of directly imaged planets exhibit ages that are younger than 100 Myr. Particularly favorable are young stellar associations with ages of even less than 20 Myr. As these are mostly farther away than 100 pc from Earth, the physical scales that can be probed with current instrumentation are usually larger than the semi-major axis of Jupiter in the Solar System: for instance, an inner working angle of $0\farcs1$ as minimum angular separation, at which state-of-the-art instruments can detect planetary-mass companions, would correspond to detectable planet semi-major axes that are larger than 10 au for systems that are farther away than 100 pc.

From these considerations it is clear that the development of observing strategies for the direct detection of planetary companions is a high-dimensional problem that cannot easily be optimized. There is certainly no perfect strategy to detect such companions, and several trade-offs have to be considered such as

- observing at longer wavelengths (as planets are brighter in this regime) versus the increasing background noise and the lower angular resolution of the imaging system (going hand in hand with a reduced sensitivity, especially at small angular separations);
- or targeting comparably older systems close to Earth (to probe physicals scales of a few Astronomical Units) versus more distant, young associations (for which the planet contrast is more favorable due to the younger ages).

The past decade has shown us that current instrumentation is capable of imaging young, self-luminous gas giant planets, yet significant technical advances were required to obtain these achievements.

Techniques required for the direct detection of exoplanets

The main deliverable required of a system for the imaging of extra solar planets is the high-contrast that has to be provided at small angular separations. The fundamental resolution limit of an imaging device with an unobstructed, circular aperture is determined by the Rayleigh criterion, which defines the minimum angular separation ϑ_{min} that can be resolved with an aperture size D when observing at wavelength λ as

$$\vartheta_{\mathrm{min}} \approx 1.22 \frac{\lambda}{D} \,. \tag{1.11}$$

This leads to the conclusion that for a fixed wavelength a large aperture size is required to probe tiny angular scales in the order of a few tens of milliarcsecond (10 mas $\widehat{=}$ 1 au at 100 pc). While this assumption is in theory correct, there are other effects that need to be considered when performing observations from the ground.

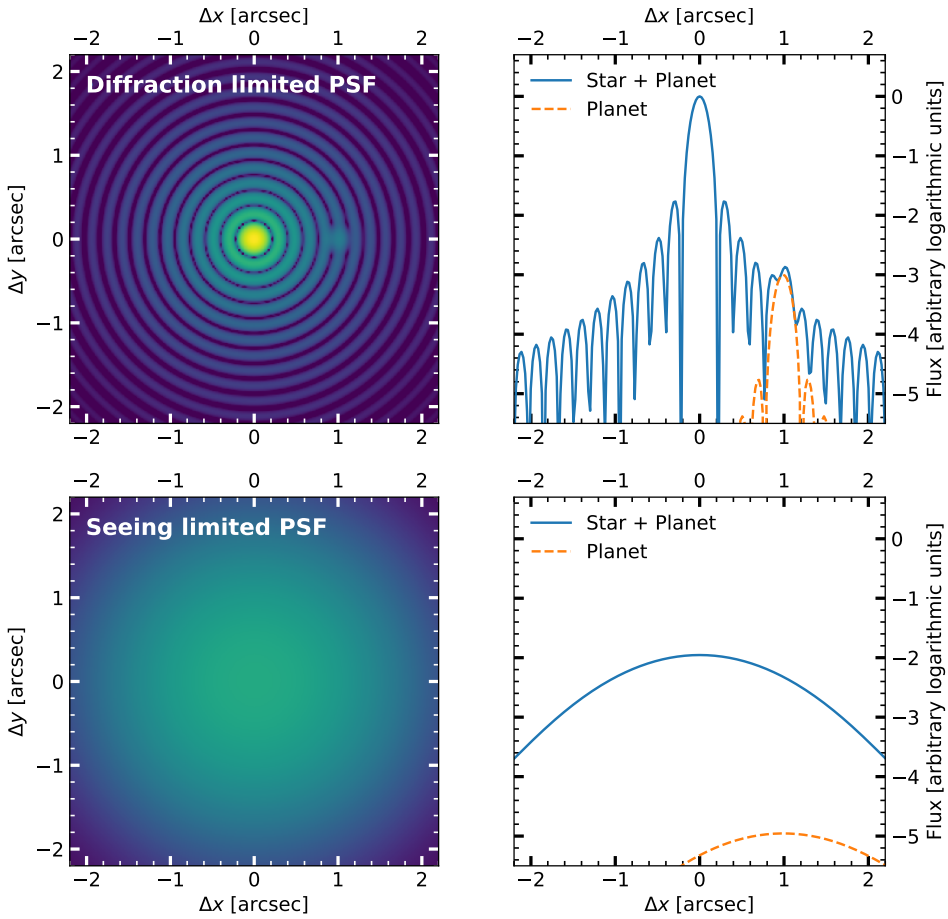

Figure 1.6: The effect of astronomical seeing on the angular resolution of an imaging system. *Top panel*: Diffraction limited PSF of a stellar point source and a planetary companion at an angular separation of $1''$ with a contrast of 10^{-3}. The left panel shows an image of both sources, and the right panel represents the intensity profile along the x direction for $y = 0''$. *Bottom panel*: Same as top panel but with a simulated seeing of $0''.8$.

Adaptive optics systems Unfortunately, the true resolution of an imaging system is usually not equal to the Rayleigh criterion in equation (1.11), which must rather be interpreted as a lower limit of physically possible resolutions. This deviation is caused by wavefront aberrations induced to the collected light before being focused onto the detector. Especially for ground-based observatories the inhomogeneous and turbulent temperature profile of Earth's atmosphere is severely perturbing the plane wavefronts that are approaching from a distant point source such as a star or an exoplanet. This effect commonly known as astronomical seeing convolves the theoretical Airy-like point spread function (PSF) with a kernel that can be approximated by a Gaussian function to first order. The full width at half maximum (FWHM) of this Gaussian function depends on the strength of the atmospheric turbulence, which can be parameterized by the Fried parameter r_0 (Fried 1966). Instead of a diffraction limited PSF that is scaling as λ/D, the collected image is seeing limited with

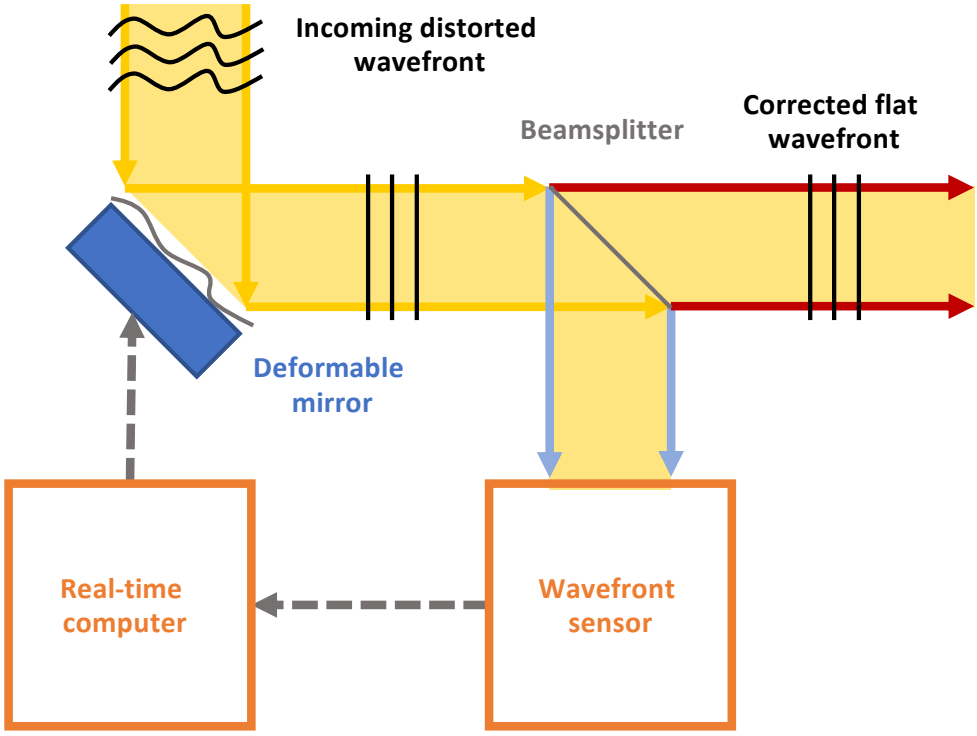

Figure 1.7: Schematic setup of an AO system. The distorted incoming wavefront is corrected by a deformable mirror. This mirror is controlled by a real-time computer that is analyzing the wavefront distortions that are measured with a wavefront sensor. This sensor is usually fed with the blue part of the incoming spectrum that is separated by a beam splitter. The aberrations of the red part of the light are corrected and focused on a detector.

$\theta \propto \mathrm{FWHM} \propto \lambda/r_o$. Accordingly, the achieved resolution does not depend on the telescope aperture size any longer, but is purely determined by the strength of the seeing. FWHMs smaller than $0\farcs8$ are usually considered as good seeing conditions; at prime sites of ground-based astronomical research these values can even go as low as $0\farcs4$ in some exceptional nights. The impact of this effect on the imaging of exoplanets is visualized in Figure 1.6. Whilst an off-axis companion at $(\Delta x, \Delta y) = (1'', 0'')$ with a contrast of 10^{-3} is visible in the diffraction-limited image (top panel of Figure 1.6), it is clearly hidden below the seeing halo of the primary star in the bottom panel of Figure 1.6.

To image faint exoplanets at close angular separations, it is therefore paramount to have an imaging system that is working close to the diffraction limit of the corresponding telescope. The effect of the seeing can be reverted by so-called adaptive optics (AO) systems (Babcock 1953; Hardy 1998). A schematic overview of such a system is presented in Figure 1.7. The blue part of the distorted, incoming light is analyzed by a wavefront sensor that sends information about the residual aberrations to a real-time computer. This computer calculates and alters the shape of a deformable mirror to correct for the remaining aberrations. This correction process has to be applied in real time to correct for quickly varying aberrations. Whereas the

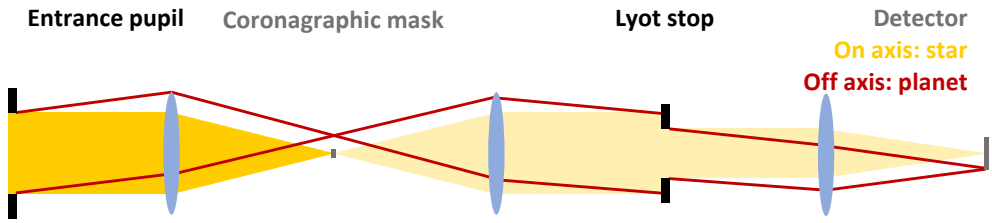

Figure 1.8: Optical design of a basic Lyot coronagraph. The light propagating through the entrance pupil is focused, and a coronagraphic mask is occulting the central part of the image (i.e. where a bright source is located). When the beam is reimaged afterwards, the majority of the remaining light from the on-axis source is distributed at the edges of the pupil image. This beam passes a Lyot stop that is blocking this residual light from the central star and transmitting photons from off-axis sources (red lines). When imaged on the detector, the bulk of stellar light is removed and faint, off-axis point sources become visible.

blue part of the corrected light is directed towards the wavefront sensor, the red part propagates as an almost flat wavefront towards the science detector.

Usually, the Strehl ratio is used as a metric to measure the performance of an AO system. Strehl is defined as the ratio between the peak intensity of a recorded PSF to the peak intensity of the theoretical diffraction-limited PSF. Extreme AO systems such as the SPHERE AO for eXoplanet Observation (SAXO; Fusco et al. 2006) can achieve on-sky Strehl ratios of more than 90 % in the H band.

Coronagraphy The use of (extreme) AO is usually not sufficient to reveal faint exoplanets with contrast in the order of 10^{-5} at angular separations of less than $1''$. As presented in the top panel of Figure 1.6, Airy rings of the stellar PSF are still brighter than the Airy core of such a potential companion. To reduce the amount of stellar flux at separations that are larger than the diffraction limit of the system, so-called coronagraphs are used. The concept of coronagraphy was developed by the French astronomer Bernard Lyot, who aimed to study the solar corona (hence the name of the device; Lyot 1939). Inspired by the concept of a solar eclipse, Lyot tried to block the bulk of solar intensity without requiring the moon to move across the line of sight. The basic concept of such a Lyot coronagraph is presented in Figure 1.8; and this setup can also be used to decrease the intensity of unresolved stars. The light that enters the entrance pupil of the telescope is focused, and the central part is blocked by an opaque mask. Afterwards, the pupil is reimaged and residual starlight is removed by a Lyot stop in the pupil plane. Eventually, the stellar intensity in the final image is reduced drastically, whereas this design does not significantly affect the throughput of off-axis sources. In this way a faint planet that is usually hidden by the PSF of the primary star can become visible. In Figure 1.9 we present two exemplary images that emphasize the importance of a coronagraph in high-contrast imaging observations. A faint off-axis point source that is barely visible in the non-coronagaphic frame (left panel of Figure 1.9) is detected at high signal-to-noise ratio when using a coronagraph (right panel of Figure 1.9). By adding this additional piece of optic to the instrument, the contrast in the science images improves significantly.

Despite many developments to improve the performance of such focal-plane coronagraphs (e.g., Guyon 2003; Soummer 2005), the underlying principles of this basic

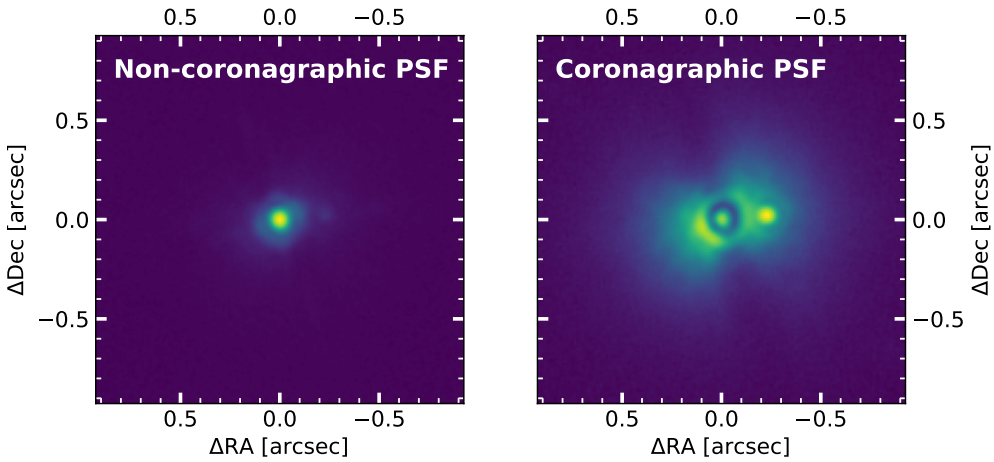

Figure 1.9: Comparison of images obtained without and with coronagraph. *Left panel*: SPHERE non-coronagraphic K_s band image of a point source with a faint companion. The exposure time is 2 s, and a neutral density filter attenuates the collected flux by a factor of 6.9. *Right panel*: SPHERE coronagraphic K_s band image of the same system. The exposure time is 8 s and no neutral density filter is placed in the optical path.

concept that was developed almost a century ago are still in use in modern high-contrast imaging instruments. Another class of coronagraphs are mounted in the instrument's pupil plane instead (e.g., Soummer et al. 2003; Codona et al. 2006; Kenworthy et al. 2007; Snik et al. 2012). These alter the stellar PSF and create dark holes around the primary star by destructive interference of light. For a recent review about past, current, and future coronagraphic concepts and technologies the reader is referred to the review articles of Mawet et al. (2012) and Ruane et al. (2018).

Observation strategies and post-processing In addition to AO systems and coronagraphs, the final contrast can further be optimized based on suitable observing and post-processing strategies. Usually, this contrast improvement is achieved by modeling and subtraction of the stellar PSF.

Angular differential imaging It is most common to carry out high-contrast imaging observations in pupil-stabilized mode, which enables so-called angular differential imaging (ADI; Marois et al. 2006a) techniques in post-processing. In this observing mode the telescope's field-derotator is disabled. This leads to a rotation of the imaged field of view throughout the observing sequence. The general idea of this setup is that the stellar PSF halo, residual aberrations, and speckles in the focal plane around the target star are (quasi) static, whereas a potential off-axis companion would rotate according to its parallactic angle. Due to this artificially created diversity, the stellar PSF can be modeled by taking the median along the time dimension. If the parallactic rotation in the datset is sufficient, this PSF model only contains a negligible amount of signal from the off-axis point source. Hence, this model can be subtracted from each individual image; afterwards, the residuals are derotated and stacked to increase the signal-to-noise ratio of a potential companion.

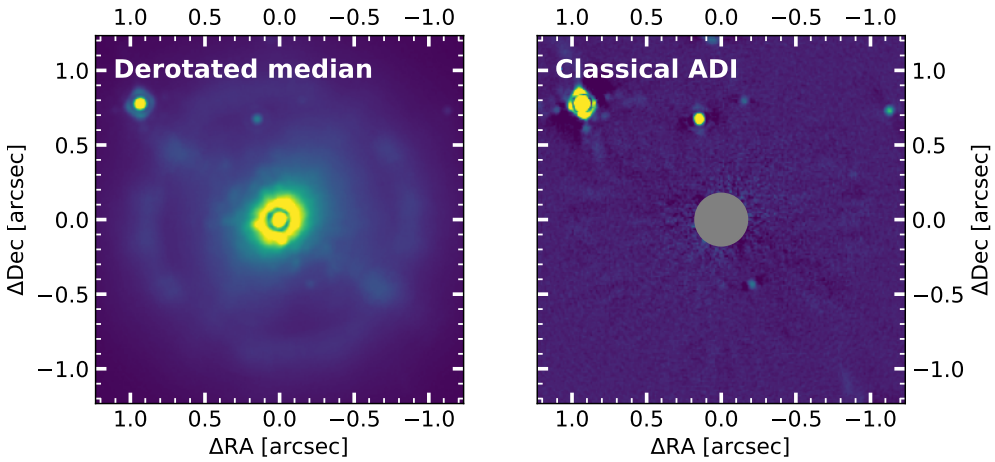

Figure 1.10: Contrast improvement with ADI. *Left panel*: Derotated and median combined image from an observing sequence collected in pupil-stabilized mode. *Right panel*: Classical ADI residuals for the same dataset. Several point sources that were hidden by stellar noise are now visible in the image. The inner region around the coronagraph is obscured by an artificial mask.

The contrast gain that can be obtained by this post-processing method is demonstrated in Figure 1.10.

Instead of just taking the median along the temporal dimension, there exist more advanced methods to obtain a PSF model from the data. The first class of these sophisticated PSF modeling algorithms relied on a Locally Optimized Combination of Images (LOCI; Lafrenière et al. 2007a); several advanced versions of the original LOCI code are available today (e.g., Marois et al. 2014; Wahhaj et al. 2015). Another class of PSF reconstruction algorithms relies on principal component analysis (PCA; Kendall 1957; Deeming 1964) that is creating an orthogonal set of basis vectors from the images via singular value decomposition. First PSF subtraction schemes that relied on this low rank approximation were `PynPoint` (Amara & Quanz 2012) and KLIP (Soummer et al. 2012); Meshkat et al. (2014) presented an optimized version of the original algorithm that includes frame rejection and different cutoff radii. Several additional PSF subtraction algorithms have been proposed in the past years, such as low-rank plus sparse decompositions (Gomez Gonzalez et al. 2016) or non-negative matrix factorization (Ren et al. 2018).

Spectral differential imaging The latest generation of extreme AO high-contrast imaging instruments such as the SPHERE, the Gemini Planet Imager (GPI; Macintosh et al. 2014), or the Subaru Coronagraphic Extreme Adaptive Optics System (SCExAO; Jovanovic et al. 2015) are all equipped with integral field spectrographs (IFS) that provide low resolution spectra ($R \sim 50$) for a field of view in the order of $2'' \times 2''$ (e.g., Claudi et al. 2008; Larkin et al. 2014; Peters et al. 2012; Groff et al. 2015). As the size of the stellar PSF, along with residual speckles, scales as a function of wavelength, the position of a planetary companion, however, is constant in the field of view, IFS observations deliver another option to discriminate between off-axis signals and contaminating stellar flux. Similar to ADI, this spectral diversity can be used to model and remove the stellar PSF without subtracting the signal of a potential planet. The

idea of using simultaneous images of different wavelengths to subtract stellar light was first introduced by Racine et al. (1999); Sparks & Ford (2002) recognized the potential of this strategy and proposed coronagraphic imaging with an IFS for an improved sensitivity to faint extrasolar planets. By now, spectral differential imaging (SDI) has become a standard technique for post-processing of high-contrast imaging IFS data, and SDI is included in several reduction algorithms (e.g., Marois et al. 2006b; Janson et al. 2008; Pueyo et al. 2012; Rameau et al. 2015; Galicher et al. 2018). SDI can also be combined with ADI, leading to improved contrasts compared to either of these methods applied by themselves (Vigan et al. 2010; Christiaens et al. 2019; Kiefer et al. subm).

Reference star differential imaging One of the oldest methods to subtract the PSF of a primary star is by observing a reference star, whose PSF is used as a template for the science target. Smith & Terrile (1984) used this technique to detect a dusty disk around the young star β Pictoris.[7] For the same primary star, Lagrange et al. (2009) later found a giant Jovian planet that was located inside the dusty debris disk. Lagrange et al. discovered the companion β Pictoris b in archival NACO data from 2004, which were reassessed by subtraction of a reference PSF acquired on HR 2435, a nearby star of similar spectral type. This basic concept of reference star differential imaging (RDI) proved to be very useful to study especially circumstellar environments (Mawet et al. 2009; Rameau et al. 2012). ADI does not work particularly well for these science cases, as the angular variety of disks observed at low inclinations is small; this leads to remaining disk signal in the subtracted PSF model and hence the over-subtraction of scattered light signal (e.g., Milli et al. 2012). RDI reductions were significantly improved by the introduction of PCA-based processing schemes. Especially the Archival Legacy Investigations of Circumstellar Environments (ALICE; Hagan et al. 2018) project revealed several circum-stellar disks from archival HST data that were previously undiscovered (e.g., Soummer et al. 2014; Choquet et al. 2014, 2016, 2017).

But RDI has also been successfully employed for the detection and characterization of extrasolar planets (e.g. Soummer et al. 2011; Wahhaj et al. 2021). Especially our Young Suns Exoplanet Survey (YSES, see Section 1.4) heavily relies on an RDI-based data reduction strategy; and with the combination of RDI and PCA we discovered disks and planets (see Bohn et al. 2019, 2021).

Other approaches Instead of traditional PSF subtraction algorithms, there are also other methods to identify planets in high-contrast imaging data. These rely either on a forward model, such as ANDROMEDA (Cantalloube et al. 2015) and TRAP (Samland et al. 2021); statistical discrimination between speckles and planet signals as proposed by Gladysz & Christou (2008), or as implemented in several variants of the PACO algorithm (Flasseur et al. 2018, 2020), the STIM map (Pairet et al. 2019), and the RSM detection map (Dahlqvist et al. 2020, 2021); machine learning algorithms (e.g., Gomez Gonzalez et al. 2018; Gebhard et al. 2020); or other iterative approaches such as MAYONNAISE (Pairet et al. 2021). For a recent performance evaluation of various post-processing algorithms the reader is referred to Cantalloube et al. (2020).

[7] Note that the final image presented in Smith & Terrile (1984) is a ratio image and not a differential image. According to the authors' description, such a difference image, which showed the circum-stellar disk, was created as well.

Results and prospects of direct detections of exoplanets

By overcoming all these challenges, direct imaging of exoplanets is indeed possible. The first planetary-mass companion that was discovered by direct imaging is 2MASS J12073346-3932539 b that is orbiting its M type primary at an angular separation of more than 55 au (Chauvin et al. 2004). One of the most emblematic environments that could be directly imaged so far is the multi-planetary system around HR 8799 (see Figure 1.11). Marois et al. (2008, 2010) revealed four gas giant planets in orbits of 15 au to 70 au around this A type[8] star. Continuous monitoring of this intriguing system showed significant amounts of orbital motion,[9] revealing that the system exhibits an almost face-on geometry (e.g., Maire et al. 2015; Wang et al. 2016).

Direct imaging is currently only sensitive to young, super-Jovian planets that are widely separated (usually $\geq 0\farcs2$) from their primary stars. This bias is clearly visible in Figure 1.1, where the directly imaged planets exclusively occupy the upper right of the mass versus semi-major axis parameter space. As this technique cannot reliably probe Solar System scales in terms of planet masses and physical separations, it seems to be inferior to other methods that can at least partly detect these Solar System analogs. Indeed, this current bias can also be interpreted as a particular strength unique to the direct imaging method as no other exoplanet detection technique is capable to probe numerous planets with semi-major axes that are larger than 5 au. Direct imaging thus provides unique insights into the wide-orbit architectures of planetary systems up to thousands of Astronomical Units (e.g. Naud et al. 2014). These extreme scales are not even accessible in our own Solar System, which is hypothesized to host an additional, undiscovered Planet 9 at several hundreds of Astronomical Units (Batygin & Brown 2016).

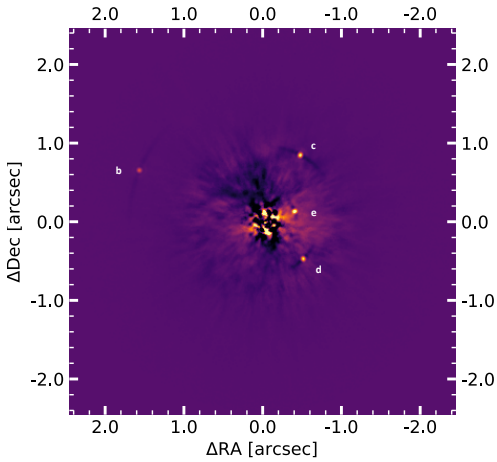

Figure 1.11: The multi-planetary system around HR 8799. The star is located at the center of the image and post-processing with ADI plus PCA is performed to remove the stellar halo. Four planets (labeled b, c, d, and e) are clearly visible in the image. North points up and east towards the left. The data was collected with VLT/SPHERE on the night of 2017 October 11 (ESO ID: 099.C-0588(A); PI: Biller) and published in Biller et al. (2021).

Direct imaging is the only technique which spatially resolves the planet and its host star. This facilitates unique opportunities of planet characterization that are hardly accessible by indirect detection methods. Even though spectra of transiting planets can be obtained by evaluation of the transit depth for several wavelengths (Brown 2001), the spectra of directly imaged companions usually exhibit

[8]The spectral type of HR 8799 is controversial as different tracers provide various estimates ranging from A5 to F0 (Gray & Kaye 1999).

[9]A nice visualization of the planets orbiting around the star is available online: `https://en.wikipedia.org/wiki/HR_8799#/media/File:HR_8799_Orbiting_Exoplanets.gif`

much higher signal-to-noise ratios and can be obtained with much higher efficiency: observations can be collected in a continuous observing sequence, whereas the combination of several transit measurements is required to perform transit spectroscopy. Despite several ground-breaking results from transit spectroscopy – such as measurements of the cloud coverage on Kepler-7b (Demory et al. 2013), the detection of Na I in the atmosphere of HD 189733b (Wyttenbach et al. 2015), and atomic iron and titanium in the atmosphere of Kelt-9b (Hoeijmakers et al. 2018a) – direct imaging is by far the most promising method to characterize the atmospheres and to evaluate the habitability of terrestrial planets in the future (e.g., Biller & Bonnefoy 2018), particularly due to the special system geometry that is required for the planet to transit and the time-intensive nature of transit-based characterizations.

One of the first directly imaged exoplanets that was characterized by consecutive spectroscopic measurements with Keck/OSIRIS (Larkin et al. 2006) was HR 8799b (Barman et al. 2011). Barman et al. found a hydrogen-dominated atmosphere, and the authors were able to constrain the planet's temperature, luminosity, and surface gravity. Further major results from these spatially resolved exoplanets are among others: water and carbon monoxide absorption features detected in the atmopshere of HR 8799c (Konopacky et al. 2013); rotation period measurements of the fast-spinning super-Jupiter β Pictoris b (Snellen et al. 2014); and the detection of two accreting protoplanets that are carving out the gap around the young T-Tauri star PDS 70 (Keppler et al. 2018; Wagner et al. 2018; Haffert et al. 2019). Interferometric observations with VLTI/GRAVITY (Gravity Collaboration et al. 2017) exhibit an enormous potential to perform medium resolution spectroscopy ($R \sim 500$) of Jovian exoplanet atmospheres even beyond the diffraction limit of current 10 m-class optical telescopes (Gravity Collaboration et al. 2019, 2020; Nowak et al. 2020; Wang et al. 2021).

Even though terrestrial planets are too faint to be directly characterized by the current generation of telescopes, future observatories and instruments might provide the required angular resolution, contrast performance, and sensitivity. These might be either space-based telescopes with large apertures such as LUVOIR (The LUVOIR Team 2019) or HabEx (Mennesson et al. 2016; Gaudi et al. 2020) or ground-based extremely large telescopes that will become available over the coming decades. The European ELT, currently under construction, will have an unprecedented primary mirror diameter of 39 m (Gilmozzi & Spyromilio 2007). METIS, one of the first-light instruments at the ELT, is expected to image and characterize terrestrial exoplanets around the nearest stellar neighbors (Quanz et al. 2015). Ambitious future projects such as LIFE might be required in the end to reveal first unambiguous signs of biosignatures outside Earth (Quanz et al. 2019, 2021).

1.3 Formation of gas giant planets

The technical advances of the past decades facilitated the direct detection of several gas giant planets outside our Solar System. Whereas some of these reside within orbits that are comparable to those of Jupiter, Saturn, Uranus, or Neptune (e.g., β Pictoris b at $\sim$8 au, Lagrange et al. 2009, 2010; Quanz et al. 2010; 51 Eridani b at $\sim$13 au, Macintosh et al. 2015; PDS 70 b at $\sim$22 au, Keppler et al. 2018; PDS 70 c at $\sim$30 au, Haffert et al. 2019), the majority of these Jovian giants exhibit significantly larger semi-major axes ($>$50 au) than all Solar System planets (e.g., Chauvin et al.

2005; Schmidt et al. 2008; Rameau et al. 2013; Bailey et al. 2014; Chauvin et al. 2017a; Bohn et al. 2020a,b). The formation pathway and early evolution of these wide-orbit systems is poorly understood. It remains unclear whether these gas giants have formed in situ by (A) fragmentation processes of the collapsing proto- stellar cloud – thus representing the lower-mass threshold of the stellar binary population (Kroupa 2001; Chabrier 2003) – or if they were born at closer separations to the star (B) via core accretion mechanisms (e.g., Pollack et al. 1996; Alibert et al. 2005; Dodson-Robinson et al. 2009; Lambrechts & Johansen 2012), followed by a potential outward migration that is driven by scattering events (e.g., Veras et al. 2009; Mustill et al. 2021). In situ formation via core accretion is thought to be very unlikely as the timescales to build up a solid core that is massive enough to accrete a gaseous atmosphere are too long at separations larger than 50 au compared to typical disk lifetimes (Haisch et al. 2001; Rafikov 2011). A third scenario that supports planet formation for a wide range of orbital separations is (C) the disk instability paradigm that postulates the formation of planetary cores from gravitational collapse of dense regions in the protoplanetary disk (Boss 1997, 2011; Rafikov 2005; Durisen et al. 2007; Kratter et al. 2010; Kratter & Lodato 2016).

There are two common approaches to evaluate the efficiencies of these planet formation scenarios:

1. In a statistical framework, simulated planet occurrence rates based on the various formation mechanisms (e.g., Mordasini et al. 2009a,b; Forgan & Rice 2013; Forgan et al. 2018) are compared to observational survey results, to place constraints on the efficiency of the corresponding formation pathway.

2. Spectral characterization of identified wide-orbit gas giants can further link the atmospheric molecular and isotopologue abundance ratios of these companions to their natal environments.

I will further discuss these scenarios in Sections 1.3.1 and 1.3.2, respectively.

1.3.1 Direct imaging surveys

Several direct imaging surveys have been conducted in the past years (e.g., Lafrenière et al. 2007b; Kasper et al. 2007; Biller et al. 2013; Nielsen et al. 2013; Galicher et al. 2016; Bowler 2016), and the analyses of the most relevant studies are still ongoing. The preliminary statistical evaluation of the first 150 stars observed as part of the the SpHere INfrared survey for Exoplanets (SHINE; Vigan et al. 2020) indicates that substellar companions (with masses in the range $1\,M_{\rm Jup} - 75\,M_{\rm Jup}$) around M-type stars likely formed via either protostellar fragmentation processes (A) or gravitational instabilities (C), whereas the occurrence rates of objects of the same mass around B and A-type stars are consistent with planet populations from core accretion simulations (B). The statistical data for the intermediate category of F, G, and K-type stars does not favor a single formation pathway, but can be explained best by a combination of all scenarios. These findings are corroborated by the Gemini PLanet Imager Exoplanet Survey (GPIES; Macintosh et al. 2018). Nielsen et al. (2019) find that giant planetary companions with masses between $1\,M_{\rm Jup}$ and $13\,M_{\rm Jup}$ are predominately created by bottom-up formation mechanisms (B), whilst the population of brown dwarf companions with masses in the range $13\,M_{\rm Jup} - 80\,M_{\rm Jup}$ is rather compatible with top-down formation mechanisms (A,C). These distinct formation pathways for

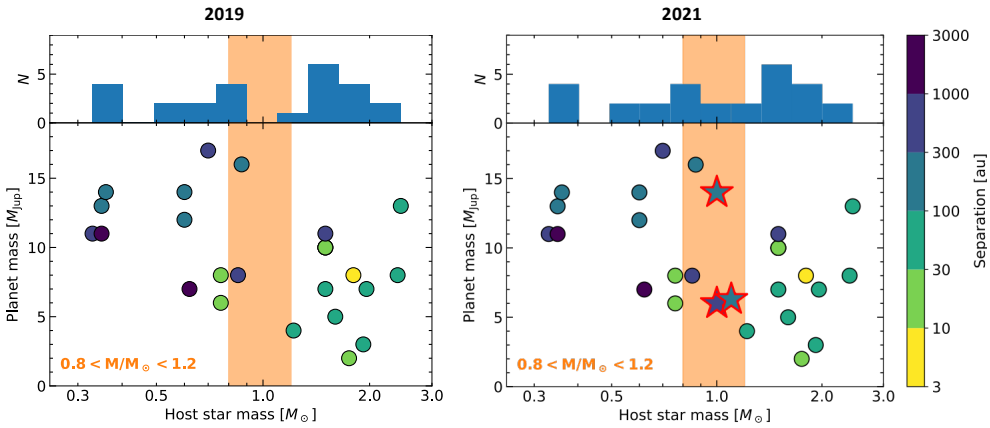

Figure 1.12: Demographics of directly imaged exoplanets in 2019 (*left panel*) and 2021 (*right panel*). The orange box highlights solar-like host stars with $0.8 < M/M_\odot < 1.2$. The stellar markers with the red outlines indicate new planetary-mass companions that were discovered as part of our Young Suns Exoplanet Survey (see Section 1.4).

the sample of giant planets and brown dwarfs are in perfect agreement with recent results from Bowler et al. (2020). Based on analyses of orbital parameters of directly imaged companions, Bowler et al. find that the eccentricity distributions show clear evidence for two distinct populations, indicative of different formation channels.[10] Both Vigan et al. (2020) and Nielsen et al. (2019) find that heavier mass stars are more likely to harbor gas giant companions with semi-major axes in the range 5 au − 100 au, which is confirmed by results from RV surveys for planets with shorter periods (Johnson et al. 2010; Ghezzi et al. 2018). To study if this relation even holds for stars with masses that are larger than 2 $M_\odot$, Janson et al. (2021) are currently carrying out the B-star Exoplanet Abundance Study (BEAST) that is targeting the most massive stellar members of the Scorpius-Centaurus association (Sco-Cen; de Zeeuw et al. 1999).

All these results indicate that planet formation depends heavily on the mass of the primary star. Especially for the intriguing sub-class of Sun-like stars, no dominant formation pathway of gas giant companions could be established so far. This can also be attributed to the small number of exoplanets that have been found around solar-type host stars. As presented in the left panel of Figure 1.12, there seems to be a void of planets around solar analogs (until 2019). However, this might also be an observational bias as young (<20 Myr), solar-type stars were largely neglected in past direct imaging searches for extrasolar planets. This is mostly due to the fact that the majority of such young targets are located at distances of more than 100 pc, and therefore exhibit R band magnitudes that are fainter than 10 mag. As the quality of the AO correction is usually degrading with a decreasing flux from the natural guide star, previous surveys focused on the significantly brighter population

[10] Note that Bowler et al. (2020) apply slight different thresholds than Nielsen et al. (2019) for objects to be considered either giant planet or brown dwarfs. In Bowler et al. (2020) giant planets are selected with masses in the range 2 M_{Jup} − 15 M_{Jup} and brown dwarfs are identified as objects with masses between 15 M_{Jup} and 75 M_{Jup}.

of B, A, and F type stars of these associations.[11] Despite the worse AO correction on young G and K type stars at distances of more than 100 pc, these targets can be observed with the latest generation of high-contrast imaging instruments, and the astrometric precision is sufficient for the required proper motions checks (see Chapters 4, 5, and 6). It is thus vital to use the capabilities of these state-of-the-art instruments to systematically study a population of young, Sun-like stars with statistical significance. Such observations might reveal important clues regarding the origin of wide-orbit Jovian companions in young, Sun-like environments.

1.3.2 Atmospheric characterization of directly imaged planets

To identify the birthplaces of individual gas giants, Öberg et al. (2011) proposed to use elemental abundances in the planetary atmosphere as a tracer. The C/O ratio might be indicative of the natal environment of a planet in a circum-stellar disk. Due to a radial temperature gradient in the disk, certain molecules freeze out at characteristic separations from the primary star. These transition regions are called icelines (or snow lines when talking especially about H_2O; Stevenson & Lunine 1988; Lodders 2004; Kuchner & Seager 2005). These processes affect the abundance ratios of molecules that are either in the solid or gas phase. For a typical young, solar-type star the H_2O snowline is located at approximately 3 au (e.g., Hayashi 1981; Podolak & Zucker 2004; Martin & Livio 2012). Farther out water molecules are predominantly present in their solid phase, whereas carbon bearing species such as CO_2 or CO are still in their gas phase. Accordingly, the C/O ratio in the gas increases when radially moving outward across the H_2O snowline, whereas the C/O ratio in the grains decreases (as the amount of oxygen in the gas phase is rising, whilst the carbon fraction remains unaltered). Similarly, the carbon to oxygen ratio is altered at the CO_2 and CO icelines, which are located at approximately 10 au and 40 au in the protosolar nebular, respectively (Öberg et al. 2011; Andrews & Williams 2007). The atmosphere of a gas giant planet should therefore exhibit imprints of this variable atmospheric abundance ratio; indicative of its natal environment with respect to the primary star. But disk evolution or enrichment by planetesimals can alter these primordial abundances and hence should be considered in the interpretation (Ali-Dib et al. 2014; Mordasini et al. 2016; Eistrup et al. 2016, 2018). Planet migration and structural inhomogeneities in the natal disk complicate establishing links between present day abundances and the planetary birth environment (e.g., Madhusudhan et al. 2014; Cridland et al. 2016). These issues are extensively discussed in recent review articles of Pudritz et al. (2018), Lammer & Blanc (2018), Madhusudhan (2019), and Öberg & Bergin (2021). Nevertheless, the study of a large sample of giant planetary atmospheres is necessary to identify potential correlations that might be indicative of planet formation history. New GRAVITY data on β Pictoris b revealed a subsolar carbon to oxygen abundance ratio for this super-Jovian gas giant (Gravity Collaboration et al. 2020). This was interpreted as an indication for its formation via the core-accretion channel (B), followed by a strong enrichment by icy planetesimals that exhibit a C/O ratio that is below the solar standard.

These unique insights into the formation history of planetary systems might even be further corroborated when measuring isotope abundance ratios such as D/H or

[11] The SPHERE user manual predicts median (H band Strehl ratio of $50-75\,\%$) to poor (H band Strehl ratio $< 50\,\%$) AO performances for stars that are fainter than $R = 9$ mag: `https://www.eso.org/sci/facilities/paranal/instruments/sphere/overview.html`

$^{12}C/^{13}C$, which impose additional constraints on the location of the natal environment (Mollière & Snellen 2019; Morley et al. 2019). Such a measurement in an exoplanet atmosphere, could not be conducted thus far. New instrumentation such as VLT/CRIRES^{+} (Dorn et al. 2014), VLT/ERIS (Davies et al. 2018), and the James Webb Space Telescope (JWST; Gardner et al. 2006), which will all become available shortly, might provide the required sensitivity to measure isotopologues contents of an exoplanet atmosphere for the first time. In addition to the large statistical surveys described in Section 1.3.1, characterization studies of known wide-orbit gas giants and measurements of their atmospheric compositions are thus promising next steps to understand the origins and formation pathways of gas giant Jovian companions.

1.4 The Young Suns Exoplanet Survey

To enhance the low number of planetary companions to Sun-like primaries and to test their formation scenarios at statistical significance, we started the Young Suns Exoplanet Survey (YSES) that targets a homogeneous sample of 70 young, solar-mass members of the Lower Centaurus-Crux (LCC) subgroup of Sco-Cen. This subgroup has an average age of 15 ± 3 Myr (Pecaut & Mamajek 2016); latest parallax and proper motion measurements provided by the third early data release of the *Gaia* space mission of the European Space Agency clearly confirm the LCC membership of our targets (*Gaia* ERD3; Gaia Collaboration et al. 2021). The YSES sample has an average distance of 118 ± 20 pc, and it consists exclusively of stars with masses in the range $0.8\,M_{\odot} - 1.2\,M_{\odot}$. Therefore, it is a unique collection of the nearest statistically significant sample of young, Sun-like stars to our Sun. There is no closer (≤ 120 pc) and younger (≤ 100 Myr) association with so many stars of about $1\,M_{\odot}$, making the YSES sample particularly significant for statistically constraining the wide-separation planetary population around young solar analogs.

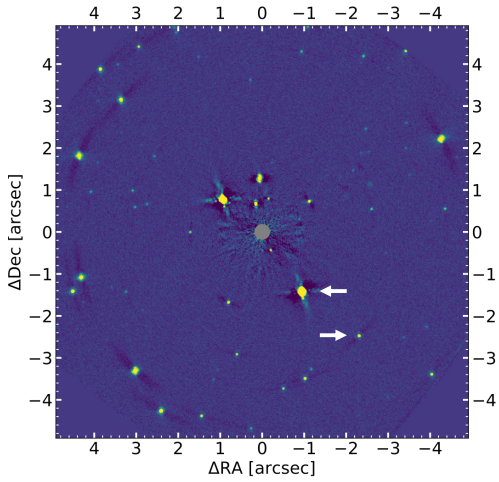

Figure 1.13: Exemplary field of view of an YSES target. As most of the stars from the sample are close to the galactic plane, there is a very high-fraction of background contaminants. From a single image as presented here, it is not possible to distinguish planetary companions (white arrows) from unassociated background stars (all other off-axis point sources in the image). The primary star is located at the image center behind an artificial mask.

YSES is still ongoing: at the moment, we have observed each of the 70 stars once with SPHERE. Our data reduction strategy is based on short snapshot observation in the order of $\sim$2 min per target and filter, in combination with RDI. As all stars within our sample are coeval, at the same distance, and of similar spectral type, they exhibit similar magnitudes at both optical and near-infrared wavelengths. This homogeneity facilitates highly consistent observing conditions in terms of recorded near-infrared flux and

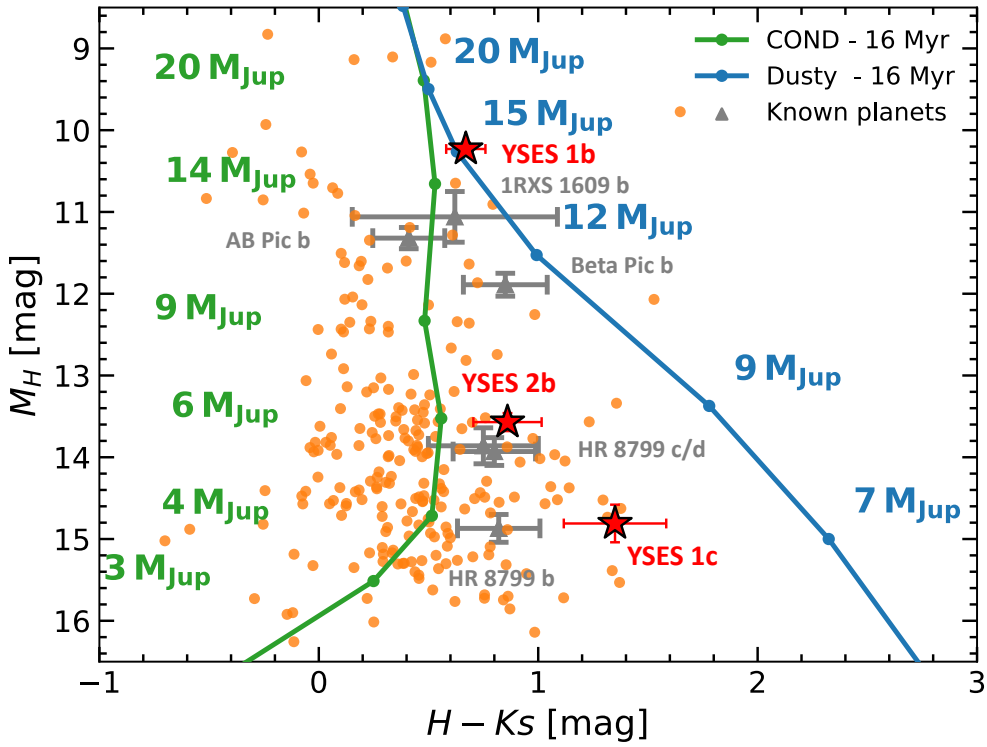

Figure 1.14: Color magnitude diagram of YSES candidate companions. The orange circles indicate off-axis objects detected within the scope of our survey. The red stars highlight actual companions that we discovered. The grey markers represent previously known directly imaged companions and the green and blue lines are from synthetic evolutionary models that were evaluated at the approximate age of LCC.

instrumental AO performance. Hence we can use all frames collected for YSES as a reference library, which we can analyze with PCA. For each target, the stellar PSF is then modeled by a certain number of principal components, and subtracted. In Chapter 6 we asses the performance of this reduction method and show that it is superior to other potential reduction schemes.

Our first epoch observations revealed a plethora of companion candidates. As LCC is close to the galactic plane, a high fraction of associated background stars is therefore expected to be revealed by our observations. This is impressively demonstrated by the image presented in Figure 1.13. More than 50 off-axis point sources are detected around the primary star. From a single snapshot like this, we cannot tell whether one of these objects is indeed a planetary companion or rather a background contaminant. For that reason, we took data in an additional filter to analyze the detected companion candidates in a color-magnitude diagram. As presented in Figure 1.14, almost all the candidate companions that we detected around the sample of 70 young suns are compatible with predictions from evolutionary models or known sub-stellar companions. Since sub-stellar objects around the L-T transition can appear colorless or even blue in the near infrared (Burgasser et al. 2002; Cushing et al. 2005), this pre-selection of interesting candidate companions by color-

magnitude analysis did not work well. Instead we established that second epoch observations and checks for common proper motion with the primary are indispensable measures to confirm gravitationally bound companions around our YSES targets (e.g., Nielsen et al. 2017).

As we did not want to bias our selection of candidate companions by angular separation cutoffs (with preference for small angular separations), we decided to follow-up all targets with candidate companions in the instrument field of view. This turned out to be the right choice as we found gravitationally bound companions at unexpectedly large separations of more than 100 au. For eight targets we could acquire second epoch-observations; these data revealed three planetary-mass companions to two of our stars. Most intriguing was the discovery of the first multi-planet system that was imaged around a Sun-like star (see Chapter 5). We found that YSES 1 (TYC 8998-760-1) hosts two giant planets of $14 \pm 3\,M_{\rm Jup}$ and $6 \pm 2\,M_{\rm Jup}$ at physical separations of more than 160 au and 320 au, respectively. Due to their moderate separations, this system is a prime laboratory to perform consecutive characterization measurements to probe their atmospheric composition and potential formation scenarios. The latest addendum to the small group of directly imaged planets to Sun-like stars is YSES 2b (see Chapter 6): a $6.3^{+1.6}_{-0.9}\,M_{\rm Jup}$ planet at a projected separation of 115 au. This gas giant is challenging current theories of giant planet formation: it is too far separated from the star to have formed in situ via core accretion mechanisms (B), yet it is not massive enough to be compatible with latest models of gravitational instabilities in circum-stellar disks (C). Follow-up measurements of this intriguing system might be able to probe further companions to YSES 2, that could have scattered YSES 2b to its current separation.

Follow-up observations for 44 remaining targets with candidate companions are still pending. We might have been extremely lucky in the selection of our 8 systems that were chosen for follow-up, but it is also likely that there are indeed more planet discoveries from YSES ahead. Even though the survey is not completed yet, YSES already contributed significantly to populate the previously sparsely sampled group of wide-orbit giant planet to solar-type host stars. As shown in the right panel of Figure 1.12, our three detections from YSES already helped to fill the apparent void of such companions; this is tentatively indicating that it was indeed an observational bias, as young Sun-like systems have not been systematically studied at statistical significance prior to our survey. Future prospects of this extremely successful and efficient program are discussed in Chapter 7.

1.5 Thesis outline

Except for Chapter 2 this thesis mainly focuses on the results from YSES. A brief description of the most important results and conclusions reported in each chapter is provided below.

Chapter 2: A multiplicity study of transiting exoplanet host stars

In this chapter we aim to study the stellar multiplicity of known systems to transiting exoplanets. We use VLT/SPHERE to study 45 transiting exoplanet host stars at unprecedented spatial resolution and contrast. We discovered new stellar companion candidates to 13 stars from the sample, which need to be validated by further proper motion measurements. Instead we assigned a likelihood for each companion to be

bound by evaluating stellar population synthesis models. In addition we could detect and confirm previously known companions to 13 other stars from our sample. Based on these data we derived a likelihood of $55.4^{+5.9}_{-9.4}$ % for transiting exoplanet host stars to be part of a multiple system. This measurement is in good agreement with previous surveys and with multiplicity estimates of stars that do not harbor any known planets.

Chapter 3: Discovery of a directly imaged disk in scattered light around the Sco-Cen member Wray 15-788

This chapter presents the first detection of YSES. Around the ∼11 Myr-old K3 star Wray 15-788 we detect clear scattered-light signals from a protoplanetary disk. SED modeling indicates that this disk is at a transition stage, and we detect clear structures of an outer and inner arc, separated by a gap of reduced scattered light flux. Even though the outer arc of the disk is detected almost face on with an inclination of 21 deg ±6 deg, we do not detect disk signal for all phase angles. We hypothesize that this lack of scattered light flux originates from a misaligned inner disk, interior to the inner working angle of our observations, that is casting a shadow on the outer structures that we see in the SPHERE images. This misalignment might be caused by a massive companion in the system. We further identify Wray 15-788 as secondary to the A type primary HD 98363 with a projected separation of about 7000 au. This massive primary hosts a debris disk that was first revealed in scattered light by GPI after the publication of our paper (Hom et al. 2020). The different evolutionary stages of these two coeval disks tentatively support the giant companion hypothesis. Such a companion could cause pressure bumps in the disk that are trapping dust grains at certain separations, leading to the observed gapped geometry (e.g., Pinilla et al. 2015). Future observations with both SPHERE and ALMA will shed light on this fascinating circum-stellar environment.

Chapter 4: Detection of a wide orbit planetary mass companion to a solar-type Sco-Cen member

In this chapter, we present the first planetary-mass companion that was detected by YSES. YSES 1b (TYC 8998-760-1 b) has a mass of $14 \pm 3\,M_{\rm Jup}$ and sits thus directly at the boundary between gas giant planets and low mass brown dwarfs. We measure a projected separation of ∼160 au which is indicative of an orbital period of more than 1'000 yr. The combination of high mass and wide separation favor formation via top down scenarios (A,C); yet further evidence to corroborate this theory is required.

Chapter 5: Two directly imaged, wide-orbit giant planets around the young, solar analog YSES 1

In data obtained to characterize YSES 1b, we identified a very red point source, which turned out to be another planet around this solar-type primary. This discovery is described in Chapter 5 of this thesis. YSES 1c has a lower mass ($6 \pm 2\,M_{\rm Jup}$) than the previously identified YSES 1b. The new planet is even farther separated from its primary star with a projected separation of ∼320 au. This discovery marked the first image of a multi-planet system around a Sun-like star. Further analysis of this intriguing environment might provide vital insights into the formation processes and the dynamical evolution that were shaping this system.

Chapter 6: Discovery of a directly imaged planet to the young solar analog YSES 2

This chapter presents the latest YSES discovery: A new planet to the 1.1 $M_{\odot}$ host star YSES 2 (TYC 8984-2245-1). YSES 2b has a mass of $6.3^{+1.6}_{-0.9}\,M_{\rm Jup}$ and a minimum physical separation of 115 au with respect to its primary star. This companion is especially fascinating due to its dubious origin: in situ core accretion (B) can be ruled out due to the large separation; yet gravitational instabilities (C) tend to produce heavier mass companions. As YSES 2b might have been scattered to its current location, a deep search for additional companions in the system is required.

Chapter 7: Outlook

In this chapter, we present the next steps that are required for the conclusion of the YSES program. We further propose a potential successor program that is observing a large sample of young, Sun-like stars with ages in the range $1 - 20$ Myr, to obtain a temporally resolved insight into the occurrence rates of gas giant companions to solar analogs. This program is likely to reveal several planetary-mass companions that allow for statistical testing of planet formation mechanisms, and detailed atmospheric characterization studies. Lastly, we come back to the initial question as to whether we are alone in the Universe, and present some speculations regarding potential insight that we might gain within the next decades.