

Exoplanets and Planetary Systems

Prof. Dr. Didier Queloz

Cavendish Astrophysics, University of Cambridge

Lent 2024 (February 16, 2024)

Contents

1	Historical Perspective	1
2	Planet Formation and Protoplanetary Disks	7
2.1	Introduction	7
2.2	Gas in the disk	7
2.3	Spectral energy distribution	10
2.4	Disk classification	11
2.5	Disk structure	12
2.5.1	Hydrostatic equilibrium	13
2.5.2	Momentum conservation	13
2.6	Planet formation	17
2.6.1	The core accretion model	17
2.6.2	Decoupling of gas and grains - looking at the disk plane	18
2.6.3	From dust to planetesimals - looking into the vertical direction	19
2.6.4	From planetesimals to planet cores	20
2.6.5	Isolation of planetary cores	27
2.6.6	Growth of thermally regulated envelopes	29
2.6.7	Disk-limited gas accretion rate	31
2.6.8	Gap formation, torque and migration	32
2.7	Timing early planet formation	38
3	Exoplanet detection	41
3.1	Elements of orbits and dynamics	41
3.1.1	Ellipse equations and Kepler's laws	41
3.1.2	Time evolution of orbit equations	43
3.1.3	Coordinate system transformation	44
3.1.4	Barycentric motion	45
3.2	Radial velocity technique	46
3.2.1	Studying the radial velocity of a star	46
3.2.2	Doppler effect	47
3.2.3	Noise limit to radial velocity measurements	48
3.2.4	Physical line broadening	50
3.2.5	The impact of stellar rotation ($v \sin i$)	50

3.2.6	Measuring the mass of a planet	52
3.2.7	Additional noise sources	54
3.3	Exoplanet transits and occultations	56
3.3.1	Impact parameter	57
3.3.2	Transit probability	57
3.3.3	Transit timing	58
3.3.4	Transit light curves	60
3.3.5	Transit shape and limb darkening	61
3.3.6	Impact parameter (cont'd)	62
3.3.7	Secondary eclipses	62
3.3.8	Transmission spectroscopy	63
3.3.9	Occultation spectroscopy	64
3.3.10	The Rossiter-McLaughlin effect	65

List of Figures

1.1	Solar System objects and artistic view of the Kuiper belt objects and Pluto eccentric orbit	1
1.2	Cumulative histogram of exoplanet discoveries trough time by various detection techniques.	2
1.3	Measured mass, radius and orbital period of all known exoplanets . Color code indicate techniques used to discover the planet (same than Fig.3). For mass measured by Doppler spectroscopy $\sin i = 1$ is considered. A grey scale density map is overlaid to locate "cluster of similar exoplanets" on these two diagrams.	3
1.4	All known exoplanets with a measurement of their mass and radius. Hatched lines indicates model of bulk density for three different compositions.	4
1.5	The orbital distribution of known Kuiper belt objects. The vertical lines mark the locations of some prominent mean motion resonances of Neptune. The cold classical and the scattered disk objects are indicated with green and orange ovals, respectively	5
2.1	The big picture from star to planet ©Stolker 2018	8
2.2	ALMA observation of the ^{12}CO	9
2.3	ALMA observation of the dusty disk of HR Tau. (Thermal signature)	10
2.4	Interstellar "dust" grain	10
2.5	Spectral Energy Distribution (SED) of star with disk at different stage	11
2.6	near IR image VLT/SPHERE proto-stellar disk PDS70 with a planet detection (Muller et al.2018)	12
2.7	Reflected light of protoplanetary disk seen edge-on measured with HST	14
2.8	The dispersion relation (solid black line) for a uniformly rotating disk, illustrating the contributions from pressure, rotation, and self-gravity (dashed blue lines). The system is unstable if, at any value of the wavenumber k , w^2 falls below the red line and is negative. (from Armitage 2020).	15

2.9	The plots show the vertically integrated density of solids, projected on the disk plane, relative to the initially uniform surface density . The x and y coordinates show the shearing-box dimensions in units of the gas disk scale height, H (the Sun is to the left; the orbital velocity vector points up). Time t increases from left to right as labelled (tsg). The circles in c depict the Hill spheres of clumps (Nesvorný 2019)	16
2.10	Four 4 stages of planetary formation related to the disappearance of the protoplanetary disk (@Williams and Cieza, 2011)	17
2.11	Geometry of the configuration	21
2.12	Schematics of the capture of the planetesimal by the embryo	22
2.13	N-body Simulation of formation of planetary cores by gravitational interactions	24
2.14	Examples of planetesimals accretion, viewed in the frame co-moving with the planet. In each panel the filled circle denotes the physical size of the planet and the dashed circle its Hill sphere. (c) and (d) qualify as pebble accretion, while (a), (b), and (e) fall in the ballistic regime. In (f) particles are so small that they follow gas streamlines (Orwell 2017)	24
2.15	25
2.16	Display of the five Lagrangian points in a rotating potential reference frame work.	26
2.17	Current combined measurements and model of the composition of the core of giant planets Jupiter and Saturn	28
2.18	Display of the Minimum Mass Solar nebula (MMSM)	29
2.19	Fraction of star with evidence of gaseous disk against age of the star (from Mamajek 2009)	31
2.20	The evolution of the core mass (solid line), envelope mass (dotted line) and total mass (dashed line) from a time-dependent calculation of giant planet formation via core accretion (Rice & Armitage, 2003).	32
2.21	Illustration of the two main regimes of aerodynamically assisted accretion. For lower mass planet cores (left) the Bondi radius r_B , within which purely gravitational 2-body interactions would lead to strong scattering, is smaller than the Hill radius r_H . Aerodynamically coupled particles enter the Bondi radius with their radial drift velocity and are accreted if their stopping time is shorter than the time it takes to traverse the Bondi radius. For higher mass planet cores (right) the Bondi radius exceeds the Hill radius, and the encounter velocity of particles approaching the planet is set by Kepler shear rather than radial drift. Only those particles that enter the Hill sphere have a chance of being accreted (from Armitage 2004)	33
2.22	Schematic of the planet-particle "impulse approximation"	34
2.23	Schematic illustration of the smoothed torque density due to angular momentum exchange between a planet and a gas from Ward (1997). The peak torque occurs at $r = r_p \pm h$, where h is the disk scale height	37
2.24	38

2.25	Allende meteorite, with zoom in on CAI's inclusions	39
2.26	Example of use of $\delta^{26}Mg$ method to get age of oldest rock formation	40
3.1	Display of conic section Family	42
3.2	Geometry of keplerian orbit	44
3.3	Astrophysical correspondance of keplerien orbit parameters	45
3.4	Coordinate system	46
3.5	The effects of e and ω on on radial velocity curves. These curves have been scaled to unit K and common P (courtesy Jason Wright)	47
3.6	Computed quality factor Q on a star function of the wavelength	50
3.7	Left: Decomposition if the $v \sin i$ effect on stellar surface, Right: geometrical definition of the $v \sin i$ lien broadening.	51
3.8	Example of stellar profile for various $v \sin i$ values	51
3.9	Effet of the $v \sin i$ on the quality factor for various stellar types (from Bouchy et al. 2001)	52
3.10	Geometrical description of the $v \sin i$ distribution	53
3.11	"historical mass function" in the range with giant planets, brow-dwarf ans stars	53
3.12	Schematics of the spatial conversion to wavelength occurring in a spectrograph	54
3.13	Power spectral distribution of long sequence of radial velocity measurements of the Sun-like star α -Cen A	55
3.14	Effect of stellar spot on the the line spectral profile	56
3.15	Geometrical configuration when a transit occurs	57
3.16	geometry of the transit	58
3.17	stellar light-curve of a transiting system	60
3.18	Left: Display of various transit shape $ \alpha - 1 $. Composite of 2346 transiting Kepler planets from the Exoplanet Orbit Database (Hippke 2019); Right: colour effect on the limb darkening	62
3.19	Transit transmission spectrum for WASP-127b (Sparke 2021). Points are measurements. Solid lines display best model fits according to retrieval of atmosphere composition.	64
3.20	Synthetical planetary spectra for the warm super-Earth GJ1214b, $R=4000$, for optically thin and thick haze. The transit depth and reflection of starlight are both given as the contrast to the host star. The thermal emission spectrum is the flux received at a distance of 1 pc (Nielsen 2016)	65
3.21	Geometry of the Rossiter-McLaughlin effect and corresponding angles	66
3.22	Angles that specify the orientation of the spin and orbital angular momentum vectors. The obliquity is ψ , the orbital inclination is i_0 , and the inclination of the stellar rotation axis is i	67
3.23	Effect on spectral line profil of a transiting planet	68

Preamble

Planets, along with the Sun and the Moon, are among the oldest celestial objects studied by mankind. The five planets visible from Earth by the naked eye¹, the Sun, and the Moon have been seen as the "Gods in the heavens", timekeepers from which we inherited the seven days of the week.

The Greeks identified these five special objects in the sky as the $\pi\lambda\alpha\nu\eta\tau\eta\varsigma$, the "wanderers". These heavenly bodies were considered by them as perfect objets of geometrical nature of spherical shape.

In the following centuries growth the understanding that planets and Earth are similar objects and together belong to a system orbiting the Sun. As it was clear that stars are simply other suns, this lead to the obvious question: what about planets on other stars - the exoplanets?

Amongst many key questions addressed by exoplanet research one finds : do other Earth exist? How can a solar system with inhabited planets form? Are such systems common? How does life form? Answering all those questions will take time. However, current studies on exoplanets are providing us a beginning of answer. The amazing diversity of planetary system architectures has already opened new perspectives about planetary systems formation mechanisms and evolution.

Answering this gigantic jigsaw is a field of science on its own, covering a large variety of physics astrophysics as well as planetary science, geophysics, high pressure mineralogy, computation, chemistry and microbiology. Today exoplanet science is mostly seen as an astrophysical topic but with increasing quality of data and the detection of exo-Earths the interest is growing Planetary and Earth science, biology and chemistry departments

This lecture addresses mostly the astrophysics side of exoplanet science. It aims at providing students with fundamental knowledge about exoplanet observations and formation theory.

The initial version of this document has been originally produced in 2015 by my student Max Günther from my hand-written notes. In course of the years it has been regularly updated thanks to comment and remarks of student attending this lecture. I thanks them all!

The material of this lecture is largely inspired by the book "Exoplanets" by Sara Seager (University of Arizona Press, 2010). For the physics of planet formation I used a lot of materials and updates from the excellent book "Astrophysics of Planet formation" by Philip

¹Mercury, Venus, Mars, Jupiter, Saturn

Armitage (Cambridge University press 2019).

Chapter 1

Historical Perspective

From historical perspective exoplanet discoveries are the continuation of discoveries that enlarged our understanding of our own solar system. With the discoveries of the first Jupiter moons (the Galilean moons) by Galileo Galilei in 1610, the discovery of Uranus (Herschel 1781), Neptune (Le Verrier and Galle 1846), the minor planets Vesta and Ceres (early 19th century) and the population of trans-Neptunian objects (TNOs)¹ such as Pluto, the solar system appeared to be way richer and more complex than the antic vision of the "five gods" suggested.

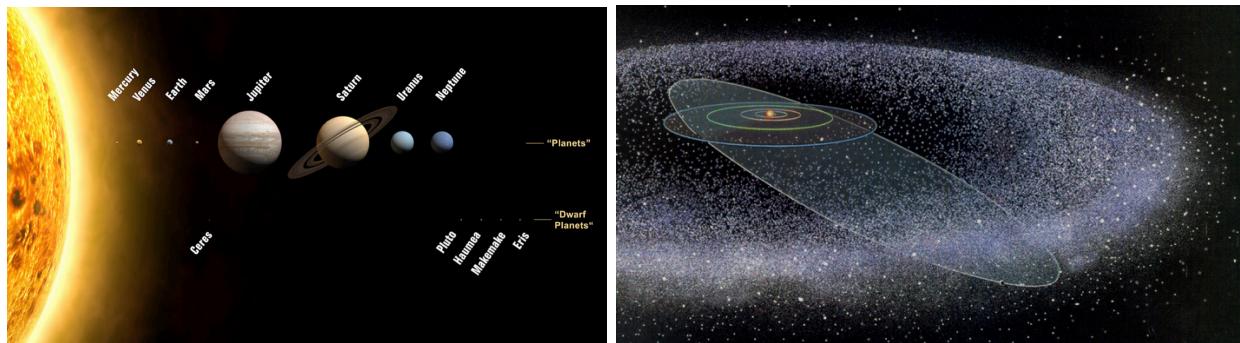


Figure 1.1: Solar System objects and artistic view of the Kuiper belt objects and Pluto eccentric orbit

From the shape of the solar system one derives the concept that planet are being form in disk shape structure. This idea was initially thought by Kant and later improved by Laplace and is known as the Kant-Laplace nebular hypothesis. It's modern developpement into "solar nebular theory" is derived from this brilliant 18th century insights and will stimulate search planet orbiting other stars.

The first claims of exoplanet detection can already be found in the early 20th century. The "infamous 24 year planet" around Barnard star, originally reported to be an exoplanet

¹minor planets that orbit the Sun at a greater average distance than Neptune, i.e. ≈ 30 AU; see also (15760) 1992 QB1, 28978 Ixion, 50000 Quaoar

by van de Kamp in 1963, is nowadays often cited as an example of detection mistake from underestimating instrument systematics.

The first serious attempt to search for exoplanets with the required precision is due to Cambell and Walker (1979). They embarked on a twelve year search for Jupiter-like planets on 21 night stars (by studying stellar light passing through hydrogen fluoride gas cells).

The first surprise in this field, in a way as an advanced notice of far more to come, was in 1992 when Wolszczan and Frail produced the evidence for two planets orbiting the pulsar PSR 1257+12. The second surprise came three years later, in 1995. First, in spring, Walker et al. published results on their 12 year search and reported no detection, challenging the prediction of Jupiter mass planets on other stars. Later, in autumn, contradicting the first study, the detection of an exoplanet was announced by Mayor and Queloz (1995) orbiting the star 51 Peg. With a mass of half the mass of Jupiter ($0.5M_J$) but an orbit 100 times shorter than Jupiter's it was challenging the foundation of planet formation and evolution. More details may be found at [Nobel Lecture: 51 Pegasi b and the exoplanet revolution](#)

At the present time thousands of exoplanets have been detected, mostly by Doppler techniques and transit measurements. In that sense we are experiencing a true "revolution". The fast and steady rise of planet discoveries since then (see on fig. 1.2) increased further the interest of the community for this new field.

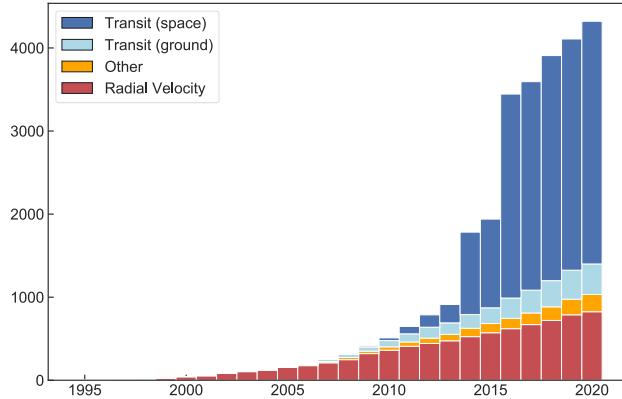


Figure 1.2: Cumulative histogram of [exoplanet discoveries](#) through time by various detection techniques.

As result of this feast of exoplanet we have learnt that our solar system architecture is far to be the norm. The wealth of diversity observed in exoplanet structures and orbital configurations is oddly contrasting with our Solar System configuration.(see fig. [1.3 on the next page](#))

So far according to our findings exoplanet configurations falls into three distinct groups of exoplanets. The group of giant planets found on short period (less than 10 days), with

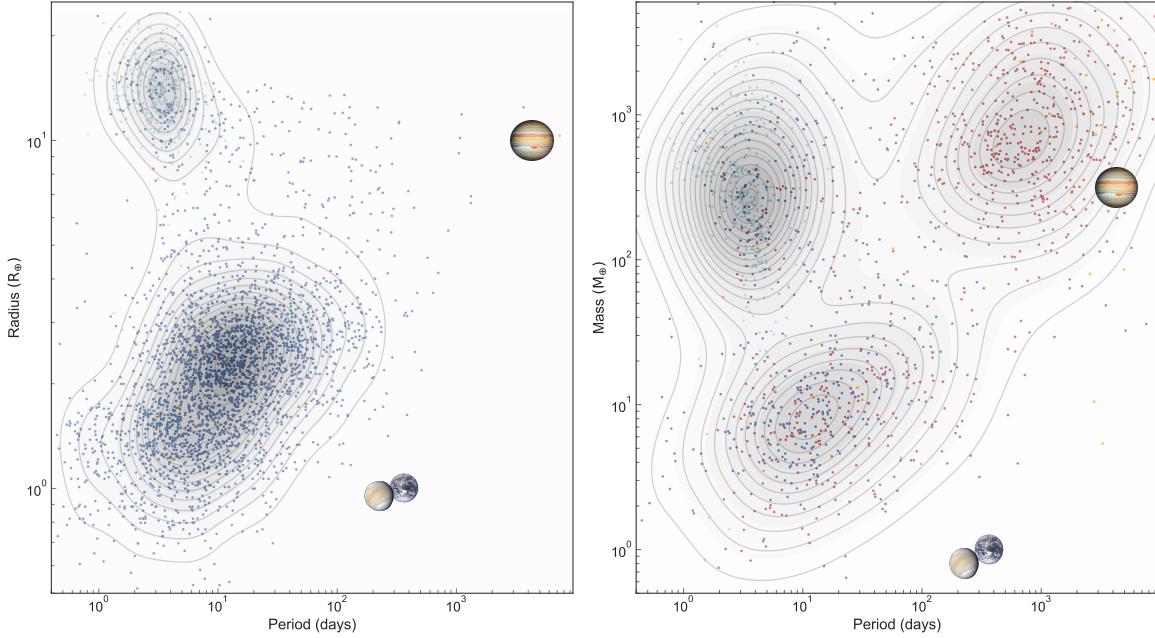


Figure 1.3: Measured mass, radius and orbital period of all known [exoplanets](#). Color code indicate techniques used to discover the planet (same than Fig.3). For mass measured by Doppler spectroscopy $\sin i = 1$ is considered. A grey scale density map is overlaid to locate "cluster of similar exoplanets" on these two diagrams.

51 Peg b its most famous member. On the colder end, further out, one finds "classical" giant planets like our own Jupiter. Then one sees a cluster of smaller exoplanets mostly on short orbit, casually named "Super Earth" or "Mini-Neptune" compact systems. When one look at the bulk density of this group of planets, one finds a mixed bag of anything fitting in a range defined on one side by Earths physical characteristics and on the other side by Neptune (see Fig. 1.4 on the following page).

In parallel during the same period the discovery of remote objects like Pluto in the TNO-population forced the International Astronomical Union (IAU) to revise the planet definition, demoting Pluto to a dwarf planet. The definition given by the IAU is technically not applicable to exoplanets since there is no clear understanding of the upper mass of a planet.

The current "gentleman's agreement" considers an exoplanet any object with a mass below the limiting mass for Deuterium thermonuclear fusion, currently predicted at $13 M_J$. Objects with a mass above this value are considered as "**brown dwarfs**", no matter how they have formed or where they are located.

Whatever their mass, planets make a negligible contribution ($\approx 0.13\%$) to the mass of the Solar System. The Sun is made up of hydrogen and helium and about 2% of heavier elements ($Z > 2$). This means that if a significant fraction of the current mass of the Sun

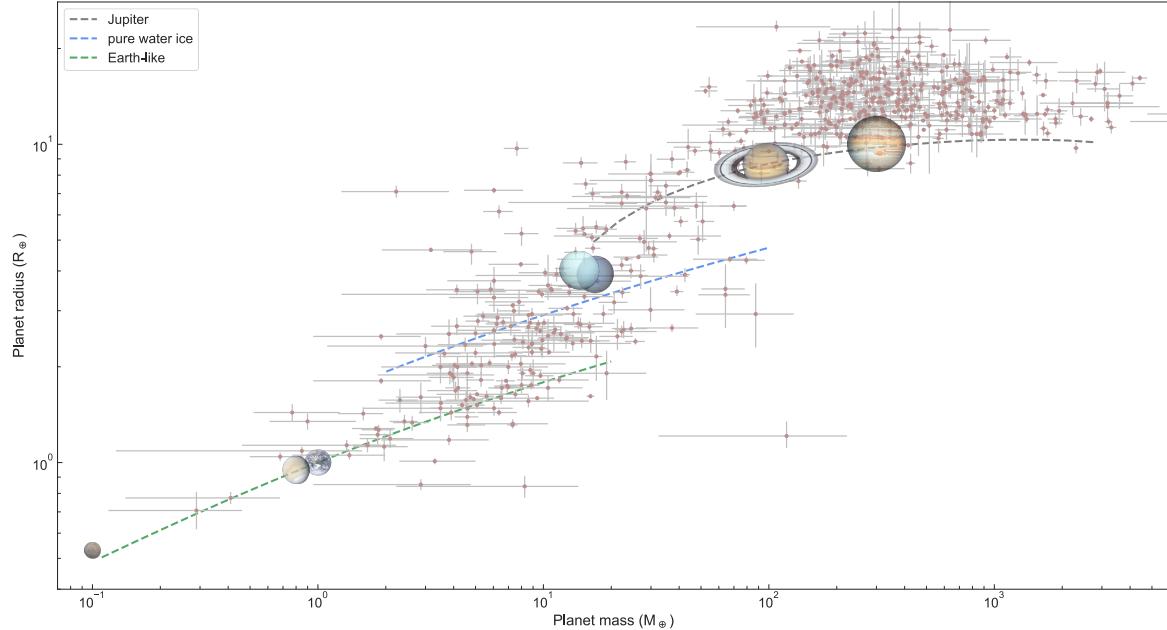


Figure 1.4: All known [exoplanets](#) with a measurement of their mass and radius. Hatched lines indicates model of bulk density for three different compositions.

transit through a disk during the formation epoch, meaning the process of planet formation need not to be efficient in converting solid material in the disk into planets, a fraction is enough.

Comparing further Solar angular momentum with the one of Jupiter one finds $J_{\odot} \sim 1\% J_J$. This means that a substantial lost of angular momentum and mass must have occurred during the star formation process.

Structure of orbit near mean-motion resonances ($P_1/P_2 \approx i/(i+j)$) is reminding us of dynamics interactions that may occur and trace back from early formation processes: Jupiter and Saturn are close to 5:2 mean-motion resonance; Pluto (and KBOs) are in 3:2 resonance with Neptune (see fig. [1.5 on the next page](#))

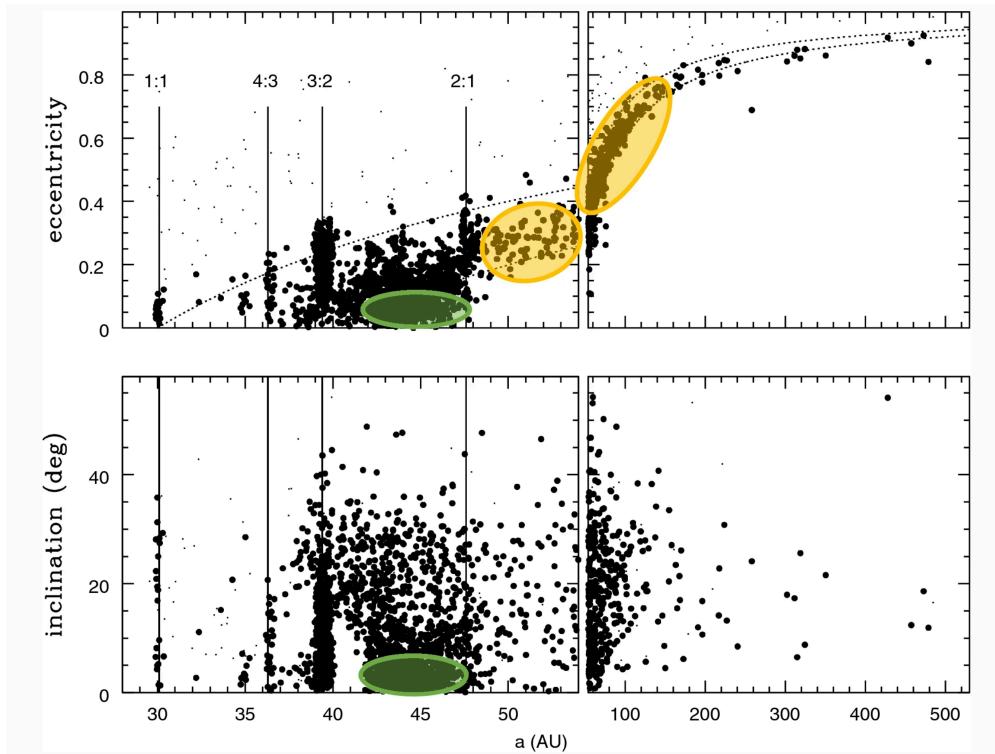


Figure 1.5: The orbital distribution of known Kuiper belt objects. The vertical lines mark the locations of some prominent mean motion resonances of Neptune. The cold classical and the scattered disk objects are indicated with green and orange ovals, respectively

Chapter 2

Planet Formation and Protoplanetary Disks

2.1 Introduction

The protoplanetary disk is a rotating disk of dense gas surrounding a young star. Planet formation mechanisms of planets are based on **core accretion** scenario. This mechanism is different from the cloud **core collapse** forming stellar object (and likely brown dwarfs). But more generally we cannot formally rule out a core collapse scenario could occurs in some heavy and cold disks.

Planet formation clocks start when the star reaches the hydrostatic equilibrium, i.e. the gravitational forces are balanced by gas pressure. The planet has to undergo a dramatic scaling, from sub-micrometer dust particles to multiple kilometre sized objects, equalling scale factor of 10^{10} - one of largest scale range to address in field of physics. In parallel the overall structure from which the planetary is formed experienced a reduction of a factor of 10^6 from large molecular cloud to planetary system scale (see Fig. 2.1 on the following page)

The protoplanetary disk is composed of mostly H₂ molecular gas (about 99%) and the rest of other molecules, usually referred to as "dust" or "grains". Earth and rocky planets in general have then to be formed out of these 1%. Refractory grains¹ of 0.1 μm size form solid are seed building blocks of "telluric" planets, while the molecular gas are the main ingredients of giant planets.

2.2 Gas in the disk

H₂ molecules are completely symmetric and have no permanent or magnetic dipole. They are only visible in the UV when pumped in a high transition level (fluorescence line). The mid-IR emission from rotational states, on the other hand, is difficult to see since hidden behind the IR emission from the dust. The molecular hydrogen is quasi a "dark matter" of

¹Refractory grains are considered primordial as they haven't been melt and be transformed

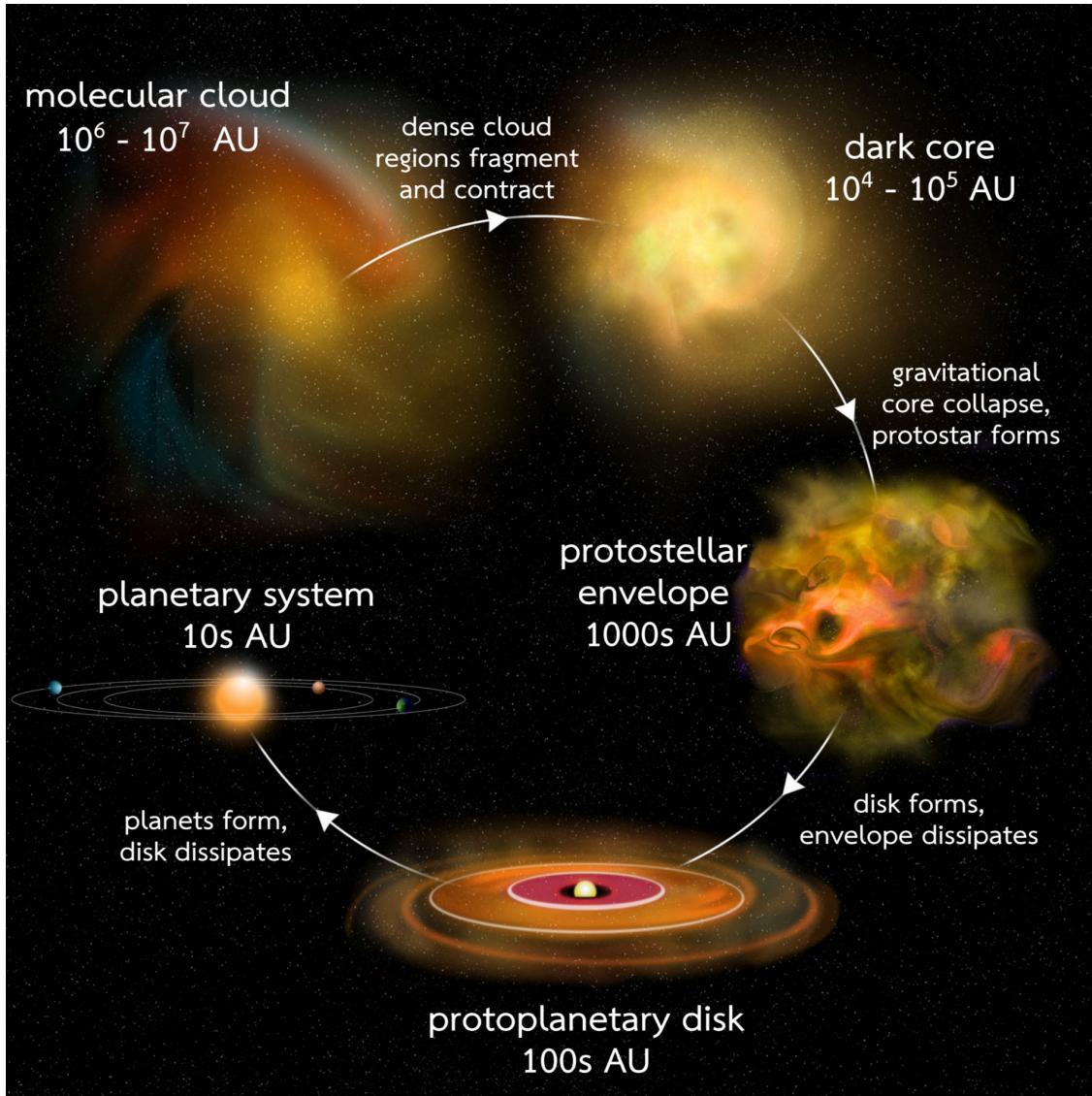


Figure 2.1: The big picture from star to planet ©Stolker 2018

planet formation. Using CO molecule as a probe is the mean alternative to measure the H₂ structure. CO gas has a strong electric dipole emitting at radio frequencies. It is frequently used to map the gas suing the following line transition:

¹²CO mean rotational states corresponds to 115 GHz ($J = 1 \rightarrow 0$), 230 GHz ($J = 2 \rightarrow 1$) and 345 GHz ($J = 3 \rightarrow 2$), corresponding to wavelengths of 2.6, 1.3 and 0.9 mm ².

²100 GHz is about 3000 μ m

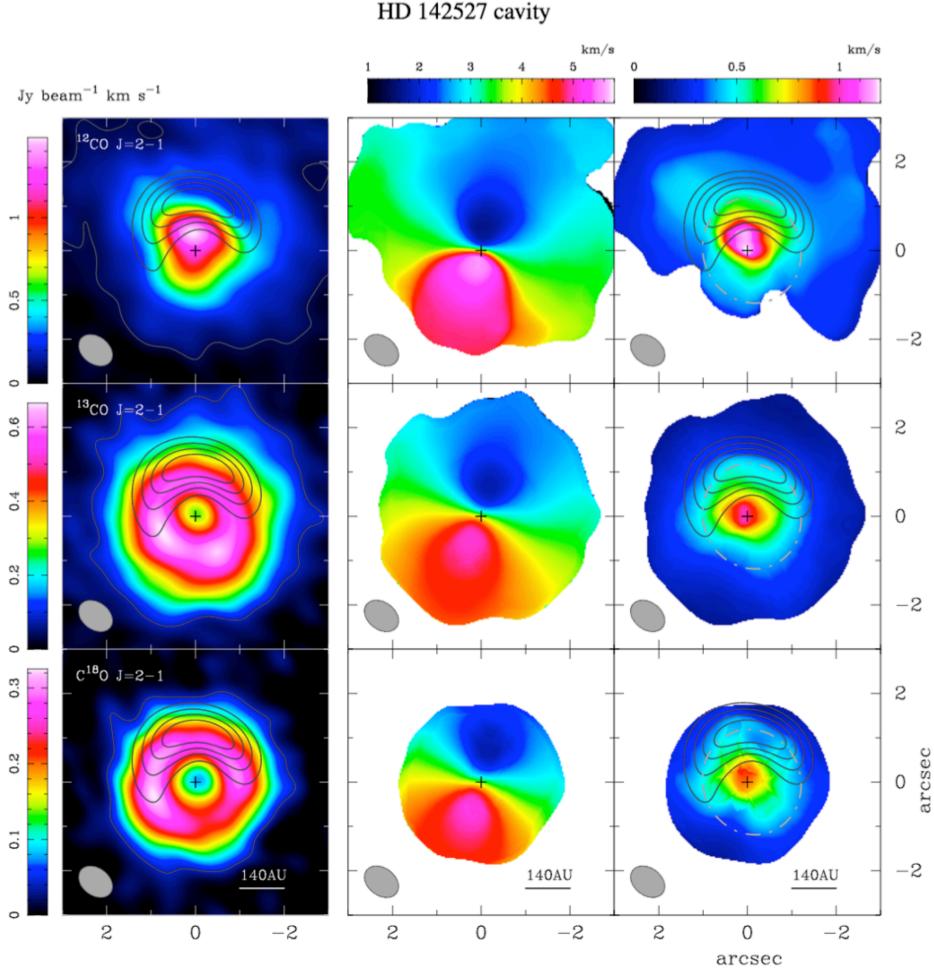


Figure 2.2: ALMA observation of the ^{12}CO .

H_2 is by far the main component of the disk gas. To measure the gaseous disk's structure one assumes a fixed abundance ratio between the two types of molecules. This assumption can lead to problems if the optical depth of the CO lines becomes too high that lines are becoming saturated. Further complication, CO freezes at 17 K (about 150 AU in the solar system), making it unsuitable for studying outer regions. Despite all these shortcomings, CO is useful for velocity mapping and temperature profiles of the disk.

Lastly, grains (or dusts) can be detected as spectral features, too. SiO for example has a broad $10 \mu m$ feature. Most of the grains are amorphous, i.e. disordered networks, but crystalline structures are observed in Silicons as well. Forsterite (Mg_2SiO_4) for example is one of them. It requires temperatures higher than 1000 K to be formed, a regime that is observed in the inner disk.

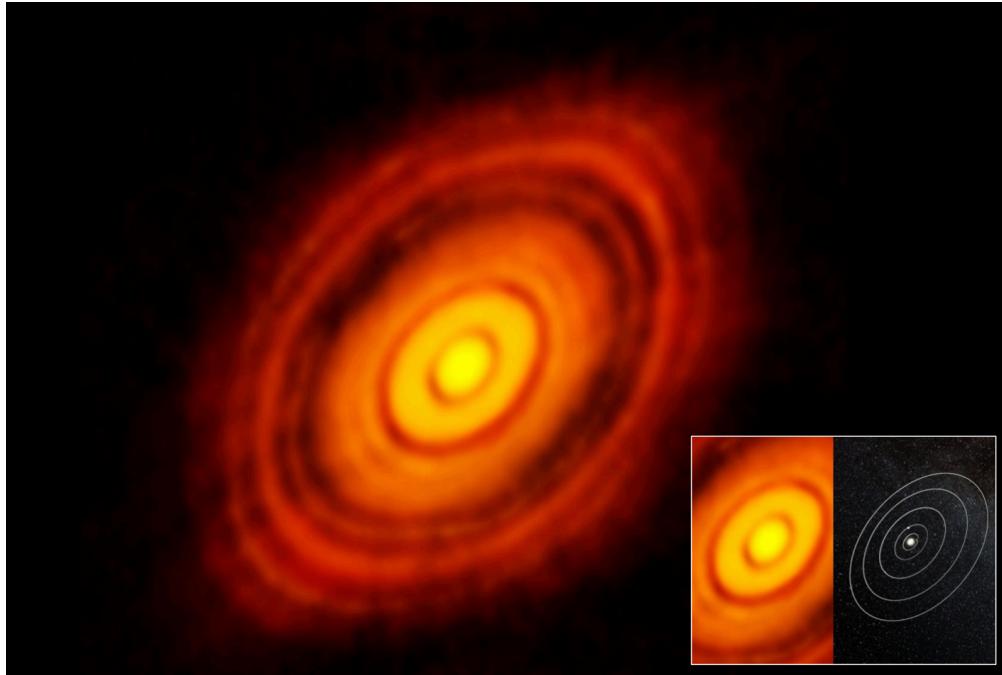


Figure 2.3: ALMA observation of the dusty disk of HR Tau. (Thermal signature)

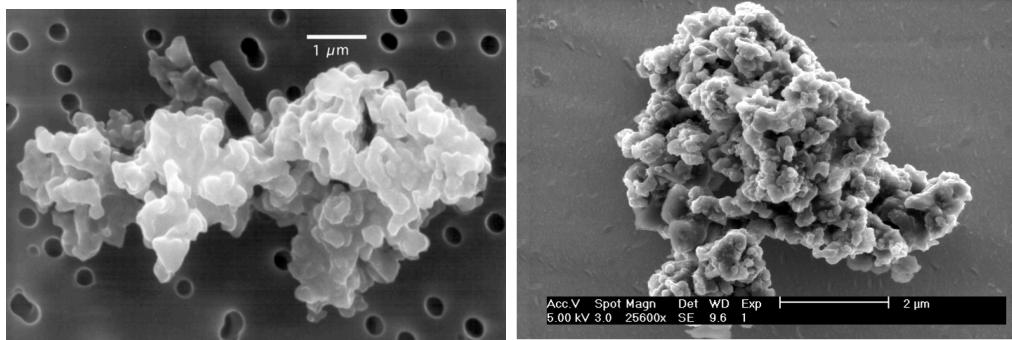


Figure 2.4: Interstellar "dust" grain

2.3 Spectral energy distribution

Practically the disk is seen from its dust content through thermal blackbody emission. The spectral energy distribution (SED) of the disk can be calculated as

$$F_\nu = \frac{\cos \vartheta}{D^2} \int_{r_{\text{in}}}^{r_{\text{out}}} B_\nu(T_d) (1 - e^{-\tau_\nu}) 2\pi r dr, \quad (2.1)$$

where D is the distance, ϑ is the inclination to the line of sight (viewing angle of the observer), $B_\nu(T_d)$ is the blackbody radiation in dependence of the disk temperature, $(1 - e^{-\tau_\nu})$ is a

term accounting for the transmission through the disk, τ_ν is the dust's optical depth, and the integral is performed from the inner to the outer radius of the disk (r_{in} and r_{out} , respectively).

In the case of an "optically thin" medium one can in first order approximate

$$1 - e^{-\tau_\nu} \approx \tau_\nu, \quad (2.2)$$

whereby (assuming iso-density)

$$\tau_\nu = \kappa_\nu \Sigma_d(r) / \cos \vartheta. \quad (2.3)$$

Thereby κ_ν is the dust opacity per mass unit and $\Sigma_d(r)$ is the mass surface density of the dust.

Consequently the SED is scaling as

$$F_\nu \sim \frac{\kappa_\nu}{D^2} B_\nu(T_d) M_d, \quad (2.4)$$

such that the SED can be directly used as a measure of the disk mass M_d .

The difficulty here lies with the fact that the disk is only optically thin when $\lambda \gg$ particle size, typically $\lambda \gg 100\mu\text{m}$, applying only to the outside regions.

2.4 Disk classification

From the observations of SEDs one can define different classes of disks (see fig. 2.5)

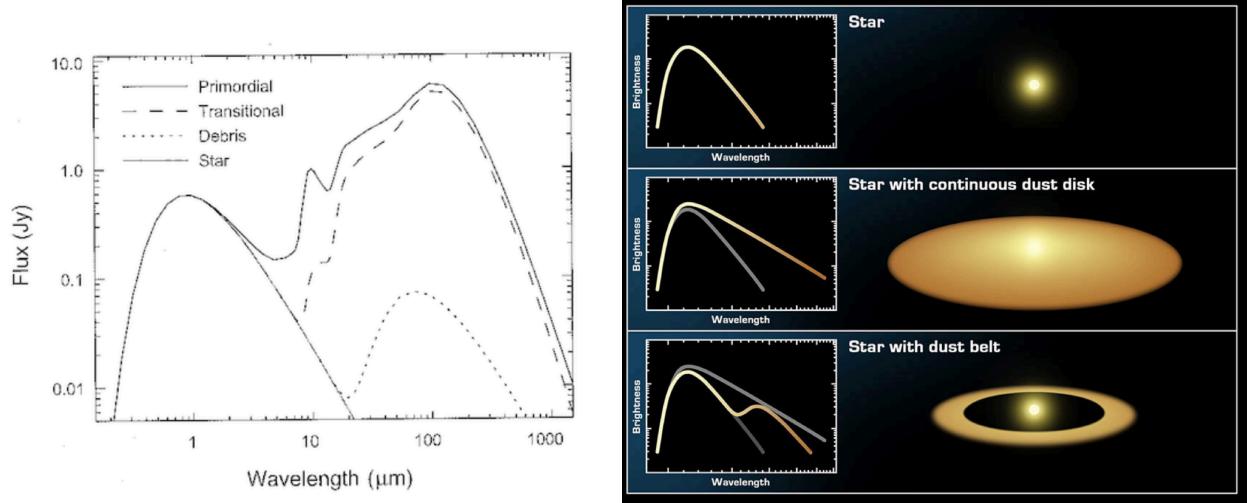


Figure 2.5: Spectral Energy Distribution (SED) of star with disk at different stage

Primordial disks typically appear around T Tauri stars and show strong sub-mm and CO emission. They typically have an optically thick disk and are very gas rich.

Transitional disks (around weak T Tauri stars) can be differed from primordial disks as they show no near-IR excess, meaning the inner regions are mainly cleared. The outer regions remain optically thick and sub-mm emission can be seen. One can see transitional disks as evolving from primordial disks to debris disks (see below), while the understanding of the "clearing out" of the inner parts remains still incomplete.

Debris disks can be found around main sequence stars. The Kuiper belt around our Sun is one example, accounting for zodiacal light. They are very poor on gas, however some gas is partially produced by icy comet-like evaporation. These disks radiate like blackbodies (with wavelengths above $10 \mu\text{m}$) in the cold outer regions.

Weak-line T-Tauri Stars (WTTS) are typical transitional disk. No gas is detected falling into the star, leading to the assumption that the gas is likely already removed from the inner disk. In the outer disk no CO is observed. The gas seems to dissipate in all places at the same time in only a few million years.

The detection of the accretion of gas depends on the L_{IR}/L_\odot detection threshold. The ratio of the 1 "zodi emission in our Solar System is about 10^7 compared the Sun. For comparison, currently, the most powerful IR space observatory (Spitzer satellite) can detect a threshold up to $L_{IR}/L_\odot \sim 1000$

Progress in IR imaging start to resolve scattered structure of stellar light coming from the surface layers of protoplanetary disks, including even planets (see fig. 2.6)

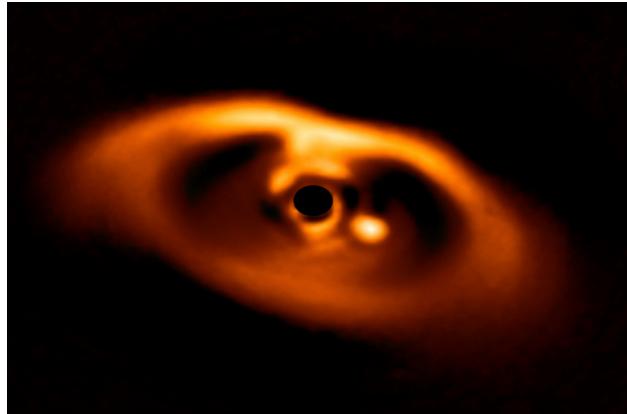


Figure 2.6: near IR image VLT/SPHERE proto-stellar disk PDS70 with a planet detection ([Muller et al.2018](#))

2.5 Disk structure

The disk model requires solving complex hydrodynamical problems involving dynamics, radiative transfer, chemistry and particle formation. To describe the structure of a disk of molecules orbiting a star one can, however, rely on two basic equations.

2.5.1 Hydrostatic equilibrium

In vertical direction one can describe the structure using the equation of **hydrostatic equilibrium**

$$\nabla P = -\rho g, \quad (2.5)$$

where P is the gas pressure, ρ is its density, and g is the gravitational acceleration. This assumes that the vertical component of the gravitational forces is balanced by the vertical gas pressure gradient. When the particle grow they loose their "fluid behaviour", the pressure support disappears, they decouple from the gas and settle down in the disk mid plane.

Using the hydrostatic equilibrium results in the equation

$$\frac{1}{\rho} \frac{\partial P}{\partial z} + \left(\frac{GM_*}{a^2} \right) \left(\frac{z}{a} \right) = 0, \quad (2.6)$$

where G is the gravitational constant, M_* is the mass of the star, a is the orbital distance. Thereby we assumed $z \ll a$, leading to $\sin \vartheta \sim \frac{z}{a}$.

Further assuming an ideal gas with $P = c^2 \rho$, c being the gas' speed, one finds

$$\frac{1}{\rho} \frac{\partial P}{\partial z} = c^2 \frac{1}{\rho} \frac{\partial \rho}{\partial z}. \quad (2.7)$$

We now consider the density distribution to follow a Gaussian profile whose width is defined as the scale height H , such that $\rho \sim \exp(-\frac{z^2}{2H^2})$. Furthermore, we assume that the star's gravity dominates. This then leads to the expression

$$c \sim H \cdot \Omega, \quad (2.8)$$

with the Keplerian angular velocity

$$\Omega^2 = \frac{GM_*}{a^3}. \quad (2.9)$$

For an ideal gas with $c^2 \sim kT/\mu m_p$ the scale height can then be expressed as

$$H^2 \sim \frac{T}{\mu m_p} \cdot a^3, \quad (2.10)$$

where T is the temperature of the gas, μ is the mean molecular weight of the gas and m_p is the proton mass. This directly shows that cold disks are in general thinner than hot disks, and that outer disk regions are thicker than inner disk regions (assuming T is constant).

2.5.2 Momentum conservation

In the radial direction, assuming a steady flow and a Keplerian-disk, the momentum conservation equation leads to

$$-\Omega r + \frac{1}{\rho_g} \frac{\partial P}{\partial r} + \frac{GM_*}{r^2} = 0, \quad (2.11)$$

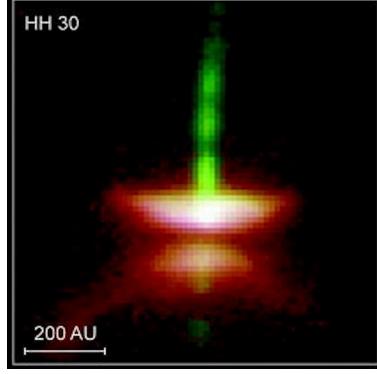


Figure 2.7: Reflected light of protoplanetary disk seen edge-on measured with HST

where v_r denotes the radial velocity.

Using the WKB approximation from perturbation theory,

$$P = P_0 + \partial P, \quad (2.12)$$

one finds a set of equations of the form

$$\partial P \sim \exp i(\omega t \pm 2\pi r/\lambda). \quad (2.13)$$

Solving the equations leads to

$$\omega^2 = 4\Omega^2 - \frac{4\pi^2 G \Sigma}{\lambda} + \frac{4\pi^2 c^2}{\lambda^2}, \quad (2.14)$$

$$\omega^2 = 4\Omega^2 - 2\pi G \Sigma k + c^2 k \quad (2.15)$$

With our strict assumption of Keplerian disk, any perturbed fluid element follows a elliptical orbit (central gravitational force dominates). The right hand side of Equation 2.15 describes the nature of waves in the disk. As long as this term is positive the term in the exponential of the Eq. 2.13 is complex and perturbation waves (∂P) oscillate and propagate radially. But when the right and side become negative, $\omega^2 < 0$, radial perturbation waves grow exponentially ($\partial P \sim \exp +\omega t$).

Examining each term reveals their respective physical meaning: when λ gets large $\omega^2 = \Omega^2$ reflecting central gravitational force dominates and stabilizes the disk. For short wavelength radial perturbation, $\lambda \rightarrow 0$, the pressure term $\sim c^2/\lambda$ dominates and a stable state is reached as well (right and side remains positive). The problem of instability originates from the negative value of the intermediate term lying between these two case (see on Fig. 2.8 on the next page)

The condition for marginal stability is that $\omega^2 \geq 0$ at all spatial scales. The most unstable scale k_{crit} can be found when $d\omega^2/dk = 0$, which yields, $k_{crit} = \pi G \Sigma$.

One can see when solving the equation for $\omega = 0$ a solution for any scale λ would implicit mean a perturbation can grow when: $c\Omega < \pi G \Sigma$.

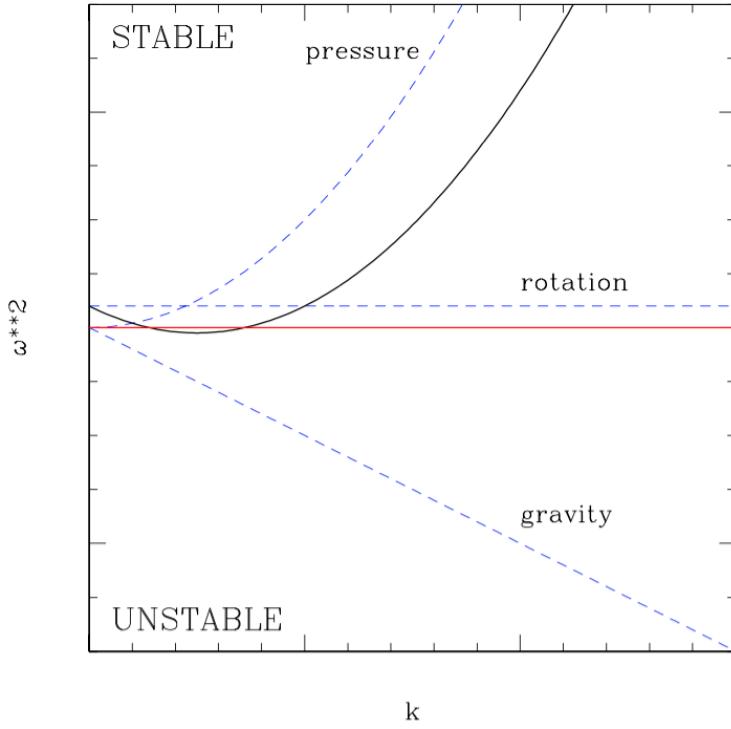


Figure 2.8: The dispersion relation (solid black line) for a uniformly rotating disk, illustrating the contributions from pressure, rotation, and self-gravity (dashed blue lines). The system is unstable if, at any value of the wavenumber k , w^2 falls below the red line and is negative. (from Armitage 2020).

To express it in simple form one usually defines the Toomre parameter and its corresponding criteria

$$Q_{cr} \equiv \frac{c\Omega}{\pi G \Sigma} \sim \frac{\sqrt{T}}{\Sigma} \Omega., \quad (2.16)$$

It is the custom to express the condition in the opposite way as a *Toomre criteria* for a stable disk where no radial perturbation can grow $Q_{cr} > 1$. From this one sees that gravitational instability are more likely to happen when disks are dynamically cold (low velocity dispersion) and dense (high surface density). Using typical parameters for protoplanetary disks: $a \approx 5$ AU, $H/a \approx 5\%$, and $T \approx 75$ K one can rewrite Toomre criteria in terms of surface density only

$$Q_{cr} \approx 3 \cdot 10^3 \Sigma^{-1} \text{ g cm}^{-2}. \quad (2.17)$$

In the specific assumption of the Minimum Mass Solar Nebulae (MMSN) (see in section 2.6.5 on page 27) a maximal surface density $\Sigma \approx 10^3 \text{ g cm}^{-2}$ at Jupiter's distance from the Sun is considered. This leads to the conclusion that Solar System at the this stage was stable against Toomre criteria $Q_{cr} > 1$.

Assuming these approximation can be extended to all systems one concludes that planets cannot be formed by **core collapse** due to gravitational instability. It would have required a more massive and cold disk. Some authors however suggest massive planets in the outer edge (cooler temperature) of planetary systems may formed that way. But this scenario cannot explain a mixed population of planets, like our Solar System

Recently, the occurrence of streaming instabilities clumping in the disk that may subsequently collapse under self-gravity increase the complexity of disk structure and the possibility of local effects. Numerical simulations have established that the streaming instability provides a pathway to forming dense clumps that can collapse gravitationally to form planetesimals. See on Fig. 2.9, though the pre- requisite particle size and local metallicity are not trivially satisfied (Armitage 2020).

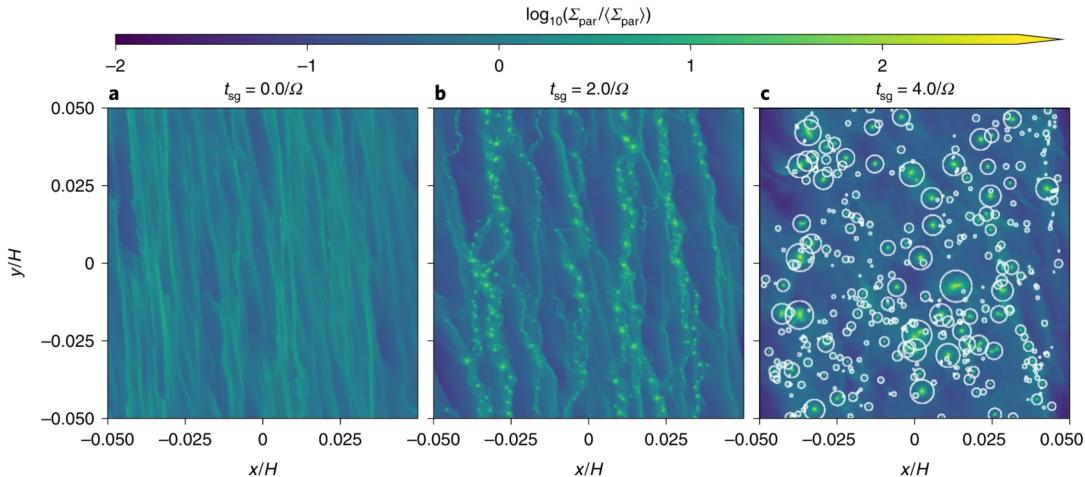


Figure 2.9: The plots show the vertically integrated density of solids, projected on the disk plane, relative to the initially uniform surface density . The x and y coordinates show the shearing-box dimensions in units of the gas disk scale height, H (the Sun is to the left; the orbital velocity vector points up). Time t increases from left to right as labelled (t_{sg}). The circles in c depict the Hill spheres of clumps (Nesvorný 2019)

In any case gaseous disks eventually disappear by viscous accretion and photo-evaporation. This mechanism leads to a "observational" and chronological sequence detailed in the next chapter.

2.6 Planet formation

This chapter describes unified theory established to describe the formation of all types of planets: **core accretion**. It is mostly built to account for the solar system case but could apply in a general manner to exo-planet as well. We do not consider planet formation by **core collapse** caused by gravitational instability in the protoplanetary disk.

2.6.1 The core accretion model

The main steps of the core accretion model is a sequence of events enabling the growth - by stages- of a planets. At the same time the protoplanetary disk evolves and disappears (see Fig. 2.10). From beginning to end it is a considerable large upscaling process from 10^{-6} to 10^8 meter, larger in scale factor by comparison to downscale collapse of molecular cloud to form a star (from 10^{18} to 10^9 m)

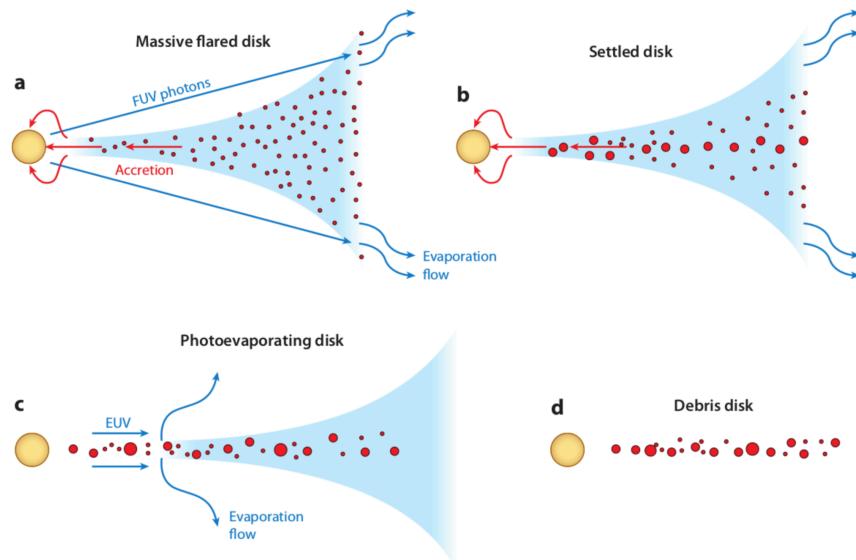


Figure 2.10: Four stages of planetary formation related to the disappearance of the protoplanetary disk ([©Williams and Cieza, 2011](#))

In the disk the following sequence of events takes place:

- Solids microscopic particles ("usually called dust") assemble into grains (larger than μ -size)
- Particles dynamically start decoupling from the gas dynamics and settle on the mid-plane
- Solids grains assembled by soft-collisions (molecular gluing) until they reach meter size. From hundred meters, when they are fully decoupled from the disk drag effect they

are called *planetesimals*

- Swarm of planetesimals grow by collisions and form planet embryos of few hundreds km
- Planet embryos attract particles and form planets or giant planet cores
- Planet core attract gas (if available) and form giant planets

The growths from kilometre sized objects to planet cores is rather well understood as a gravity effect leading to accretion. On smaller scale the process that efficiently "glues" grains together to form the planetesimals, is still an open question and currently not well understood.

The initial phases to make gas giants and rocky planets are following about the same scenario; they are controlled by accumulation of solid material, i.e. dust and small grains. Then without gas around rocky planets only grow further by collisions with other solid bodies. Giant planet cores keep their growth by accreting a large gas envelope.

The final outcome depends on the amount of gas the planet core is able to capture from the disk before it vanishes. But that is not that "easy" for the planet. The gas accreted to a planet core then forms an envelope (usually called atmosphere). When the atmosphere becomes radiatively thick gas kinetic energy increases and slows down the mass build-up. At some point gravitation becomes strong enough to take over and all mass around is accreted (runaway effect). Eventually the growth stops when no more gas is available to capture.

2.6.2 Decoupling of gas and grains - looking at the disk plane

We can understand the different influence on gas and particles in the disk by going back to the radial momentum conservation equation in the disk that we have already studied in section [2.5.2 on page 13](#):

$$-r\Omega^2 + \frac{1}{\rho} \frac{\partial P}{\partial r} + \frac{GM}{r^2} = 0 \quad (2.18)$$

When we take a closer look at the pressure term for gas we get $\frac{\partial P}{\partial r} < 0$, since the pressure decreases when moving radially to the outside of the disk. For particles (e.g. grains in the disk), on the other hand, we get $\frac{\partial P}{\partial r} = 0$ due to their incompressibility. If we rearrange the equation

$$r\Omega^2 = \frac{1}{\rho} \frac{\partial P}{\partial r} + \frac{GM}{r^2}, \quad (2.19)$$

one can directly see that due to the pressure term $\Omega_{gas} < \Omega_{particle}$, meaning the orbital velocity of the gas is smaller than that of the particles in the same orbit. The gas is therefore said to be "sub-Keplerian" and we speak of a **decoupling** between gas and particles behaviour. This results in friction on the particle caused by the slower gas ("gas drag"). In other terms each particle suffers a "head-wind", loses angular momentum and drifts towards the star. This phenomenon is known as the cm-size problem and suggests non-linear effects local effect, preventing this happening and all particles to disappear into the star. As soon as the particles grow bigger than meter size the head-wind drag becomes negligible. In reality the scenario is complex, still not understood, and includes loss of non-linear turbulent effects.

2.6.3 From dust to planetesimals - looking into the vertical direction

While we can assume that the vertical motion of the gas is negligible, particles in the disk are affected by vertical friction as well and hence also move out of the disk plain. This can be described by

$$\frac{\partial^2 z}{\partial t^2} + \frac{1}{\tau_f} \frac{\partial z}{\partial t} = -\frac{GM}{a^2} \cdot (z/a), \quad (2.20)$$

where z is the distance over the disk mid plane, τ_f is the friction time-scale. The second term corresponds to the frictional force exerted by the gas, while the third term results from the vertical gravitational acceleration.

Expressed in terms of Ω (the angular velocity of the particle around the star), Keplerien orbit ($a\Omega^2 = GMa^{-2}$), we then find the relation

$$\frac{\partial^2 z}{\partial t^2} + \frac{1}{\tau_f} \frac{\partial z}{\partial t} + z\Omega^2 = 0. \quad (2.21)$$

This equation represents a typical damped oscillator.

For $\tau_f \rightarrow 0$, meaning that gas and particles are perfectly coupled, $\frac{dz}{dt}$ equals the vertical gas velocity and is therefore zero, so that no sedimentation can happen.

For $\tau_f \rightarrow \infty$ the term becomes friction-less and results in an oscillation of the particles around the disk mid plane with $P = 2\pi/\Omega$.

As established by Epstein (1924) the friction time for solid spheric particles bathed in a gas can be modelled by

$$\tau_f = \frac{\rho_{sp}}{\rho} \frac{R_{sp}}{c}, \quad (2.22)$$

where ρ_{sp} and R_{sp} are i density and radius of solid spheric particles, and ρ and c respectively the density and typical speed of the gas where they move and suffer a friction effect.

In see section 2.5.1 on page 13 from the condition of the hydrostatic equilibrium in vertical direction, with the assumption of an ideal gas with $p = c^2\rho$ and gaussien profile of height H , one derived $c = H\Omega$. When Replacing c with this relation into Eq. (2.22) we find

$$\tau_f = \frac{\rho_{sp} R_{sp}}{\rho H \Omega}, \quad (2.23)$$

$$(2.24)$$

leading to a derive vertical damping timescale

$$\Omega\tau_f = (a/H) \frac{1}{a\rho} \rho_{sp} R_{sp}. \quad (2.25)$$

Assuming typical protoplanetary disk values of $(A/H) \approx 5\%$ and $\rho \approx 10^{-10} \text{ g cm}^{-3}$ for $a \approx 5 \text{ AU}$ one finds:

$$\Omega\tau_f \sim 10^{-3}\rho_{sp}R_{sp} \quad (2.26)$$

We can solve differential Eq. 2.21 on the preceding page to compute the envelop of the decay for small value of $\Omega\tau_f$:

$$z \sim z_0 e^{-t\Omega^2\tau_f} \sim z_0 e^{-t\Omega(\Omega\tau_f)}. \quad (2.27)$$

We see that to reach 1 % of the initial distance from the mid plane z_0 , i.e. $z \approx z_0 e^{-5}$, one needs to satisfy the relation

$$\Omega t > \frac{5}{\Omega\tau_f} \sim \frac{5 \cdot 10^3}{\rho_{sp}R_{sp}}. \quad (2.28)$$

with a typical particle values would be around $\rho_{sp} \sim 1 \text{ g cm}^{-3}$ (about a snow flake) and $R \sim 1 \mu\text{m}$, one derives a typical time-scale of

$$t > \frac{10^6}{\Omega}, \quad (2.29)$$

After typically a million orbital periods (10^7 yrs at Jupiter distance) most of grains are settled on the mid-plan. The time is even shorter when the particle gets bigger. The settling and the decoupling are appendng together. The growth from dust to planetesimals is poorly understood and is topic of laboratory experiments on smaller scales. According to current models the best growth rates can be achieved given a low collision speed, a high dissipation and a high binding energy. (See video)

2.6.4 From planetesimals to planet cores

The foundation of modern planet formation theory started with the legacy work of the Russian mathematician Ott Schmidt and his team known through the monumental treatise published in 1969 by his pupil Victor Safronov. In this publication they were the first to come up with an explanation of planetesimal built up mechanism.

In the so-call Safronov Theory planetesimals are rocky or icy objects with a size of about 1 km and masses in the range of 10^{12} kg and more³. At the beginning there is a large number of planetesimals, interacting and colliding with the assumption that they dynamical interactions lead to merging rather than fragmentation, and consequently a mass enhancement of

$$\frac{dM}{dt} = \pi R^2 v_{rel} \rho_s F_g, \quad (2.30)$$

³In comparison, Cera has a size of 500km and mass of 10^{20} kg ; Earth is 10^{24} kg

where R is the radius of the accreting object, v_{rel} is the relative velocity between the colliding objects, ρ_s is the volume density of solid material in the disk.⁴ The gravitational enhancement factor

$$F_g = \frac{\pi R_{eff}^2}{\pi R^2} \quad (2.31)$$

is determined by the ratio of the effective cross section (due to gravitational attraction) of the planetesimal to its geometrical cross-section. The factor $v_{rel}\rho_s$ describes a mass flux, being a measure of the material hitting the target per unit time and unit area.

Assuming v_{rel} between the objects are isotropic we can express $\Omega_{rel} = \Omega \sin \vartheta$ (see fig. 2.11).

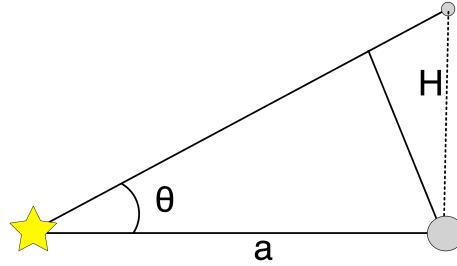


Figure 2.11: Geometry of the configuration

With the usual small angle approximation $\sin \vartheta \sim H_s/a$ following from geometry (H_s being the vertical thickness of the planetesimal disk) we get⁵

$$v_{rel} = \Omega H_s, \quad (2.32)$$

and with the surface density of the planetesimal disk (of solids) $\Sigma_s \sim H_s \rho_s$, we can express Eq. (2.30 on the facing page) as

$$\frac{dM}{dt} = \pi R^2 \Omega \Sigma_s F_g. \quad (2.33)$$

Assuming $\Sigma_s F_g \approx \text{const.}$ we find

$$\frac{dM}{dt} \sim \Omega \sim a^{-3/2}, \quad (2.34)$$

indicating that the accretion rate is slower the further the planetesimal is from the star.

If we want to compare the accretion rates of Jupiter and Uranus (assuming they formed at their current orbital distance) we get a ratio of

$$\left(\frac{a(\text{Jup})}{a(\text{Ur})} \right)^{-3/2} \approx \left(\frac{5}{19} \right)^{-3/2} \approx 7, \quad (2.35)$$

⁴Note that R has to be replaced by $R_1 + R_2$ in case the radius of the accreted object is not negligible.

⁵Note the similarity with $c = \Omega H$ for the gas.

expressing that a planetesimal at Jupiter's location has a seven times higher accretion rate than one at Uranus position.

Neglecting all the surroundings and assuming a two body problem where the accreting object is significantly more massive (e.g. a planet embryo "swallowing" a planetesimal), one can assume that the accreting body is located at the center of mass of the system and translate into its rest frame. With this assumption the conservation of the angular momentum leads to

$$R_{eff}v_{rel} = Rv_T, \quad (2.36)$$

where v_T is the relative tangential velocity (i.e. the planetesimal grazing the embryo). The size of the embryo is neglected).

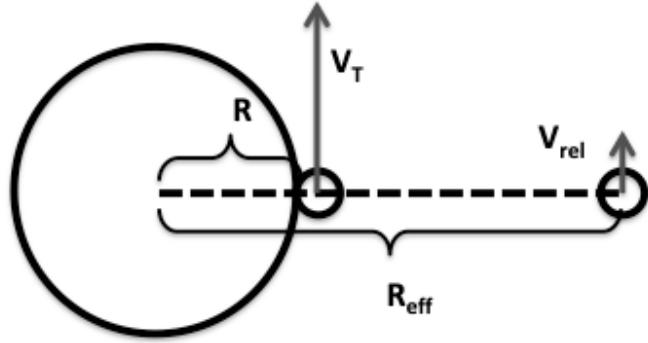


Figure 2.12: Schematics of the capture of the planetesimal by the embryo

We can use this to express the gravitational enhancement factor as

$$F_g = \frac{\pi R_{eff}^2}{\pi R^2} = \frac{v_T^2}{v_{rel}^2}. \quad (2.37)$$

Further, energy conservation gives us

$$v_{rel}^2 = v_T^2 - \frac{2GM}{R} = v_T^2 - v_{esc}^2, \quad (2.38)$$

with v_{esc} being the escape velocity from the surface of the accreting object (the potential energy at R_{eff} is considered as negligible). Now we can further express the gravitational enhancement factor in dependency of the relative and escape velocities as

$$F_g = 1 + \frac{v_{esc}^2}{v_{rel}^2}. \quad (2.39)$$

We directly see that for $v_{rel} \gg v_{esc}$, i.e. high radial velocities, $F_g \approx 1$, meaning that the effective cross section of the accreting object equals its geometric cross section. The growth rate is then given by

$$\frac{dM}{dt} \sim R^2 \sim M^{2/3}. \quad (2.40)$$

In the opposite assumption $v_{rel} \ll v_{esc}$, $F_g \approx \frac{v_{esc}^2}{v_{rel}^2}$ (with $(v_{esc}^2 \sim M/R \sim \rho R^3/R)$, one sees that $F_g \sim R^2$, meaning the growth rate is then

$$\frac{dM}{dt} \sim R^4 \sim M^{4/3}. \quad (2.41)$$

The growth time-scale (expressed in exponential terms) may be described by

$$\tau_s = M \left(\frac{dM}{dt} \right)^{-1}. \quad (2.42)$$

When looking at the two cases in more detail, one sees a different time-scale behaviour:

When $v_{rel} \gg v_{esc}$, $\tau_s \sim M^{1/3}$, meaning when the mass increases, rate of growth (τ^{-1}) slows down.

When $v_{rel} \ll v_{esc}$, $\tau_s \sim M^{-1/3}$, meaning when the mass increase, rate of growth rate increase as well. It corresponding to a *runaway* situation often call *oligarchic groth*.

Numerical N-body simulations displayed in fig. 2.13 on the next page similar behaviour and can produce planet core similar to the rocky planets of our solar system in about 200 Myr.

The mechanism suggests the relative velocity plays a crucial role in the efficiency of accretion rate and hence the formation of planets. Another way to see this is the concept of Hill radius applicable to giant planet accretions. It is discussed in the next chapter.

When gravitational effect described above is combined with aerodynamic forces during their encounter we have an aerodynamically assisted regime or "pebble accretion". This mechanism is particularly efficient when the accreted body (the pebble) experienced aerodynamic forces leading to enhanced accretion rate. It is a qualitatively distinct mechanism to growth planetesimals. In some case it can be more efficient than planetesimal accretion 'Safronov type'. This mechanism is not treated in the lecture but Orwell 2017 produced a comprehensive review. An visual illustration of various possible accretion mechanisms are displayed in fig. 2.14 on the following page.

A relevant parameter to describe the outcome of an impact is the specific energy $Q_D \equiv mv^2(2M)^{-1}$ with the impactor m having a speed v hitting the target M . On fig. 2.15 on page 25 The minimum typical specific energy for a catastrophic disruption of solid bodies (Assuming different material: icy or unfractured rocky) is plotted as a function of the bodys radius. With thsi model, practically the most vulnerable bodies are generally those with radii in the 100m to 1km with $Q_D = 10^5 - 10^6$. For comparison asteroids in main belt have typical velocities of the order of 2 km s^{-1} . For a mass ratio $m/M = 0.1$ the specific energy of the collision is then around $Q_D = 2 \times 10^9 \text{ erg s}^{-1}$, which would easily destroy large solid bodies up to 100 km.

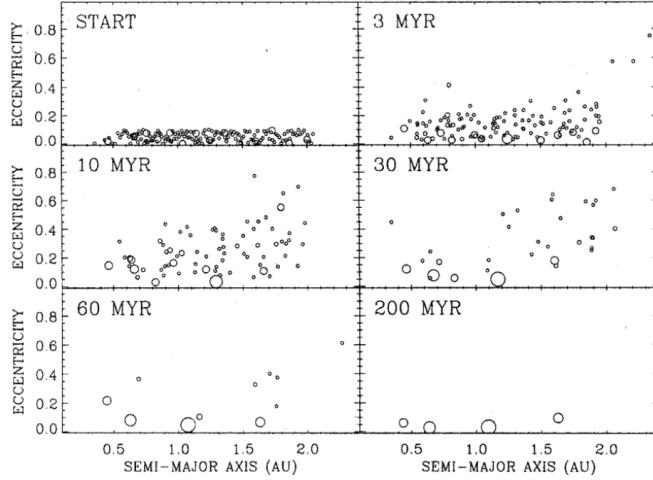


Figure 2.13: N-body Simulation of formation of planetary cores by gravitational interactions

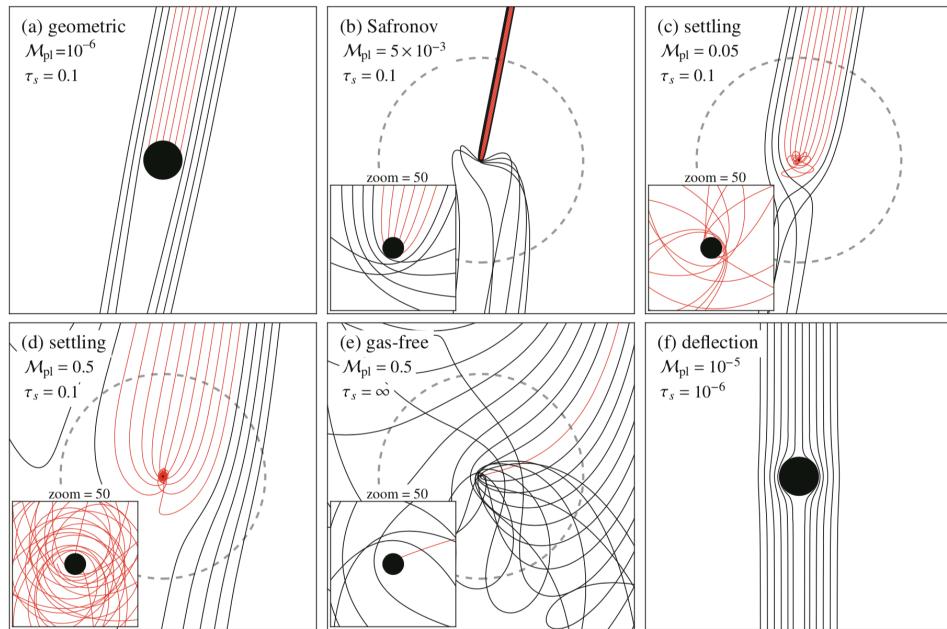


Figure 2.14: Examples of planetesimals accretion, viewed in the frame co-moving with the planet. In each panel the filled circle denotes the physical size of the planet and the dashed circle its Hill sphere. (c) and (d) qualify as pebble accretion, while (a), (b), and (e) fall in the ballistic regime. In (f) particles are so small that they follow gas streamlines (Orwell 2017)

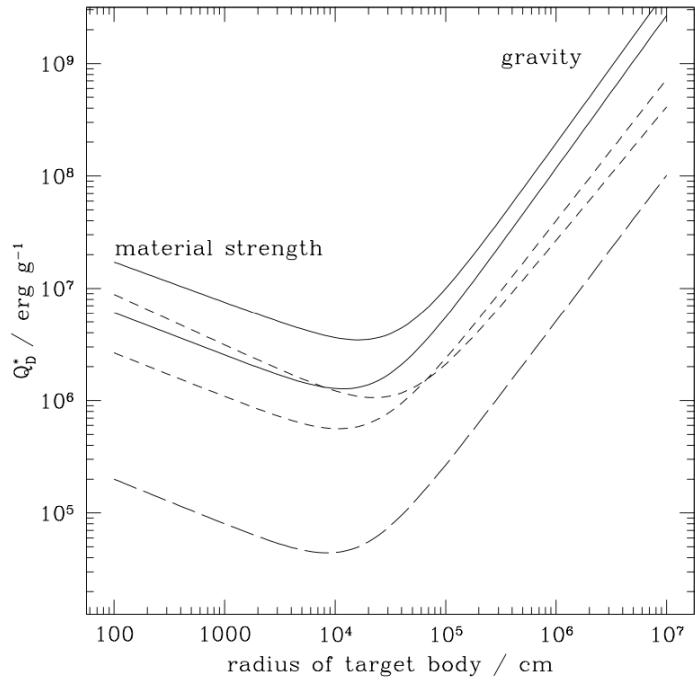


Figure 2.15:

Hill radius

In the solution of the "restricted three body problem (one object gravitation dominates the two others), there are five stable points (or Lagrange stability point) where a third object can be placed in a scenario of a smaller object (planet) orbiting a larger object (star). The **Hill radius** defines the distance of the second Lagrange point L_1 from the orbiting object (see fig. 2.16 on the following page).

The Lagrangian equation:

$$G \frac{M_s}{R_H^2} - G \frac{M_*}{(a - R_H)^2} + \Omega^2(a - R_H) = 0, \quad (2.43)$$

can be solved (with $\Omega = (GM_s)^{-1/2}a^{-3/2}$) to obtain the definition of the Hill radius

$$R_H = a \left(\frac{M_s}{3M_*} \right)^{1/3}. \quad (2.44)$$

When planet embryo (M_s) orbits a star (M_*) and planetesimal are found within its *Hill radius* the accretion is velocity-shear dominated rather than dictated by random velocity of

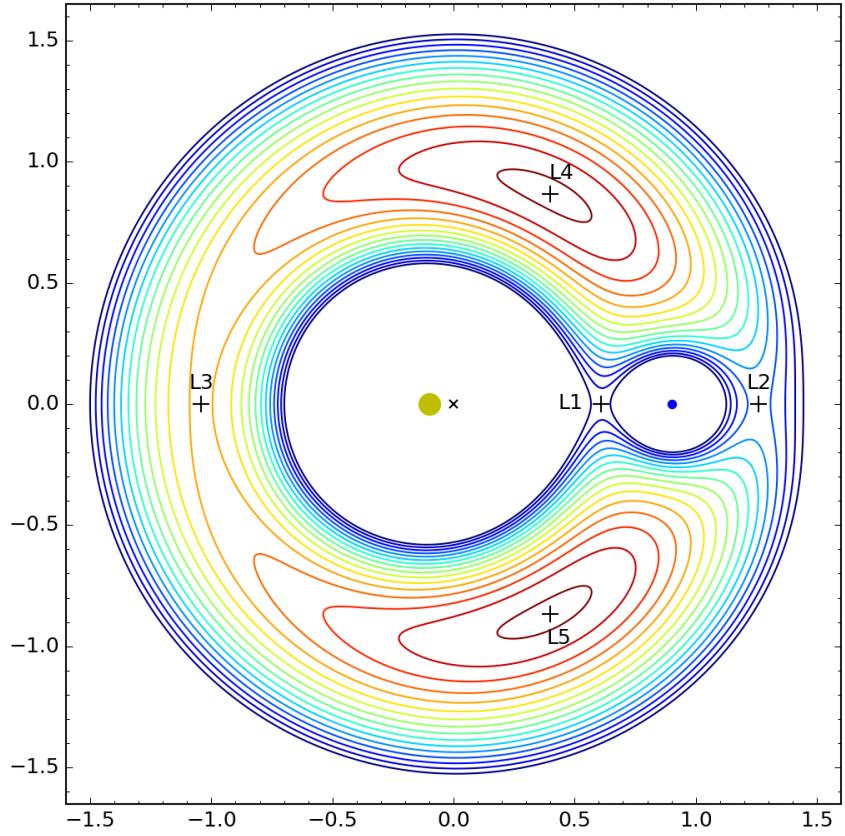


Figure 2.16: Display of the five Lagrangian points in a rotating potential reference frame work.

planetesimals. The Hill radius allows to estimate the regime in which the gravitation of the accreting object dominates over any influence from the star or centrifugal effects.⁶

The relative velocity between the accreting and accreted object on a circular orbit can be expressed as

$$v_{rel} = |(a + \Delta a)\Omega(a + \Delta a) - a\Omega(a)|, \quad (2.45)$$

When expanding $\Omega(a + \Delta a)$ to the first order one easily finds:

$$|v_{rel}| \approx \frac{1}{2}\Omega\Delta a. \quad (2.46)$$

⁶Note that the Hill radius does not determine a sphere. The total volume of this regime equals only roughly 1/3 of the one a sphere with the radius R_H would have.

When expressing Ω in terms of the Hill radius by substituting a ,

$$\Omega^2 = \frac{GM_*}{a^3}, \quad (2.47)$$

$$= \frac{GM_s}{3R_H^3}. \quad (2.48)$$

The one can write v_{esc}^2/v_{rel}^2 considering the escape velocity ($v_{esc}^2 = 2GM_s R^{-1}$) in terms of R_H and Δa , using Equ. 2.48 and Equ. 2.46 on the facing page and one finds

$$\frac{v_{esc}^2}{v_{rel}^2} \sim \frac{R_H^3}{R(\Delta a)^2} = \left(\frac{R_H}{\Delta a} \right)^2 \frac{R_H}{R}. \quad (2.49)$$

We directly see that R_H acts as a scaling factor controlling the growth rate regime of the embryo.

It is important to note that the growth scenario in reality is by far more complex and dominated by non-linear and stochastic processes. Turbulent effects, for example, can lead to a bicyclic accretion orbit of the smaller particles, leading to a larger influence zone⁷. More advanced studies of these scenarios have therefore to rely on computational N-body simulations.

Note that "Peeble accretion" mentioned earlier combines self-accretion of planetesimal with disk dynamic friction. Hill radius remains a valid scaling reference

2.6.5 Isolation of planetary cores

The oligarchic growth continues until its feeding zone has been cleared of planetesimals. Current N-body simulations show that this feeding zone can reach a size of up to around four times the Hill radius. In general one can write it as

$$\Delta a = b \cdot R_H. \quad (2.50)$$

If the planetary embryo has reached this state, its mass equals the so called isolation mass

$$M_{iso} \sim 4\pi a \Delta a \Sigma_s \quad (2.51)$$

$$\sim 4\pi ab R_H \Sigma_s \quad (2.52)$$

$$\sim \sqrt{\frac{(4\pi a^2 b \Sigma_s)^3}{3M_*}}, \quad (2.53)$$

with Σ_s describing the initial value of the surface density.

⁷Note that in calculating the Hill Radius we relied on several assumptions and neglected any turbulent effects

Assuming $b = 4$, an isolated planetary core of $11M_{\oplus}$ at $a = 5.2\text{AU}$, one would expect an initial surface density of

$$\Sigma_s \approx 10 \frac{\text{g}}{\text{cm}^2}. \quad (2.54)$$

The use of $11M_{\oplus}$ as a reference for the core of giant planets comes from estimated values based on internal models of Jupiter and Saturn (see fig. 2.17)

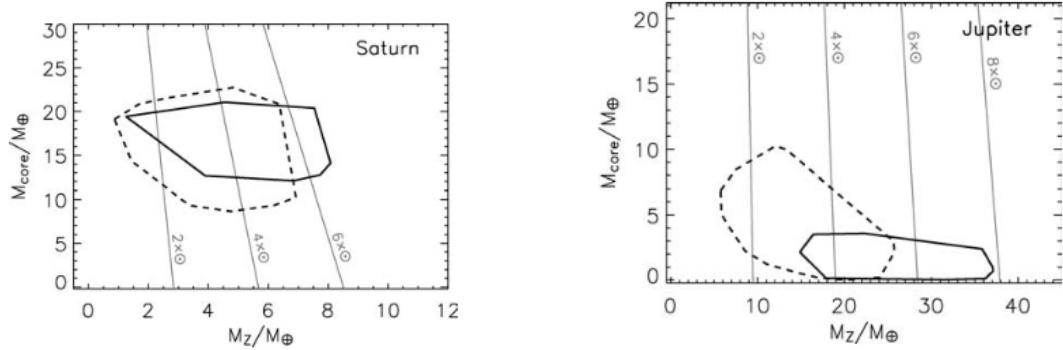


Figure 2.17: Current combined measurements and model of the composition of the core of giant planets Jupiter and Saturn

To understand the significance of derived value for the surface density one compares with the predictions for a minimum mass solar nebula (MMSN). The MMSN is a crude idea (originally by Weidenschilling 1977 and further develop by Hayashi) to estimate how the gas and dust that formed the Solar System was distributed with the distance from the Sun. It is computed by "spreading" the mass of each planet into annulus extended half-way through orbit gravitation range of neighbouring planets.

By definition MMSN is a minimum estimation considering a some fraction of the disk material didn't end up captured into planet. $\Sigma = \Sigma_{r,r+i,g}$ is the overall surface density containing the typical values for the rocky (r), rocky and icy ($r+i$), and gaseous (g) planets in our solar system, given by

$$\Sigma_r \approx 7 \frac{\text{g}}{\text{cm}^2} \left(\frac{r}{\text{AU}} \right)^{-3/2} \quad \text{for } r < 2.7 \text{ AU} \quad (2.55)$$

$$\Sigma_{r+i} \approx 30 \frac{\text{g}}{\text{cm}^2} \left(\frac{r}{\text{AU}} \right)^{-3/2} \quad \text{for } 2.7\text{AU} < r < 36 \text{ AU} \quad (2.56)$$

$$\Sigma_g \approx 1700 \frac{\text{g}}{\text{cm}^2} \left(\frac{r}{\text{AU}} \right)^{-3/2} \quad \text{gaseous component} \quad (2.57)$$

$$(2.58)$$

Integrating the gas component of the MMSN up to 30 AU leads to $13M_J$ which is comparable with estimated mass of protoplanetary disks observed around other stars. It suggests as well most of the gas was not captured by giant planets.

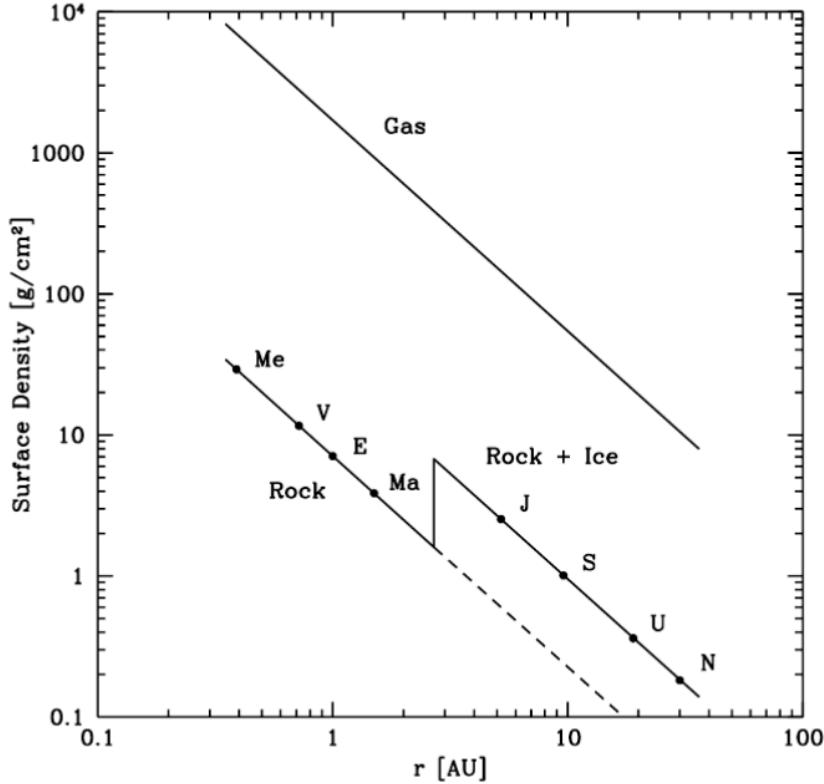


Figure 2.18: Display of the Minimum Mass Solar nebula (MMSM)

We can estimate the time-scale required to reach the isolation mass by looking at the ratio of M_{iso} and the planetesimal accretion rate derived earlier. Assuming values of $a = 5.2$ AU and $\Sigma_s = 10$ (gcm^{-2}) we find

$$\tau_{iso} = M_{iso} \cdot \left(\frac{dM_s}{dt} \right)^{-1} \quad (2.59)$$

$$\sim 3 \cdot 10^8 F_g^{-1} \Omega^{-1}. \quad (2.60)$$

If we assume a typical enhancement factor $F_g \approx 100 - 1000$ over the entire accretion epoch, we see that τ_{iso} is tens of thousands of orbital periods, meaning

$$\tau_{iso} \sim 10^5 - 10^6 \text{ yr.} \quad (2.61)$$

2.6.6 Growth of thermally regulated envelopes

Gas accretion occurs when thermal kinetic energy of a gas (c) is smaller than the gravitational energy from the planetary embryo (V_{esc}). This may be expressed by escape velocity $c < v_{esc}$.

With $v_{esc} = \sqrt{2GM_s/R}$ and using $c \sim H\Omega = H\sqrt{GM/a^3}$ one can obtain

$$M_s > \frac{c^2 R}{2G} \quad (2.62)$$

$$> \frac{H^2 M_* R}{2a^3}. \quad (2.63)$$

Assuming typical values of $H/a \sim 10^{-2}$, $R \sim 1000\text{km}$, and $a \sim 10^8\text{km}$ (about 1 AU) we find $M_s > 10^{-9}M_*$. It is about 10% mass of the Moon. This shows that small bodies like planet cores are accreting gas in the disk early in the process.

As long as the atmosphere is optically thin all thermal energy from collisions can be released into space, allowing it to grow by accreting more gas. When the atmosphere is too dense and becomes optically thick, it cannot radiate all the thermal energy away any more. A pressure gradient in the envelope is build up, reducing the accretion of large amounts of gas. An equilibrium is reached between gravitational contraction and released energy. The process becomes dominated by the nature of the atmosphere and the amount of energy that still can be radiated into space (assuming unlimited supply of gas for accretion). This state of quasi static contraction is defined by the Kelvin-Helmholtz time-scale

$$\tau_{KH} = \frac{E_{grav}}{L} \quad (2.64)$$

$$= \frac{GM_c M_e}{\bar{R}L}, \quad (2.65)$$

where E_{grav} is the gravitational energy released by contraction, L is the envelopes luminosity⁸, M_c and M_e are the masses of the core and the envelope, and \bar{R} is the average radius of the protoplanet.

This so called "slow contraction phase" represents the main uncertainty in the time-scale of the gas capture by giant planets, since the gas dissipates in about 10^7 yr. The value of τ_{KH} hence describes whether a core can accrete enough gas to evolve into a giant gas planet before the disk's gas dissipates. (see fig. 2.19 on the facing page)

Detailed atmospheric models indicate that when the core mass passes a critical value $M_c^{cr} > 16 M_\oplus$, the hydrostatic equilibrium cannot be sustained and the envelope collapses. From this time on the accretion is rapid, and governed by nothing else than the disk - planet interaction. See on fig. 2.20 on page 32 for a computed model of giant planet growth.

Note that the diversity of planets we find nowadays is intrinsically build into this model. It naturally explains how gas giants like Jupiter and Saturn can have formed, while Uranus and Neptune failed to gain a massive envelope of gas. They are often seen as "failed" giant planets.

⁸The luminosity is related to the opacity of the envelope and its thorough computation necessitates detailed atmospheric modelling.

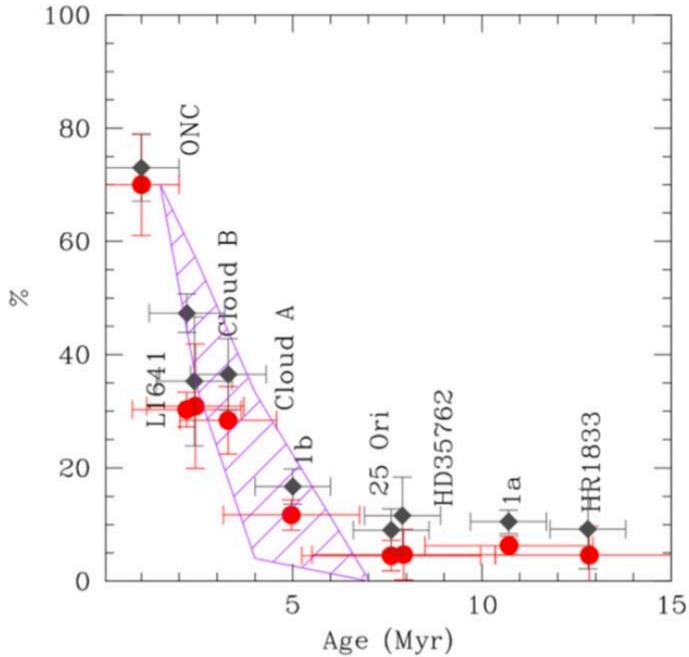


Figure 2.19: Fraction of star with evidence of gaseous disk against age of the star (from Mamajek 2009)

2.6.7 Disk-limited gas accretion rate

Similarly to the Safronov theory for the growth of solid bodies, the gas accretion rate can be estimated by

$$\frac{dM}{dt} \sim \pi R_{gc}^2 \rho v_{rel}, \quad (2.66)$$

in which R_{gc} is the radius within which gas can be captured by the planet core. With $\rho \sim \Sigma/H$, $v_{rel} \approx \Omega \Delta a$, and $\Delta a \sim R_{gc}$ we find

$$\frac{dM}{dt} \sim \frac{\Sigma}{H} \Omega R_{gc}^3. \quad (2.67)$$

We already know that the Hill Radius R_H describes the distance up until which the gravitational attraction of the protoplanet dominates over the star's gravity and centrifugal forces. In the case of a gas, to compute R_{gc} one should as well consider gas thermal energy, expressed in term of Bondi radius:

$$R_B = \frac{GM}{c^2} \quad (2.68)$$

It is the distance where gravitational attraction of the protoplanet dominates over the thermal energy of the gas.

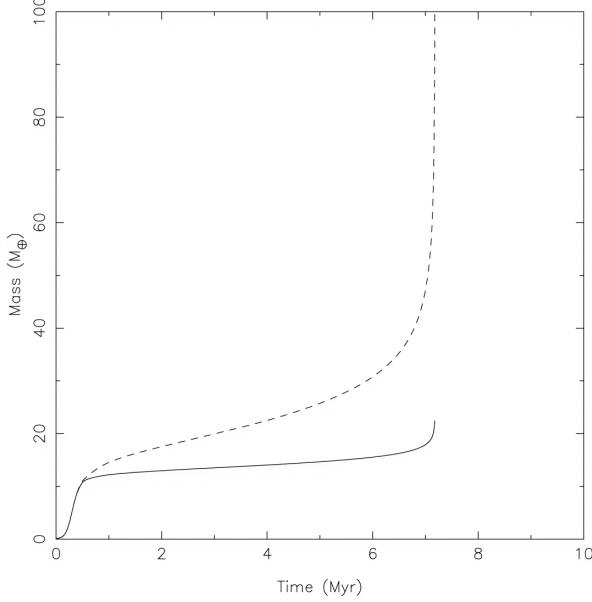


Figure 2.20: The evolution of the core mass (solid line), envelope mass (dotted line) and total mass (dashed line) from a time-dependent calculation of giant planet formation via core accretion (Rice & Armitage, 2003).

When $R_B < R_H$ the internal gas pressure within the Hill sphere only may prevent gas from being captured by protoplanet, meaning that the Bondi Radius is the maximal relevant distance for gas accretion, not the Hill Radius (Bondi-type gas accretion rather than Hill-type gas accretion). See Illustration of two competing effects on fig. 2.21 on the facing page.

On practical terms accretion radius is defined as

$$R_{gc} = \min(R_B, R_H). \quad (2.69)$$

We can formally derive the distinct time-scales (τ_B and τ_H) of each physical process with as usual $\tau = M(dM/dt)^{-1}$ and define the transition between the two regimes, at $\tau_B = \tau_H$, leading to transition mass

$$M_{tr} \sim M_* \left(\frac{H}{a} \right)^3. \quad (2.70)$$

A practical example of such a transition is illustrated on figure 2.21 on the next page.

2.6.8 Gap formation, torque and migration

The interactions in the disk produce a torque on the planet. Considering a Keplerian disk, i.e. $\Omega = \sqrt{GM/r^3}$, the gas outside of the planets orbit moves faster/slower than the planet

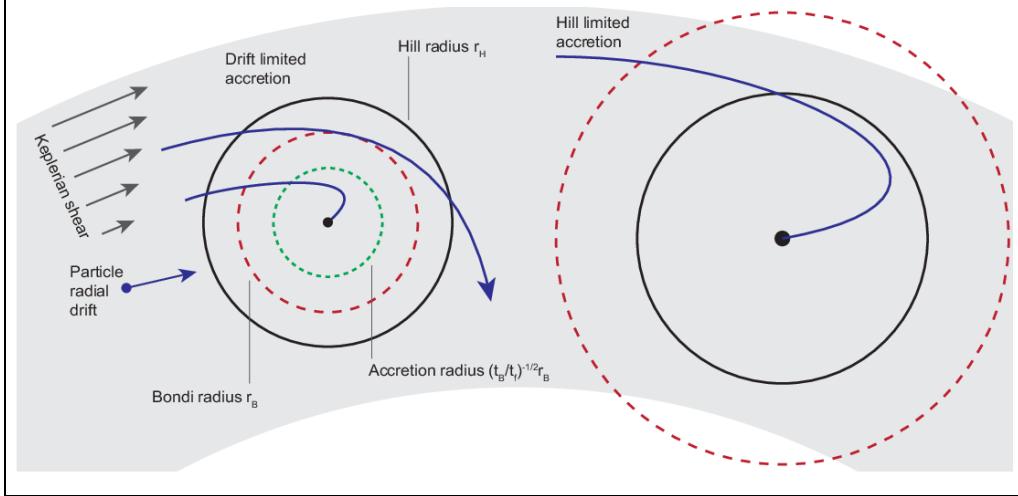


Figure 2.21: Illustration of the two main regimes of aerodynamically assisted accretion. For lower mass planet cores (left) the Bondi radius r_B , within which purely gravitational 2-body interactions would lead to strong scattering, is smaller than the Hill radius r_H . Aerodynamically coupled particles enter the Bondi radius with their radial drift velocity and are accreted if their stopping time is shorter than the time it takes to traverse the Bondi radius. For higher mass planet cores (right) the Bondi radius exceeds the Hill radius, and the encounter velocity of particles approaching the planet is set by Kepler shear rather than radial drift. Only those particles that enter the Hill sphere have a chance of being accreted (from Armitage 2004)

(for $r = a - \Delta a$ and $r = a + \Delta a$, respectively). When passing by close to the planet, the gas is gravitationally attracted and both exchange angular momentum⁹.

A simple description of the scenario is given by the "impulse approximation", developed by Lin and Papaloizou in 1979, and can be seen in figure 2.22 on the following page.

In the following we treat the problem purely as a two body interaction. The coordinate system is set at the center of the protoplanet, such that $x = r - a$. Before the impact the particle's velocity is $v = -v_0 \cdot \hat{y}$, with $v_0 \approx \Omega x$ (see Eq.(2.46 on page 26)). The total effect is approximated to happen as a single impulse Δp happening at closer approach moment when $y = 0$, and changing the trajectory direction of the particle in the disk

$$\Delta p_x = M \cdot \Delta v_x = F_x \cdot \Delta t. \quad (2.71)$$

F_x is the gravitational force exerted by the planet on the particule with $t = 0$ corresponding to the closest approche, the impulse moment). We can write the perpendicular force, F_x , on the particle as

⁹In reality this effect is again complicated by various turbulent effects

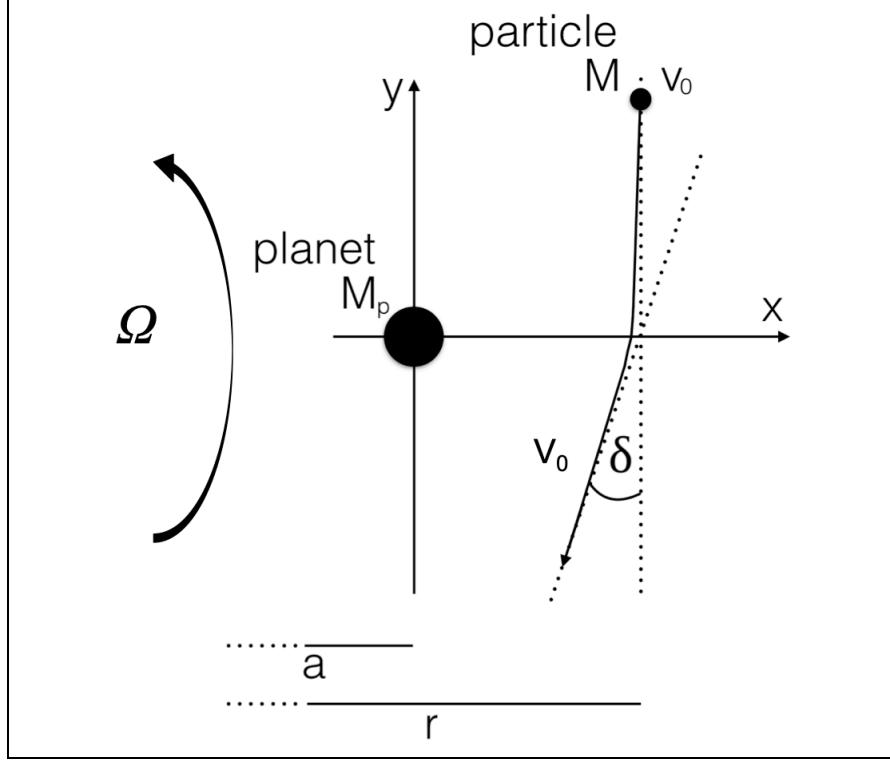


Figure 2.22: Schematic of the planet-particle "impulse approximation"

$$F_x = -\frac{GM M_p}{x^2} \left(1 + \left(\frac{v_0 t}{x} \right)^2 \right)^{-1} \quad (2.72)$$

Using Eq. 2.71 on the preceding page and integrating over the duration of the encounter yields the corresponding effect of the impulse felt by the planetesimal:

$$\Delta v_x = \int_{-\infty}^{+\infty} \frac{F_x}{M} dt \quad (2.73)$$

$$= \frac{GM_p}{xv_0} \int_{-\infty}^{+\infty} (1/(1+t'^2)) dt' \quad (2.74)$$

$$= \frac{2GM_p}{xv_0}. \quad (2.75)$$

Then we use the impact approximation to compute the deflection $\delta = \Delta v_x / v_0$, leading to

$$\delta \sim \left(\frac{M_p}{x v_0^2} \right) \quad (2.76)$$

$$\sim \left(\frac{M_p}{M_*} \right) \left(\frac{a}{x} \right)^3 \quad (2.77)$$

$$\sim \left(\frac{R_H}{x} \right) \quad (2.78)$$

With this change of the trajectory v_y is modified by

$$\Delta v_y = v_0(1 - \cos \delta), \quad (2.79)$$

(note Δv_y for small trajectory changes is positive). Then using the first order approximation with $\cos \delta \approx 1 - \delta^2$ we find

$$\Delta v_y \sim v_0 \delta^2 \quad (2.80)$$

$$\sim x \Omega \left(\frac{M_p}{M_*} \right)^2 \left(\frac{a}{x} \right)^6 \quad (2.81)$$

$$\sim a \Omega \left(\frac{M_p}{M_*} \right)^2 \left(\frac{a}{x} \right)^5 \quad (2.82)$$

This means that the interacting particle seen from the planet referential *gains speed* (in y -direction). The corresponding gains in angular momentum *by the interacting particle* is given by

$$\Delta J \sim Ma \Delta v_y \quad (2.83)$$

$$\sim Ma^2 \Omega \left(\frac{M_p}{M_*} \right)^2 \left(\frac{a}{x} \right)^5. \quad (2.84)$$

Since the system planet-particle keeps a constant angular momentum, the planet's angular momentum decreases accordingly. By speeding up the particle the planet spirals in. From the planet perspective it may be understood in the same way as a friction mechanism.

Applying this equation in the continuum case of a disk with a *flux of particles distributed along x -axis* leads to $dM \sim v_0 \Delta t \Sigma dx$.

When substituting M by its integral one can rewrite equation 2.84 as

$$\Delta J \sim \int a^3 \Delta t \Sigma \Omega_P^2 \left(\frac{M_p}{M_*} \right)^2 \left(\frac{a}{x} \right)^4 dx. \quad (2.85)$$

From this equation one can compute the torque on the planet such that $\mathcal{T}_{\text{planet}} = -\mathcal{T}_{\text{disk}}$,

$$\mathcal{T}_p^{out} \sim \frac{-\Delta J}{\Delta t}. \quad (2.86)$$

$$(2.87)$$

The torque on the planet resulting from the **outer parts** (at $r \geq a + \Delta r$) of the disk is then

$$\mathcal{T}_p^{out} \sim -a^3 \Sigma \Omega_P^2 \left(\frac{M_p}{M_*} \right)^2 \int_{\Delta r}^{\infty} \left(\frac{a}{x} \right)^4 dx \quad (2.88)$$

$$\sim -\Sigma \Omega^2 a^4 \left(\frac{M_p}{M_*} \right)^2 \left(\frac{a}{\Delta r} \right)^3. \quad (2.89)$$

Equally we have to calculate the torque resulting from the inner parts of the disk, i.e. where $r \leq a - \Delta r$. In doing so we find a similar expression with signed inverted. Since disk structure is not symmetrical the residual is the resulting acting torque $\mathcal{T}_p = \mathcal{T}_{out} + \mathcal{T}_{in} \sim -\Sigma^{out} + \Sigma^{in}$. We see from this that the direction of the migration depends of degree of asymmetry in the disk. To get a sens of the main direction on can use another point of view, assuming only a few orbit matters when computing the torque. This idea is based on the original work of (Goldreich and Tremaine in 1976) to decompose the resulting torque into a sum of partial torques exerted at resonant locations with the disk. In disk dynamics yields, the conditions for resonances exists for radii in the disk where the angular frequency is a fractional number of the angular frequency Ω_p . For a Keplerian disk one define the *Lindblad resonance* when

$$\Delta m(\Omega - \Omega_p) = \pm \Omega. \quad (2.90)$$

In the disk they are located at radius (from the star) r_L

$$r_L = \left(1 \pm \frac{1}{m} \right)^{2/3} r_p. \quad (2.91)$$

Each orbit $r = r_L$ experiences strong perturbation to the motion. Practically for a gaseous disk, angular momentum exchange between the planet and the disk occurs essentially at resonant locations r_L . On one hand the planet gains angular momentum from interacting with the gas disk at the interior Lindblad resonances (ILR) ($r_L < r_p$) and symmetrically loses angular momentum from interacting with the gas disk at exterior (outer) of Lindblad resonances (OLR) ($r_L > r_p$). Ward in 1997 employed this framework to compute the *smoothed torque density* due to angular momentum exchange between a planet and a gas disk at the location of Lindblad resonances and found the interaction is almost invariably asymmetric (see figure 2.23 on the facing page), such that when integrated over the entire disk the planet loses angular momentum and migrates inward. Changes to the planets orbit as a result of this net torque are called **Type I migration**.

In addition to the effect of Lindblad resonances, there is also an additional co-orbital torque \mathcal{T}_{CR} from the gas co-orbiting with the planet. The gas, near the co-orbital orbit, follows an horseshoe trajectory when viewed in a frame co-rotating with the motion of the planet (see figure 2.24 on page 38). As the gas executes the U-shaped turns at the ends of the horseshoe, changes in the gas density occur that are not symmetric between the disk trailing and leading the planet. The density variations is the source of the corotation torque.

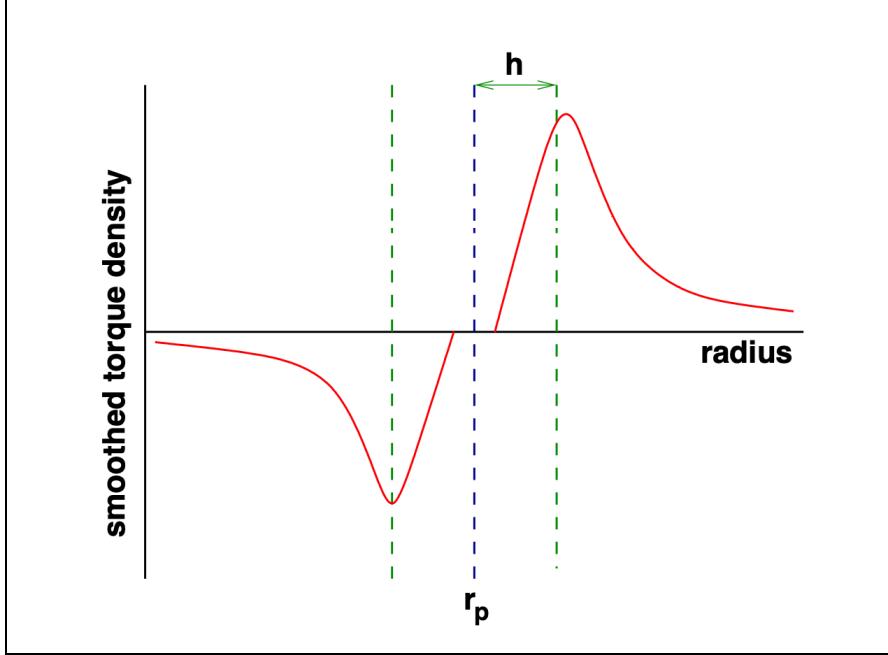


Figure 2.23: Schematic illustration of the smoothed torque density due to angular momentum exchange between a planet and a gas from Ward (1997). The peak torque occurs at $r = rp \pm h$, where h is the disk scale height

It can be as large or larger than the torque from the Lindblad resonances. The total torque experienced by planets is determined by the sum of Lindblad and corotation torques. Type I migration rate is specific to disk model and depends on the planet mass. They all suggest that Type I migration can rarely be ignored in any circumstance where substantial gas disks co-exist with planets. For massive planets there are radial zones of outward migration interspersed with radii where migration is inward (Armitage 2020)

The gravitational interaction of a planet with a gas disk tends to repel gas from the vicinity of the planets (high- m Lindblad resonances) creating a gap. Practically the Hill sphere (R_H) of the planet needs to be comparable to the thickness of the gas disk (H). With $R_H > H$, using Eq. 2.44 on page 25 and $q \equiv M_p/M_\star$ on finds

$$q > 3 \left(\frac{H}{r} \right)^3 \quad (2.92)$$

This condition is satisfied for typical protoplanetary disk parameters for $q \sim 4 \cdot 10^{-4}$ (Saturn mass range).

A second condition to meet is to prevent the gap to fill back by viscosity of the disk. In the α disk model the viscosity is given by $\nu = \alpha c H$, where α is a characteristic scaling parameter for the viscosity. Practically a gap opens when acting torque is larger than

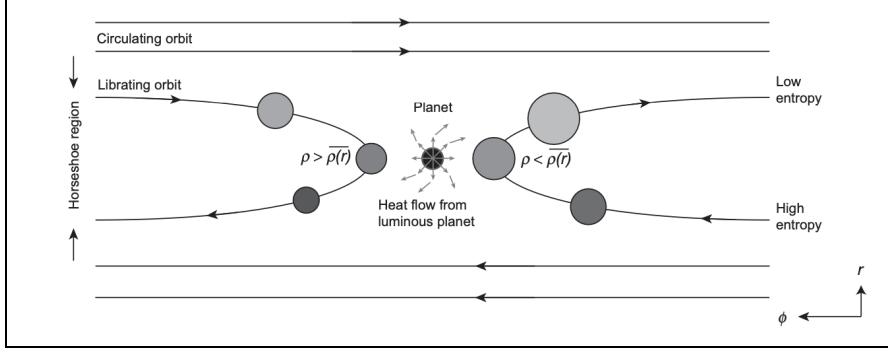


Figure 2.24:

viscosity effect (think about your finger moving around in a pot of jam). The typical time-scale for the viscosity to close the gap may be estimated by $\tau_\nu \approx a^2/\nu$. If one equals it with the time-scale to open it ($\Delta J/|dJ/dt|$) one finds a critical mass ratio $(M_p/M_*)^2 \approx (H/a)^5\alpha$, about $60 M_\oplus$ for typical disk parameters ($\alpha = 0.01$, $H/a = 0.05$). When viscosity balances gap opening depletion one speaks of a **type II migration** now. The transferred angular momentum is related to the viscous time-scale (like for tidal effects). With $\tau_\nu \approx a^2/\nu$ and $c \sim H\Omega$ one can get the following description of the migration:

$$\Omega\tau_{II} \sim \frac{1}{\alpha} \left(\frac{a}{H} \right)^2, \quad (2.93)$$

where the factor α can vary by many orders of magnitudes depending on the assumption made on the physical mechanisms at the origin of the viscosity effect.

With the assumption of a/H constant over the entire disk one finds

$$\tau_{II} \sim \frac{1}{\Omega} \sim a^{3/2}, \quad (2.94)$$

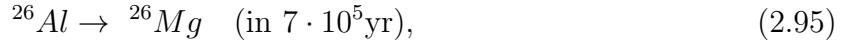
telling us that the migration rate $\frac{1}{\tau_{II}}$ increases the closer the planet is to the star. In order for it to stop, effects like tides, magnetic fields and gaps will then play a role.

As discussed earlier, the entire process is highly dynamic and can only be understood using highly detailed simulations. Migration effects are considered an important building block in planetary formation to explain the many planets of all sort detected with short orbits.

2.7 Timing early planet formation

In the solar system the decay of unstable isotopes can be used to track the formation age of refractory (non-volatile) elements, **assuming the element was injected at the time of the formation of the solar system**. Elements that are often used to map initial stages

are short lived isotopes like ^{26}Al :



respectively reefers as parent and daughter isotopes. ^{26}Al is particularly interesting because its half-life decay matches planet formation time-scales (few Myr) and well sample this short period. Moreover for measurements they can be scaled against their corresponding stable isotopes ^{27}Al and ^{24}Mg to make ratio-ratio isochron diagrams.

Calcium-Aluminium-rich refractory inclusions (CAIs) are found in meteorites (see fig. 2.25). They exist in a wide diversity of compositions, mineralogy, structure, and size (spinel, melilite, hibonite, anorthite, pyroxene,...). They are the oldest known objects from the solar system: 4567.2 ± 0.1 Myr. They are alike time capsule that can be used to trace back the time of their formation.

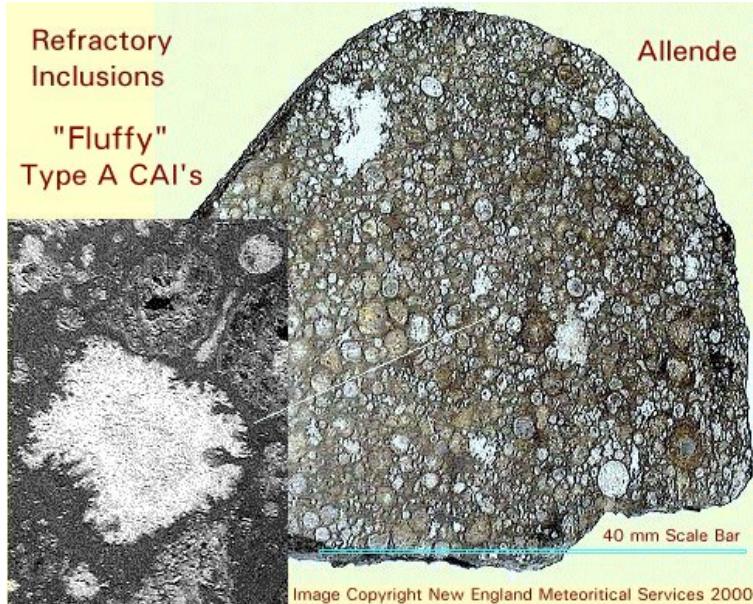


Figure 2.25: Allende meteorite, with zoom in on CAI's inclusions

The isotopic that is observed today is related to its original parent isotope by

$$(^{26}Mg)_{\text{today}} = (^{26}Mg)_{\text{original}} + (^{26}Al)_{\text{original}} (1 - e^{-t/\tau}) \quad (2.96)$$

When (^{26}Mg) is compared with its stable daughter isotope (^{24}Mg)

$$\underbrace{\left(\frac{^{26}Mg}{^{24}Mg} \right)_{\text{today}}}_{\delta^{26}Mg} = \underbrace{\left(\frac{^{26}Mg}{^{24}Mg} \right)_{\text{original}}}_a + \left(\frac{^{27}Al}{^{24}Mg} \right)_{\text{today}} \times \underbrace{\left(\frac{^{26}Al}{^{27}Al} \right)_{\text{original}} (1 - e^{-t/\tau})}_{b(t)}, \quad (2.97)$$

relating to an equation of the form

$$\delta^{26}Mg = a + \left(\frac{^{27}Al}{^{24}Mg} \right)_{today} b(t), \quad (2.98)$$

where the slope b depends only of the time sample studied was formed (rock solidifying). Each mineral containing the ratio of stable isotope of $^{27}Al/^{24}Mg$ varies from sample to sample. For short lived isotope like ^{26}Al , most of it is being converted into ^{26}Mg during the first million years a time when lots happen for planet formation, making Al an interesting element to trace. The uncertainties on timing of planet formation is negligible compared to age of the solar system meaning that t in Eq. 2.96 on the preceding page may be considered as a constant and $b(t)$ essentially given by ratio of $^{26}Al/^{27}Al$

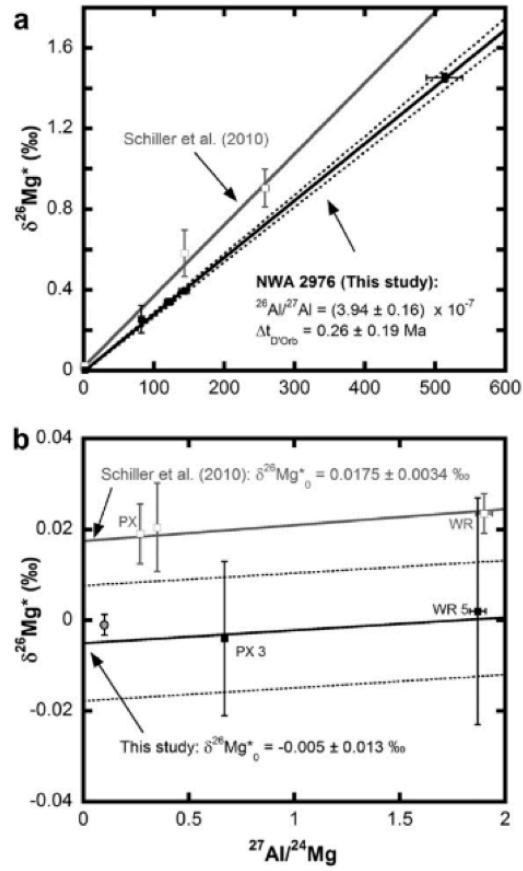


Figure 2.26: Example of use of $\delta^{26}Mg$ method to get age of oldest rock formation

Chapter 3

Exoplanet detection

3.1 Elements of orbits and dynamics

3.1.1 Ellipse equations and Kepler's laws

In a two-body problem with two orbiting mass (m_1 and m_2), the dynamic of the system is given by

$$\mathbf{F}_1 = +\frac{GM_1M_2}{r^3}\mathbf{r} = M_1\ddot{\mathbf{r}}_1 \quad (3.1)$$

$$\mathbf{F}_2 = -\frac{GM_1M_2}{r^3}\mathbf{r} = M_2\ddot{\mathbf{r}}_2, \quad (3.2)$$

where

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1. \quad (3.3)$$

$\mathbf{F}_1 + \mathbf{F}_2 = \mathbf{0}$, leads to the equation

$$\ddot{\mathbf{r}} + G(M_1 + M_2) \cdot \frac{\mathbf{r}}{r^3} = 0. \quad (3.4)$$

By the nature of the potential and its resulting force $\mathbf{r} \times \mathbf{F}_{1,2} = 0$, then $\mathbf{r} \times \dot{\mathbf{r}} = 0$. This leads to $\mathbf{r} \times \dot{\mathbf{r}} = \mathbf{h}$, with \mathbf{h} a constant usually called *angular momentum integral*. Finally when expressed in polar coordinates (r, Θ) , using $\mathbf{r} = r\hat{r}$ and $\dot{\mathbf{r}} = \dot{r}\hat{r} + r\dot{\Theta}\hat{\Theta}$, one finds

$$r^2\dot{\Theta} = h. \quad (3.5)$$

If one defines the area of the sector defined by the motion as

$$\delta A = \int_0^r r dr \delta\Theta = \frac{1}{2}r^2\delta\Theta,$$

one obtains the famous *Kepler's second law*

$$\dot{A} = \frac{1}{2}h = \text{Cte.} \quad (3.6)$$

Writing the Equ. (3.4 on the previous page) in polar coordinates using $\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\Theta}^2)\hat{r} + \left[\frac{1}{r}\frac{d}{dt}(r^2\dot{\Theta})\right]\hat{\Theta}$, and looking at the solution in the radial direction one can write

$$\ddot{r} - r\dot{\Theta}^2 = -\frac{G(M_1 + M_2)}{r^2}. \quad (3.7)$$

Then substituting $u = 1/r$ one obtains a second-order linear differential equation known as Binet's equation, with the following general solution to describe the orbit:

$$r = \frac{h^2}{G(M_1 + M_2)} \cdot \frac{1}{1 + e \cos(\Theta - \omega)}, \quad (3.8)$$

in which e and ω appear as integration constants. This solution corresponds to conic section equation family (circle, ellipse, parabola or hyperbola) (see fig. 3.1)

If one considers the ellipse solution one finds

$$r = \frac{a(1 - e^2)}{1 + e \cos f}. \quad (3.9)$$

In this equation a is the semi-major axis of the ellipse, and e is the eccentricity, and $f = (\Theta - \omega)$, called **true anomaly**, is the angle between the current position of the orbiting object and the periastron¹ w_0 . Connecting the physical parameters with geometrical description in the case of the elliptical motion one finally finds

$$\frac{h^2}{G(M_1 + M_2)} = a(1 - e^2). \quad (3.10)$$

Further, when considering the ellipse properties, $A = \pi ab$ and $b^2 = a^2(1 - e^2)$ and one defines the period of the orbit P

$$P = \frac{2}{h} \cdot \int \frac{h}{2} dt = \frac{2}{h} \cdot \int \dot{A} dt = \frac{2A}{h}, \quad (3.11)$$

one finds

$$h = \frac{2\pi ab}{P}, \quad (3.12)$$

$$h = \frac{2\pi}{P} a^2 \sqrt{1 - e^2}. \quad (3.13)$$

Finally using Eq.(3.10) and Eq.(3.13) one derives the **Kepler's third law**

$$P^2 = \frac{4\pi^2}{G(M_1 + M_2)} a^3. \quad (3.14)$$

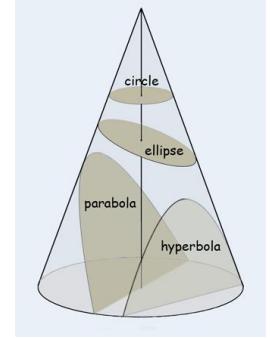


Figure 3.1:
Display of conic
section Family

¹location in the orbit closest to the central body

3.1.2 Time evolution of orbit equations

To express the solution as a function of time one needs to express $r(t)$ instead of $r(\Theta)$. Starting from the orbital motion one can demonstrate that

$$\dot{r} = \frac{2\pi}{T} \cdot \frac{a}{r} \sqrt{a^2 e^2 - (r - a)^2}. \quad (3.15)$$

To solve the equation one makes the following transformation

$$r(t) = a(1 - e \cos E(t)), \quad (3.16)$$

The differential equation is transformed to

$$\dot{E}(t) = \frac{2\pi}{P} \cdot \frac{1}{1 - e \cos E(t)}. \quad (3.17)$$

with the following solution:

$$\frac{2\pi}{P}(t - t_0) = E(t) - e \sin E(t). \quad (3.18)$$

with $E(t)$ is called the **eccentric anomaly**. The left side is defined as the **mean anomaly** $M(t) = \frac{2\pi}{P}(t - t_0)$, with t_0 time at the periastron. Finally the orbital motion equation is given by

$$M(t) = E(t) - e \sin E(t). \quad (3.19)$$

This equation cannot be solved directly, but just iteratively, starting with $E_0 = M_0$.

Although we cannot have an explicit relation between the angle f and M , combining eq. 3.13 on the preceding page with the integral of the motion (eq. 3.5 on page 41) it is possible to obtain a differential relation between the time and the angular frequency

$$df = 2\pi/P \sqrt{1 - e^2} \left(\frac{a}{r}\right)^2 dt. \quad (3.20)$$

This expression is useful to average physical quantities over a complete orbit, for instance

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{dM}{r^2} = \frac{1}{a^2 \sqrt{1 - e^2}}. \quad (3.21)$$

Combining Eq. (3.16) and Eq. (3.9 on the facing page) one can express the true anomaly f with the eccentric anomaly E .

$$\tan(f/2) = \sqrt{\frac{1+e}{1-e}} \tan(E/2) \quad (3.22)$$

An additional useful angle is the mean longitude $\lambda = M + \omega$

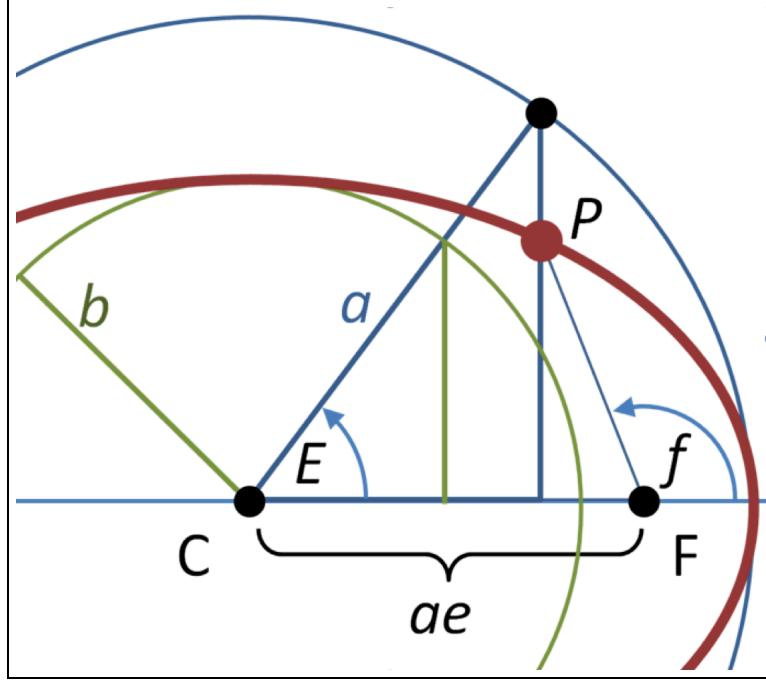


Figure 3.2: Geometry of keplerian orbit

3.1.3 Coordinate system transformation

If one wants to express orbital motion in a more general form to account for three-dimensions one needs to consider direction perpendicular to orbital plan (z -axis), and then transform our system (x, y, z) , with $x = r \cos f$, $y = r \sin f$, into a new system (X, Y, Z) using three rotations matrix $\mathbf{P}_z(\Omega)$, $\mathbf{P}_x(i)$, $\mathbf{P}_z(\omega)$. Now in addition to ω , one has two more angles: i and Ω (see figure).

Practically, for astronomical application (X, Y, Z) coordinate system is defined such that (X, Y) represents the plane of the sky (seen from the Earth perspective with Y pointing North), and Z -axis points directly towards the observer. We shall use this ref in this lecture notes. When considering an Earth referential the equatorial plan represents X-Y and vernal equinox X-axis direction. In all case The intersection between the orbital plan and the reference plan is **line of nodes**. In the astronomical case the **ascending node** is the intersection of the orbit (on the side moving towards the observer) with the reference plan (the sky).

$$(X, Y, Z)^T = \mathbf{P}_z(\Omega)\mathbf{P}_x(i)\mathbf{P}_z(\omega)(x, y, z)^T$$

Looking at the Z direction (towards the observer) we find the expression

$$Z(t) = r(t) \sin(\omega + f) \sin i, \quad (3.23)$$

showing that the transformation operator Ω has no effect here, which becomes obvious if we look at the figures 3.3.

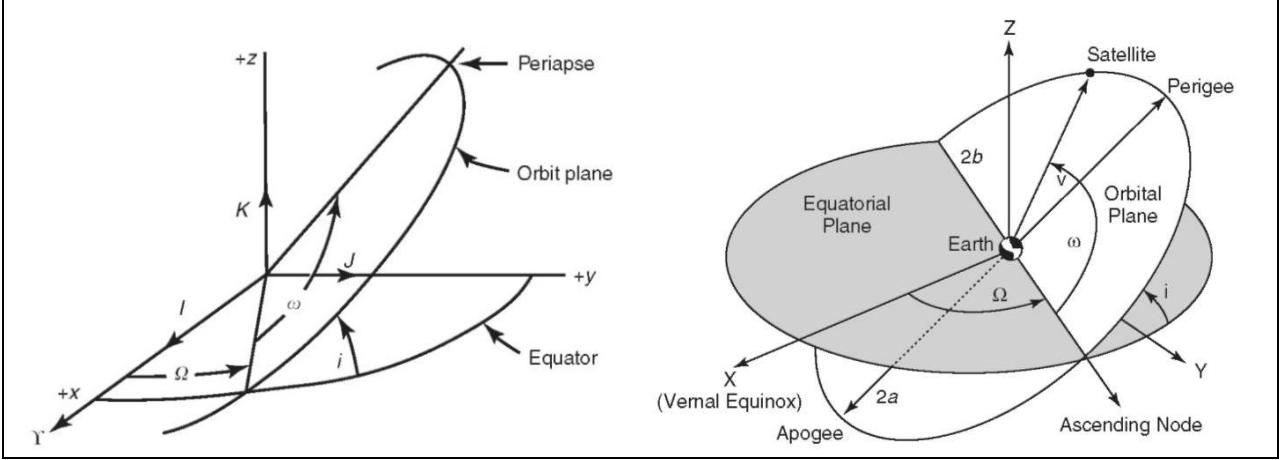


Figure 3.3: Astrophysical correspondence of Keplerian orbit parameters

3.1.4 Barycentric motion

All the above equations are based on the definition of the coordinate system to be centered in one of the two bodies (fig. 3.4 on the following page). However, in real cases we want to study the motion of the objects around their common center of mass, i.e. their barycentric motion:

$$\mathbf{R} = \frac{M_1 \mathbf{r}_1 + M_2 \mathbf{r}_2}{M_1 + M_2}, \quad (3.24)$$

We further define $\mathbf{R}_{1,2} = \mathbf{r}_{1,2} - \mathbf{R}$, and $\mathbf{r} = \mathbf{R}_2 - \mathbf{R}_1$ (with $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$).

On obtains:

$$\mathbf{R}_1 = -\frac{M_2}{M_1 + M_2} \cdot \mathbf{r} \quad (3.25)$$

$$\mathbf{R}_2 = \frac{M_1}{M_1 + M_2} \cdot \mathbf{r}, \quad (3.26)$$

meaning the relative orbital motion described by \mathbf{r} applies as well to each individual bodies scaled by the mass of the other normalised by the system system. In term of elliptic motion one can scale the semi major axis the following way:

$$a_1 = \frac{M_2}{M_1 + M_2} \cdot a \quad (3.27)$$

$$a_2 = \frac{M_1}{M_1 + M_2} \cdot a. \quad (3.28)$$

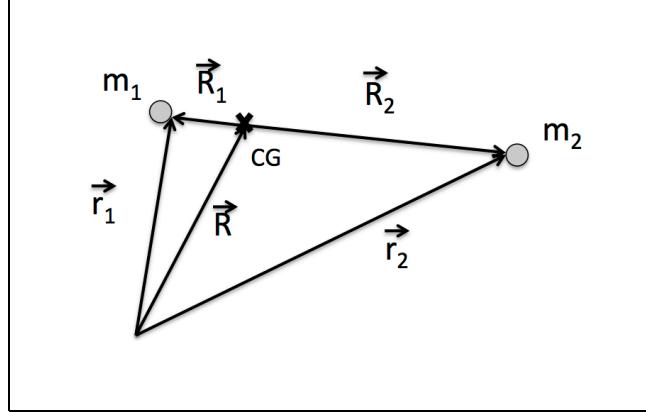


Figure 3.4: Coordinate system

Note that this is a very interesting and important point - if you want to find out the mass of object 2 (planet for instance) you need to look at the motion of object 1 (its star)!.

3.2 Radial velocity technique

3.2.1 Studying the radial velocity of a star

The radial motion \hat{z} corresponds to the radial velocity. For the star we get²:

$$v_{z_1} = \dot{\mathbf{r}}_1 \cdot \hat{z} \quad (3.29)$$

$$= (\dot{\mathbf{R}} + \dot{\mathbf{R}}_1) \cdot \hat{z} \quad (3.30)$$

$$= v_z + \frac{M_2}{M_1 + M_2} \dot{\mathbf{r}} \cdot \hat{z}. \quad (3.31)$$

Using the general expression for z from Eq. (3.23 on page 44) one finds

$$\dot{\mathbf{r}} \cdot \hat{z} = \dot{r} \sin(\omega + f) \sin i + r \dot{f} \cos(\omega + f) \sin i. \quad (3.32)$$

Using equation (3.15 on page 43) for \dot{r} and $h = r^2 \dot{f}$ (3.10 on page 42), one finds:

$$\dot{\mathbf{r}} \cdot \hat{z} = \frac{2\pi}{P} \frac{a \sin i}{\sqrt{1 - e^2}} (\cos(\omega + f) + e \cos \omega). \quad (3.33)$$

Finally one obtains:

$$v_{r_1} = v_z + K_1 \cdot (\cos(\omega + f) + e \cos(\omega)), \quad (3.34)$$

²Note that the suffix 1 is usually given to the more massive component in a two body problem.

with K being defined as

$$K_1 = \frac{2\pi}{P} \frac{a \sin i}{\sqrt{1-e^2}} \frac{M_2}{M_1 + M_2}. \quad (3.35)$$

Playing with w and e allows a large diversity of shapes as displayed in fig. 3.5

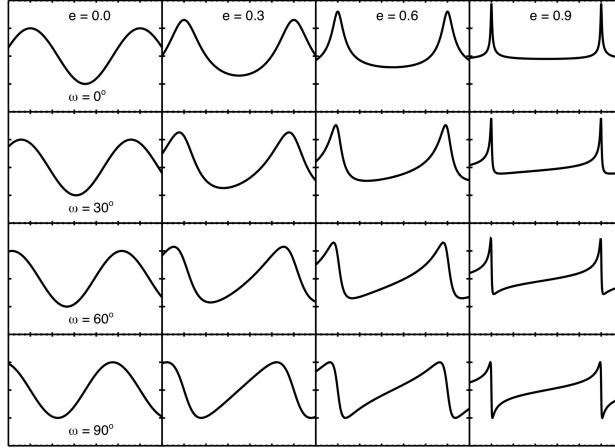


Figure 3.5: The effects of e and ω on on radial velocity curves. These curves have been scaled to unit K and common P (courtesy Jason Wright)

To obtain a better idea of teh effect in the case of a planet orbiting a star, one can assume M_1 is significantly larger than M_2 and evaluate some numbers to get the following handy equation

$$K_1 = 28.4 \left[\frac{m}{s} \right] \cdot M_2 \sin i [M_{Jup}] \cdot (P[yr])^{-1/3} \cdot (M_1[M_\odot])^{-2/3} \cdot \frac{1}{\sqrt{1-e^2}}. \quad (3.36)$$

Considering a star similar to our Sun, a planet like Jupiter at 5 AU would produce a maximum amplitude of 12.7 m s^{-1} . In the case of the Earth at 1 AU it is 9 cm s^{-1} . This explains why in the early years of planet search with radial velocity techniques the goal was to reach a precision of about 10 ms^{-1} , to detect giants planets (Jupiter like) around sun-like stars. Nowadays, best precision reach is around 10 cm s^{-1} gettign close to actually detect Earth-analogues. This is an improvement of precision by a factor of 1000 just in twenty years!

3.2.2 Doppler effect

The Doppler effect is given by

$$\lambda = \lambda_0 \frac{1 + \frac{1}{c} \mathbf{k} \cdot \mathbf{v}_*}{1 - \frac{\Phi}{c^2} - \frac{v_*^2}{2c^2}}. \quad (3.37)$$

Thereby λ_0 is the wavelength at which the source (host star in our case) emits, \mathbf{v}_* (the host star) is the velocity of the source, \mathbf{k} is the direction vector pointing from the observer to the source and λ the wavelength measured by the observer. Further, Φ is the gravitational reddening of the source, which is significant and constant for a dwarf star ($\sim 0.5 \text{ km s}^{-1}$) and can be considered as an offset. Note the radial velocity can never be measured *in absolute geometrical sens* without considering the general relativistic correction (Φ/c^2)!

Earth-bound Doppler-shift measurements must be corrected from the effects of local motions of the observer, i.e. Earth rotation and revolution around the Sun. The maximum amplitude of the radial velocities from Earth's orbit is about 30 km/s (along the axis of the Earth motion) and 500 m/s from Earth's rotation at the equator. Therefore we have

$$\lambda_{obs} = \lambda_0 \frac{1 + \frac{1}{c} \mathbf{k} \cdot (\mathbf{v}_* + \mathbf{v}_\oplus)}{1 - \frac{\Phi}{c^2} - \frac{v'^2}{2c^2}} \quad (3.38)$$

where v' is the projected amplitude of the observed stellar velocity seen by the observer on Earth.

The common reference frame in astronomy is the barycentre of our solar system. Neglecting the gravitational reddening and considering the small amplitude of radial velocity effect, non-relativistic approximation applies and one can write

$$\lambda_B = \lambda_0 \left(1 + \frac{1}{c} \mathbf{k}_B \cdot \mathbf{v}_* \right), \quad (3.39)$$

Practically the observed wavelength is "corrected" by applying a *barycentric correction*. In that case correction from the Earth gravitational potential may be neglected and if non-relativistic approximation is used residual effect is of the order of sub-mm/s and is usually considered as negligible. One obtains the following radial velocity equation:

$$\mathbf{k}_B \cdot \mathbf{v}_* = c \frac{\lambda_B - \lambda_0}{\lambda_0}. \quad (3.40)$$

The emitted signals detected with radial velocity techniques are line emissions of the stars. Knowing λ_0 of spectral lines provides an reference frame to mesure the radial velocity.

3.2.3 Noise limit to radial velocity measurements

We want to measure a Doppler shift the most optimal way to benefit from all information available in stellar spevtra. Each shift of a spectral line in wavelength, $\delta\lambda$, corresponds to velocity variation of δv , given by

$$\frac{\delta v}{c} = \frac{\delta\lambda}{\lambda}. \quad (3.41)$$

We now consider a calibrated reference spectrum with an amplitude A^0 , considered as being noise free. For each given pixel i we know the intensity A_i^0 of the measured target flux

(in photoelectrons) in dependency of the wavelength λ_i . At another epoch, the spectrum is stretched by a Doppler shift and becomes A . At this point we do not consider any external noise sources, such that these two spectra have the same intensity level.

For a Doppler shift that is small compared to the linewidth, the observable intensity change at a given pixel i is

$$A_i - A_i^0 \approx \frac{\partial A_i^0}{\partial \lambda_i} \delta \lambda_i, \quad (3.42)$$

or

$$\delta \lambda_i = (A_i - A_i^0) \left(\frac{\partial A_i^0}{\partial \lambda_i} \right)^{-1}. \quad (3.43)$$

Expressing $\delta \lambda_i$ in term of radial velocity changes leads to

$$\delta v_i = (A_i - A_i^0) \frac{c}{\lambda_i} \left(\frac{\partial A_i^0}{\partial \lambda_i} \right)^{-1}. \quad (3.44)$$

Assuming Poisson distribution of the measurements (with A_0 noise-free) we can estimate the uncertainty

$$\sigma(\delta v_i) = \sqrt{A_i} \frac{c}{\lambda_i} \left(\frac{\partial A_i^0}{\partial \lambda_i} \right)^{-1} \quad (3.45)$$

This expression is valid for one spectral line, or actually one pixel, i . When considering all lines one gets $\sigma_{RV} = \sqrt{\sum \sigma^2(\delta v_i)}$

Practically to compute the global effect one needs to compute the weighted sum of all contributing pixels δv_i with a weight $w_i = \sigma^{-2}(\delta v_i)c^2$. When making the assumption $A_i \approx A$ one obtains in that case

$$\sigma_{RV} = c \left(\sum_i A \right)^{-1/2} \cdot Q^{-1} \quad (3.46)$$

With Q a dimension less quality factor, defined as

$$Q^2 = \left(\sum_i \lambda_i^2 \left(\frac{\partial A_i^0}{\partial \lambda_i} \right)^2 \frac{1}{A_i^0} \right) \left(\sum_i A \right)^{-1}. \quad (3.47)$$

The quality factor Q is a measurement of the "strength" of the line structure variation when computing the RV line. It corresponds to the amount of information available in the spectrum.

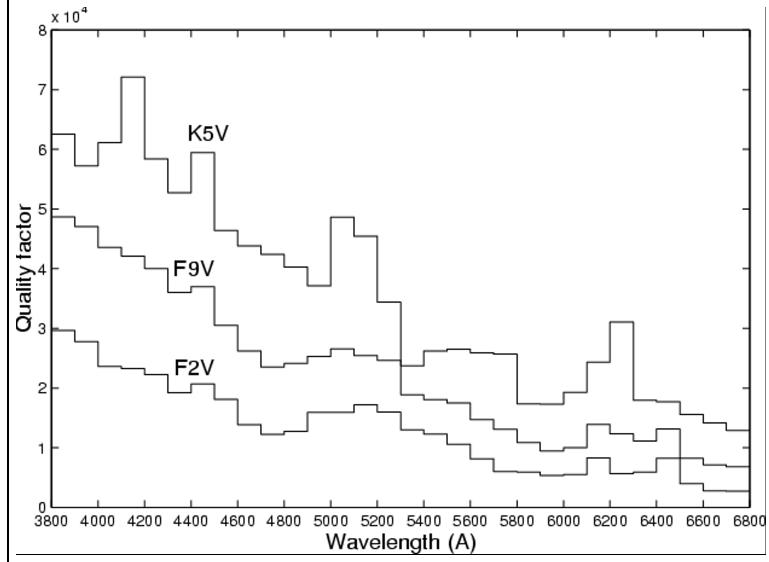


Figure 3.6: Computed quality factor Q on a star function of the wavelength

3.2.4 Physical line broadening

Physical line broadening can affect the measurement precision significantly. It can result from different causes, such as

- natural line broadening (quantum mechanics, generally far beyond resolution limits in astrophysics),
- Stark effect (for ions),
- Van der Waals forces (for neutral ions and molecules),
- thermal effects.

If we assume a Gaussian profile for the line shape, the final product of all these effects is a convolution of all their Gaussian profiles, i.e.

$$\gamma = \sum_i \gamma_i \quad (3.48)$$

For a star similar to the Sun, the typical width (spectral resolution of lines) is of the order 1 km s^{-1} . But this value is "blurred" by the Doppler effect broadening due to the rotation of the star.

3.2.5 The impact of stellar rotation ($v \sin i$)

A rotating star can be sketched by vertical stripes of different velocities, such that the x axis directly translates into a $\Delta\lambda$ axis (see fig. 3.7 on the next page). Note $i_{rotation}$ of the stellar rotation axis in that case is defined as the complementary angle of the i_{orbit} . The reason is that for the former we don't consider the plane of the stellar equator but the stellar rotation vector perpendicular to it.

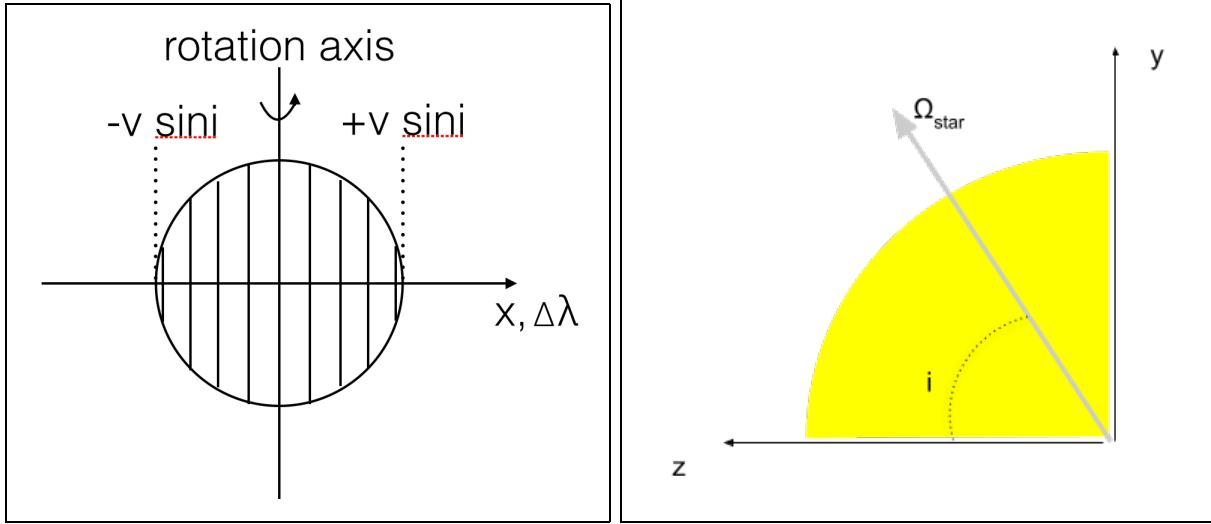


Figure 3.7: Left: Decomposition if the $v \sin i$ effect on stellar surface, Right: geometrical definition of the $v \sin i$ line broadening.

Considering a spectral profile $H(\lambda)$ the effect of adding up all stripes can be written as

$$\int_{-\infty}^{\infty} H(\lambda - \Delta\lambda) \cdot G(\Delta\lambda) d(\Delta\lambda), \quad (3.49)$$

whereby $G(\Delta\lambda)$ is a weighing function describing the change of the amplitude of the flux along the x direction (see also "sphere projection" and "limb darkening"). This equation expresses a convolution

$$H(\lambda) \star G(\lambda). \quad (3.50)$$

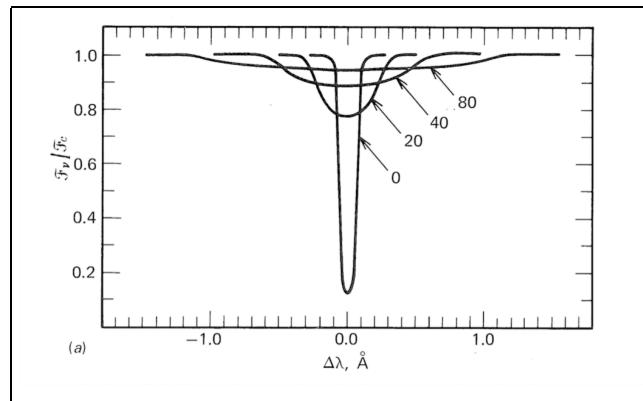


Figure 3.8: Example of stellar profile for various $v \sin i$ values

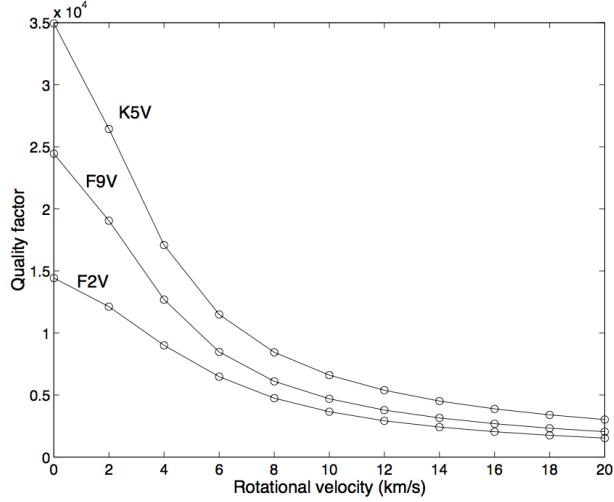


Figure 3.9: Effet of the $v \sin i$ on the quality factor for various stellar types (from Bouchy et al. 2001)

3.2.6 Measuring the mass of a planet

The detection of a RV period signal provides a direct measurement of $M_2 \sin i$. For $\sin i$ we have (see fig. 3.10 on the facing page)

$$P(i)di = 2\pi \sin i di, \quad (3.51)$$

$$\langle \sin i \rangle = \frac{\int_i P_i \sin i di}{\int_i P_i} = \frac{\pi}{4} \sim 1. \quad (3.52)$$

The fact the statistic of $v \sin i$ is severely skewed towards 1, from a statistical point of view, it is often implicitly assumed that $M_2 \sin i \sim M_2$.

With enough detection when many $M_2 \sin i$ are being measured, one can get a good perspective on the statistical distribution of the mass of planets. With only a single measurement of $M_2 \sin i$ available, one can usually assume that the mass is likely to be within a factor of 2-3 of this value. When considering the stellar rotation v (from rotation period detection) and measurement of the spectroscopic $v \sin i$, we can retrieve information about $\sin i$ itself and therewith potentially restrain the limits for $M_2 \sin i$ assuming both rotation axis (star and orbit) are parallel (that is not always true...).

Consequently the discovery of 51 Peg with $M_2 \sin i \approx 0.5M_J$ was well within planetary range and hence not consider suspicious on that matter. On the contrary, this was not the case when a companion to HD114762 was found with $M_2 \sin i = 11M_J$. We know from recent astrometric measurement that the companion orbiting the star HD114762 is stellar and not even a brown-dwarf. The confusion between high mass planet and brown dwarf is a recurrent theme of debate and poisoned the early days of planet discoveries. Luckily for planet interpretation brown-dwarfs are rare and the combined probability to pick a rare

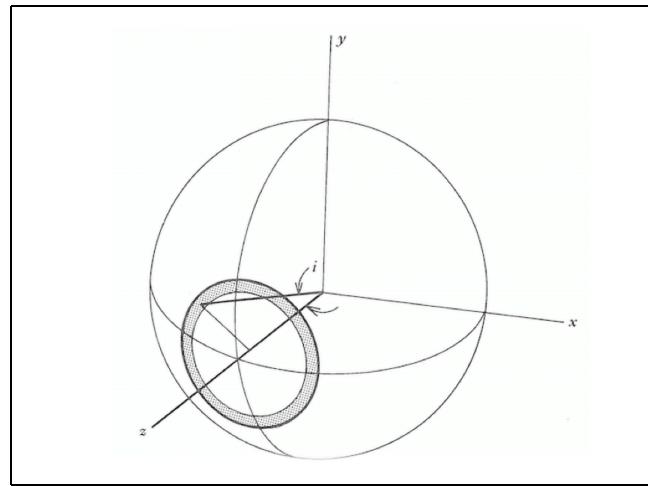


Figure 3.10: Geometrical description of the $v \sin i$ distribution

object with a rare statistical occurrence is almost negligible when one has companion of few Jupiter mass.

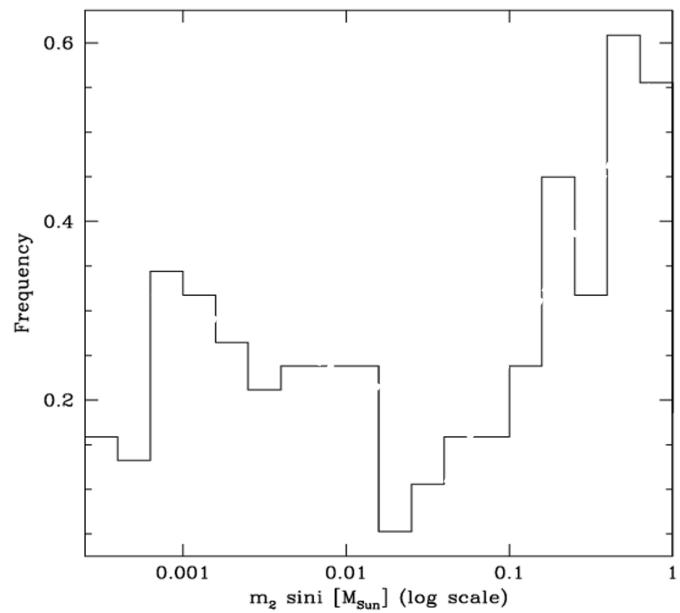


Figure 3.11: "historical mass function" in the range with giant planets, brown-dwarf ans stars

3.2.7 Additional noise sources

Instrumental "noise"

A spectrograph is an imaging device where the directional axis of the detector is "converted" into a wavelength scale. This means, any motion of the entrance slit of the detector will also be "converted" into a wavelength (see fig. 3.12). The effect of the geometrical projection of the entry slit into wavelength leads to "instrumental" radial effects.

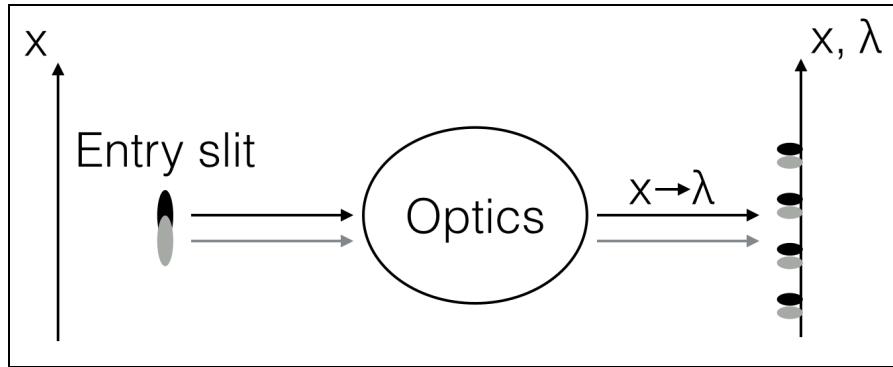


Figure 3.12: Schematics of the spatial conversion to wavelength occurring in a spectrograph

For a resolution of $R \sim 100.000$ one usually considers 2-3 sampling pixel (Nyquist theorem). This corresponds about to 1 km s^{-1} per pixel on the detector. With a pixel size of around $13 \mu\text{m}$, a doppler effect of 1 m s^{-1} corresponds to 13 nm . This indicates that to reach the precision necessary for the detection of Earth-like planets 10 cm s^{-1} , i.e. one needs a spatial accuracy the order of 1 nm . (An atom of silicon has a diameter of 0.2 nm . More details on nuts and bolts and tricks used to tame natural spectrograph tendency to drift away from its calibration zero point may be found in the following publication [Nobel Lecture: 51 Pegasi b and the exoplanet revolution](#)

Stellar photospheric noise

The stellar envelope of cool stars -like the Sun- is convective, meaning the photosphere is not stable. Convection is the main heat transportation mechanism. Acoustic modes can propagate (p-mode pulsations) on timescales of several minutes. About a million of gas cells in motion with typical vertical velocities of kilometre-per-second: hotter when emerging and cooler when sinking down generates photosphere granulation. In addition magnetic field (generated from the sheer motion of the convection) produces active regions on the photosphere (solar spots). Magnetic flux tubes form and decay on time-scales typically comparable to the stellar rotation period (few weeks) and long-term magnetic cycles modify the convection patterns. The combined result of all these effects, due to the convective nature of the stellar atmosphere, is to produce spectral lines of variable shape with an underlying

periodic pattern that is related to the rotation period and which evolves slowly over time. (see [\(See video\)](#))

Typically the following typical time-scale range may be identified on a Sun-like star (see fig. 3.13):

- short-timescales (about 5 min): P-mode $\sim \rho^{1/2}$
- medium-timescales (typically from 20 min to few hours) from the evolution granulation pattern structures (super cell)
- long-timescales (order of the day scale): stellar spots and rotation
- very long timescales (month to years): stellar cycle

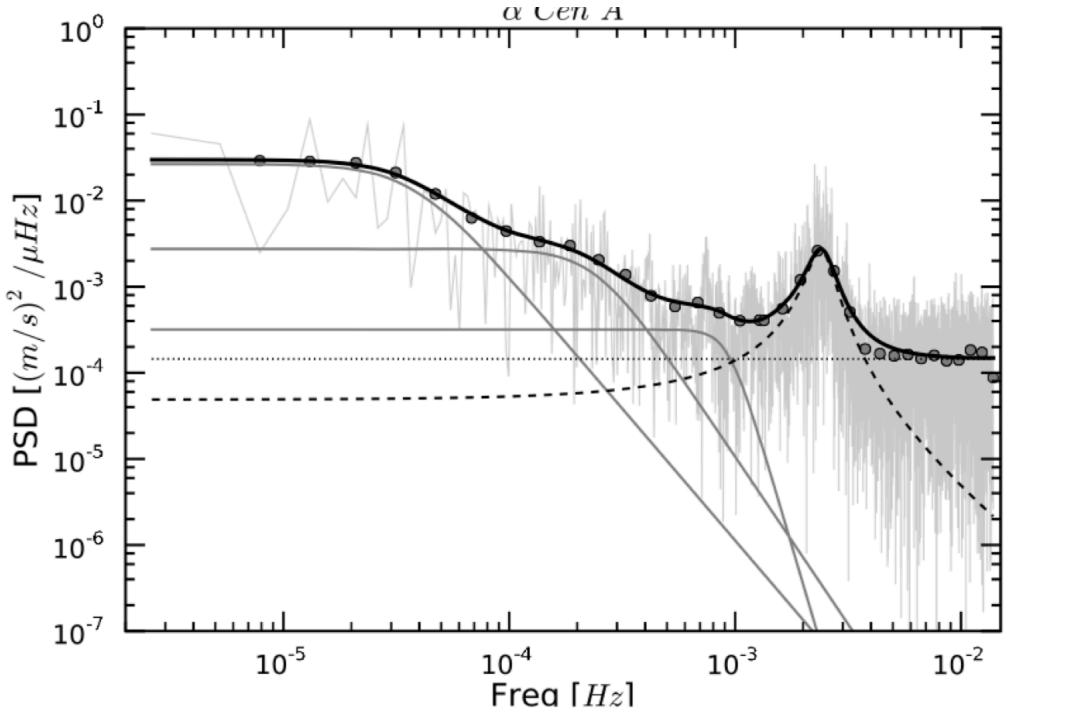


Figure 3.13: Power spectral distribution of long sequence of radial velocity measurements of the Sun-like star α -Cen A

In the particular case of stellar spots can modify $G(\Delta\lambda)$ their seemingly occultation effect of shape of the line profile depends on their size and $v \sin i$ of the star. It has an effect of a few m/s on the radial velocity (mean center gravity of line profile), with a distinctive periodic rotation pattern corresponding to the stellar rotation. (see fig. [3.14 on the next page](#) for display)

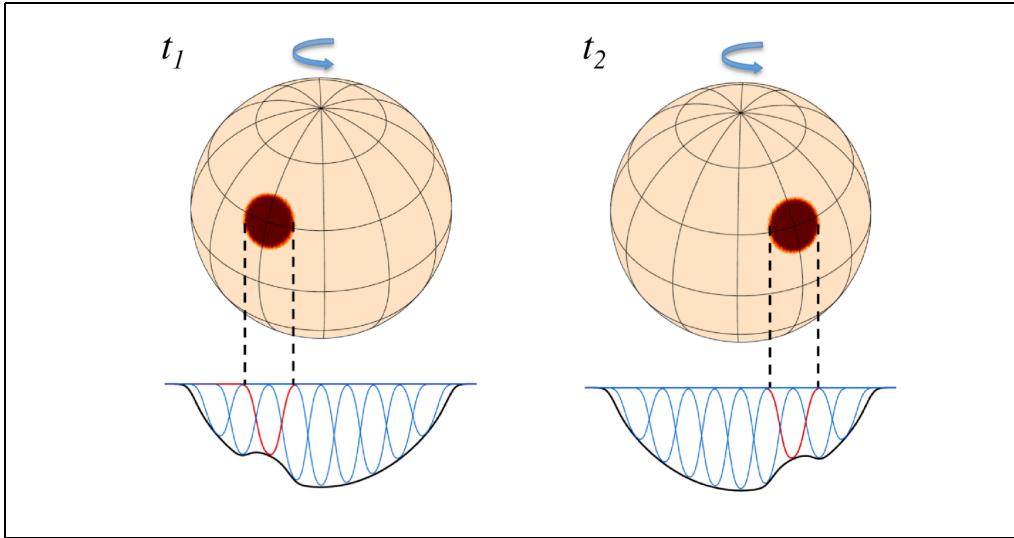


Figure 3.14: Effect of stellar spot on the the line spectral profile

3.3 Exoplanet transits and occultations

The orientation of Ω (see section 3.1.3 on page 44) is not relevant for the transit, the phenomenon is symmetric around the line of sight axis (z-axis). It is usual to place the ascending node $\Omega = 0$ of the stellar orbit on the X-axis. In that case the corresponding symmetric planet orbit *descending node* is at $\Omega = 180$ and can be described as

$$X = -r \cos(\omega + f) \quad (3.53)$$

$$Y = -r \sin(\omega + f) \cdot \cos i \quad (3.54)$$

$$Z = r \sin(\omega + f) \cdot \sin i \quad (3.55)$$

The eclipse occurs by definition around the conjunction where $x = 0$ when the orbit of the **planet** crosses the line of sight. We can identify two cases: (1) *The inferior conjunction* when the planet is in front the star - a transit occurs; (2) *The superior conjunction* when the star is in front the planet - an occultation occurs. One can write this as

$$f_{transit} = \frac{\pi}{2} - \omega, (z = r) \quad (3.56)$$

$$f_{occultation} = -\frac{\pi}{2} - \omega, (z = -r). \quad (3.57)$$

3.3.1 Impact parameter

To determine where the transit occurs one defines the impact parameter

$$b_{\text{tra}} = \frac{y}{R_*} \Big|_{x=0} \quad (3.58)$$

$$= \frac{a \cos i}{R_*} \left(\frac{1 - e^2}{1 + e \sin \omega} \right), \quad (3.59)$$

since (see Fig. 3.3 on page 45 and Equ. 3.9 on page 42)

$$y_{\text{tra}} = r_{\text{tra}} \cos i, \quad r_{\text{tra}} = \frac{a(1 - e^2)}{1 + e \sin \omega}. \quad (3.60)$$

3.3.2 Transit probability

A transit occurs if the observer is in the same plane as the star-planet system. With this configuration one can define a maximum configuration angle Θ , such as

$$\sin \Theta = \frac{R_* \pm R_p}{r_{\text{tra}}}, \quad (3.61)$$

corresponding to a geometrical configuration displayed in fig. 3.15.

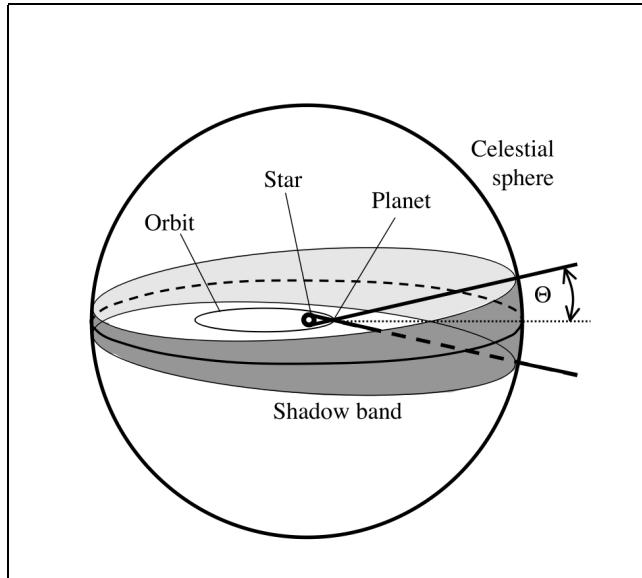


Figure 3.15: Geometrical configuration when a transit occurs

Averaging over a random distribution of ω leads to

$$\langle P \rangle = \int_0^{2\pi} \sin \Theta d\omega, \quad (3.62)$$

then using Equ. 3.60 on the preceding page and Equ. 3.61 on the previous page one finds

$$\langle P \rangle = \frac{R_* \mp R_p}{a} \frac{1}{1 - e^2} \quad (3.63)$$

$$\approx \frac{R_*}{a} \frac{1}{1 - e^2} \quad (3.64)$$

With 1 AU corresponding to about $150 \cdot 10^6$ km and diameter of the Sun $1.4 \cdot 10^6$ km one finds 1 AU is about $200 R_*$. The probability to detect a transit of a planet 1 AU is 0.5% and at 0.05 AU (51 Peg) 10%. Discoveries of short period planets acted as a trigger to promote development of transit search programs.

3.3.3 Transit timing

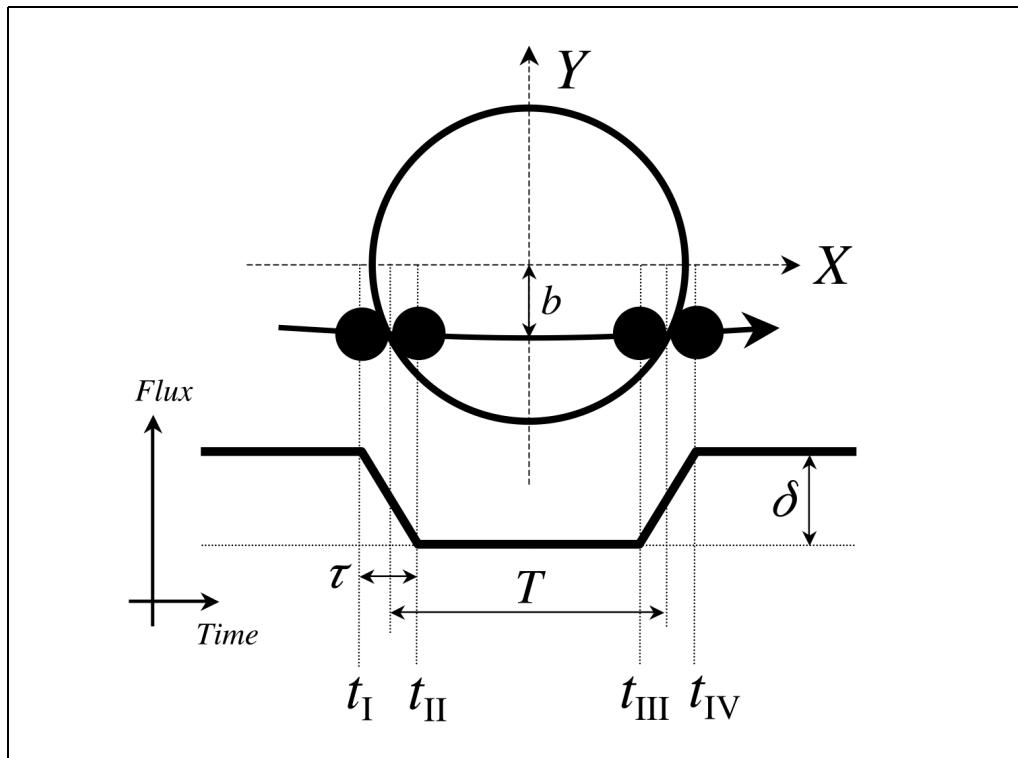


Figure 3.16: geometry of the transit

On see from Fig. 3.16 we can compute the x-axis locations of the significant points, e.g.

$$x_{I,IV} = x \mp R_* \sqrt{\left(1 + \frac{R_p}{R_*}\right)^2 - b^2} \quad (3.65)$$

The mid-transit duration ΔT (T on the figure) can be calculated as

$$\Delta T = \int_{t_i}^{t_e} dt = \int_{f_i}^{f_e} \frac{r^2}{h} df, \quad (3.66)$$

Using Eq. (3.20 on page 43) with suffixes *ig* and *eg* denote 'ingress' and 'egress', we get

$$\Delta T = \frac{P}{2\pi\sqrt{1-e^2}} \int_{f_{ig}}^{f_{eg}} \left(\frac{r(f)}{a} \right)^2 df. \quad (3.67)$$

We now assume a circular orbit. In that case $\Delta T = P(2\pi)^{-1}\Delta f$. Further with $f_{tra} = 0$ and $\omega = \frac{\pi}{2}$ one has the following configuration:

$$x(t) = -a \cos(f + \frac{\pi}{2}) \quad (3.68)$$

$$= a \sin f. \quad (3.69)$$

With respectively the in-transit projection of the motion Δx and Δf its angular distance, we can further express

$$\Delta x = 2R_*\sqrt{1-b^2} \quad (3.70)$$

or expressed with true anomaly f

$$\Delta f = \sin^{-1} \left(\frac{2R_*}{a} \sqrt{1-b^2} \right). \quad (3.71)$$

Using these expressions, we find:

$$\Delta T = \frac{P}{2\pi} \Delta f = \frac{P}{2\pi} \sin^{-1} \left(\frac{2R_*}{a} \sqrt{1-b^2} \right) \quad (3.72)$$

With the assumption $R_*/a \ll 1$ one finally finds

$$\Delta T = \frac{R_* P}{\pi a} \sqrt{1-b^2} \quad (3.73)$$

$$= T_0 \sqrt{1-b^2} \quad (3.74)$$

with transit time duration scale

$$T_0 = \frac{R_* P}{\pi a}. \quad (3.75)$$

We know that a and P are connected by Kepler's third law such as $a \sim M_*^{1/3} P^{2/3}$, then we have $T_0 \sim \rho_*^{-1/3} P^{1/3}$. The transit duration is hence a measure of the stellar density. This property combined with a prior information on the stellar mass and radius can be used to constraint transit solutions. Note there is no information about the planet size or mass in the equation. The transit duration scale is essentially a property of type of star and period of the orbit.

3.3.4 Transit light curves

The observation of a transit light curve a rich information on the star-planet system: the transit the (secondary) eclipses and the phase-function. Eclipses are the only moment when only the stellar flux star is (during the occultation -or eclipse- of the planet), otherwise we measure a combination of flux from the star and from the planet (see Fig. 3.17).

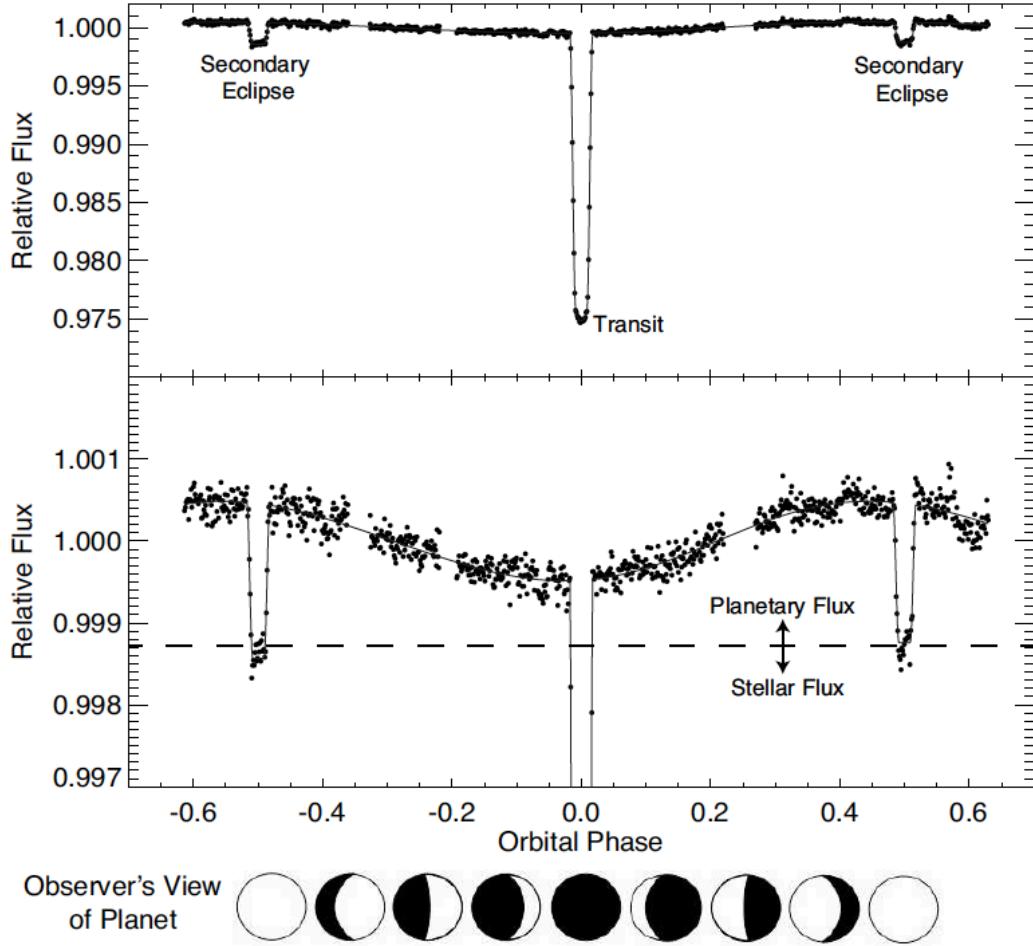


Figure 3.17: stellar light-curve of a transiting system

We can describes observed the light-curve with the following equation:

$$F(t) = F_*(t) + F_p(t) - \begin{bmatrix} k^2 \alpha_{tr}(t) F_* \\ 0 \\ \alpha_{oc}(t) F_p \end{bmatrix} \begin{array}{l} (transit) \\ (eclipse) \end{array} \quad (3.76)$$

The parameter $\alpha(t) \in [0, 1]$ is describing the functional shape of the transit or occultation.

At mid-transit time $\alpha = 1$ (with the exception of an extremely grazing transit) and out-of-transit $\alpha = 0$. The parameter k describes what fraction of the area (and hence flux) of the star the planet blends out, and is given by

$$k = \frac{R_p}{R_*}. \quad (3.77)$$

The emerging flux F and the intensity I for planet or star component is related by $F \sim R^2 I$, such that we get

$$\frac{F_p}{F_*} = k^2 \frac{I_p}{I_*}. \quad (3.78)$$

The measured light curve of a transit, it is usually normalised to 1, such that we usually refer the normalised flux using $\langle F(t) \rangle \approx F_\star$:

$$\frac{F(t)}{F_\star} = 1 + \frac{k^2 I_p(t)}{I_*} - k^2 \begin{bmatrix} \alpha_{\text{tr}}(t) \\ 0 \\ \frac{I_p}{I_*} \alpha_{\text{oc}}(t) \end{bmatrix} \begin{cases} \text{(transit)} \\ \text{(eclipse)} \end{cases}. \quad (3.79)$$

At the "deepest depth" of the transit $\alpha(t_{[\text{tr}]}) = 1$ we get the transit depth:

$$1 - \frac{F(t_{[\text{tr}]})}{F_\star} = k^2 \left(1 - \frac{I_p(t_{[\text{tr}]})}{I_*} \right) \approx \delta_{\text{tr}}, \quad (3.80)$$

whereby δ_{tr} is the transit depth, i.e. the value that is commonly referred to when working in the field of exoplanet transits.

Similarly one can define the analogue for the occultation and one gets

$$\delta_{\text{oc}} = k^2 \frac{I_p(t_{[\text{loc}]})}{I_*}, \quad (3.81)$$

corresponding to a direct measurement of the emerging intensity radiation from the planet atmosphere.

3.3.5 Transit shape and limb darkening

The α function in the case of transit contains physical information about the stellar photosphere and detail geometry of the impact. It primarily depends on the impact parameter b and on the change of apparent brightness of the star across its diameter the 'limb darkening' that is wavelength dependant. Example of various transit shape and limb darkening effects are displayed on Fig. 3.18 on the next page. To the second order it also depends on the detail atmosphere of the planet. This effect will be described further in Sect. 3.3.9 on page 64.

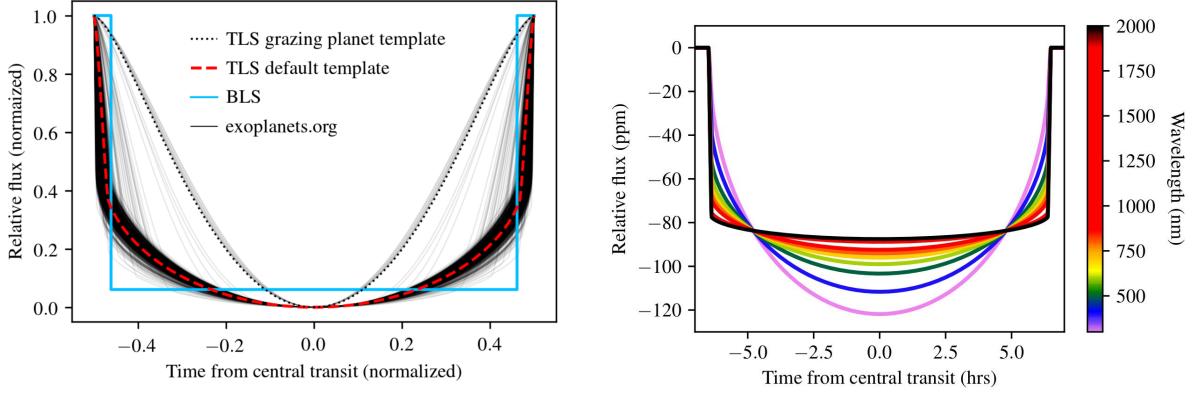


Figure 3.18: **Left:** Display of various transit shape $|\alpha - 1|$. Composite of 2346 transiting Kepler planets from the Exoplanet Orbit Database (Hippke 2019); **Right:** colour effect on the limb darkening

3.3.6 Impact parameter (cont'd)

We established earlier that $\Delta T = T_0 \sqrt{1 - b^2}$ and T_0 depends on stellar density. So technically the equation may be solved. Additional information to further constrain the impact parameter b may be retrieved from the geometry of the transit itself displayed on Fig. 3.17 on page 60 with the assumption the speed of transit corde-crossing is constant (case $e = 0$).

$$\frac{2R_*\sqrt{1 - b^2}}{\Delta T} = \frac{2R_p}{\tau} \quad (3.82)$$

By reorganizing the terms one obtains

$$b^2 + \left(\frac{R_p}{R_*}\right)^2 \left(\frac{\Delta T}{\tau}\right)^2 = 1 \quad (3.83)$$

$$\delta \cdot \left(\frac{\Delta T}{\tau}\right)^2 = 1 - b^2. \quad (3.84)$$

Not that we can rearrange the equation and substitute $\sqrt{1 - b^2}$ in Eq. 3.75 on page 59 and we obtain a simple relation relating τ and δ .

$$\tau = T_0 \sqrt{\delta} \quad (3.85)$$

3.3.7 Secondary eclipses

Another useful parameter when investigating the orbit of the planet can be given by secondary eclipses. By integrating dt/df (see eq. Eq. 3.20 on page 43) between two conjunctions, one obtains the following approximate expression for the time difference Δt_c between

the two conjunctions

$$\Delta t_c \approx \frac{P}{2}(1 + 4\pi^{-1}e \cos \omega). \quad (3.86)$$

The time between transit and occultation gives us a direct measure of $e \cos \omega$.

Similarly, we can also study the ratio of the transit and occultation duration to find additional information about the orbit:

$$\frac{\Delta T_{tra}}{\Delta T_{occ}} \approx 1 + e \sin \omega. \quad (3.87)$$

However, this measurement generally is not practicable, since T_{occ} can currently not be measured with high enough accuracy. It is more practicable to rely on the first method, especially since the measurement error for $\Delta t_c/P$ is expected to be significantly smaller. Note also that in neither case we can extract direct information about e but a combination of e and ω .

3.3.8 Transmission spectroscopy

The edge of a planet with a significant atmosphere is not sharp, but fuzzy due to the wavelength dependent transmission of the atmosphere. If we look at single wavelengths (e.g. using spectroscopy) we can gain important information about the atmosphere of the planet. The change of the transit depth can approximately be written as

$$\Delta\delta = \frac{(R_p + n \cdot H)^2}{R_*^2} - \frac{R_p^2}{R_*^2}. \quad (3.88)$$

H describes the atmospheric scale height (varying with λ) and n is a multiplication factor for the scale height (usually set to few). The atmospheric scale height is given by the hydrostatic equilibrium (see to Sect. 2.5.1). Similarly assuming a ideal gas H may be expressed as

$$H = \frac{k_B T}{\mu_m g}, \quad (3.89)$$

where μ_m is the mean molecular mass and g the gravitational acceleration of the planet.

The effect of the atmosphere is high if H is large, e.g. for hot planets and low surface densities. Hot Jupiters typically planet with large H value. The transit depth variation due to their atmosphere can be on the order of $\Delta\delta \sim 10^{-4}$, corresponding to relative planet radius variation of the order of 1% with observation wavelength considered. See for example the multi-wavelength observation of the Wasp-137 transiting hot Jupiter in fig. 3.19 on the following page.

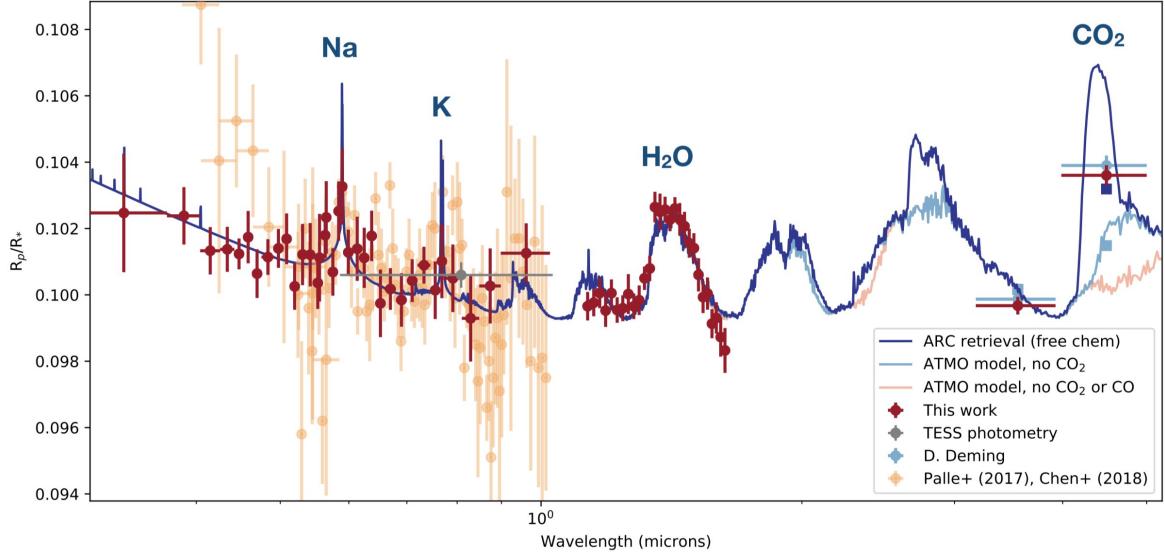


Figure 3.19: Transit transmission spectrum for WASP-127b (Sparke 2021). Points are measurements. Solid lines display best model fits according to retrieval of atmosphere composition.

3.3.9 Occultation spectroscopy

Assuming a pure thermal regime emission (no stellar reflection), we can approximate the planet and star as blackbodies, such that $I \sim B(T)$. The occultation depth (see Eq. 3.81) can then be written as

$$\delta_{\text{oc}} = k^2 \frac{B_\lambda(T_P)}{B_\lambda(T_*)} \approx k^2 \frac{T_P}{T_*}, \quad (3.90)$$

whereby we made use of the Rayleigh-Jeans limit (for regimes where $\lambda \gg hc/kT$) for convenience but the ratio of blackbodies remains valid in any case. In that regime when the occultation is observed it is actually a direct measurement of the thermal emission of the planetary atmosphere.

Assuming at the contrary that the emerging flux from the planet is purely due to reflection of stellar light, one introduces the geometric albedo A_λ to quantify the efficiency of the reflection and we can calculate

$$L_p = \frac{L_*}{4\pi a^2} \pi R_p^2 A, \quad (3.91)$$

with $L_p = I_p 4\pi R_p^2$ and $L_* = I_* 4\pi R_*^2$ respectively the luminosity (total energy emitted per unit of time) of the planet and the star. One obtains $I_p 4a^2 = I_* R_*^2 A_\lambda$ and therefore

$$\frac{I_p}{I_*} = A_\lambda \left(\frac{R_*}{2a} \right)^2. \quad (3.92)$$

Using the definition of the occultation depth Eq. 3.81) one finally obtains

$$\delta_{\text{oc}} = A_\lambda \left(\frac{R_p}{2a} \right)^2 \quad (3.93)$$

The detection (essentially in the visible light) of stellar light reflection by the planet is mostly done only when the planet is large and orbiting close to its star. A simulation of JWST observation with both effect considered is displayed on Fig. 3.20.

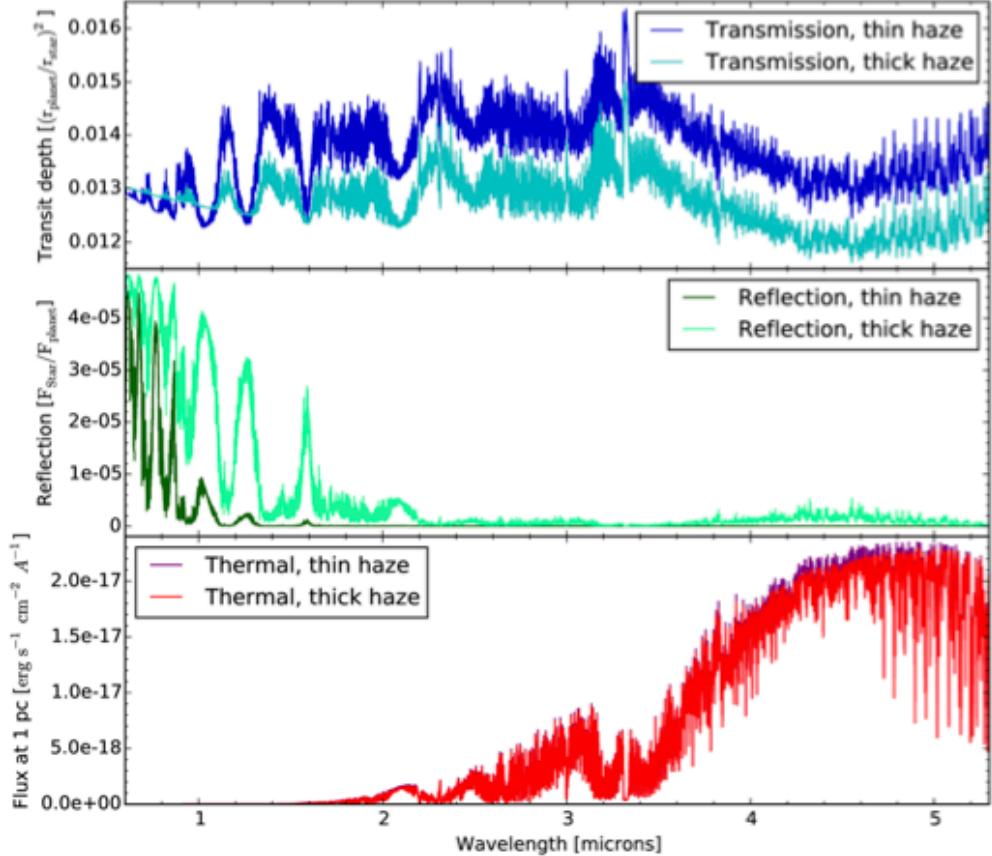


Figure 3.20: Synthetical planetary spectra for the warm super-Earth GJ1214b, $R=4000$, for optically thin and thick haze. The transit depth and reflection of starlight are both given as the contrast to the host star. The thermal emission spectrum is the flux received at a distance of 1 pc (Nielsen 2016)

3.3.10 The Rossiter-McLaughlin effect

A transit observed in spectroscopy with enough spectral resolution will affect the shape of the spectral line profile and will create an anomalies that can be detected. It is called the Rossiter-MacLaughlin effect. It was firstly observed on 1924 on eclipsing binaries.

It is somewhat similar to the effect of a rotating "dark" spot on the stellar photosphere (display on Fig. fig: 3.14 on page 56), but in the case of the transit event it is due to the crossing of an object in front of the stellar disk. Considering the effect of the rotation of the star on its photospheric emission, light emerging from the approaching half of the stellar disk is practically blue-shifted and light from the receding half part is red shifted. When the planet covers parts of the blue-shifted side, the overall emerging flux will be changed and stellar emission will look slightly redshifted - and vice versa in the other case. The character and time-evolution of the spectral distortions depend on stellar $v \sin i$ and angle λ , in addition to the usual eclipse parameters. Fig. 3.21 illustrates the transit geometry and its schematic corresponding effect.

7

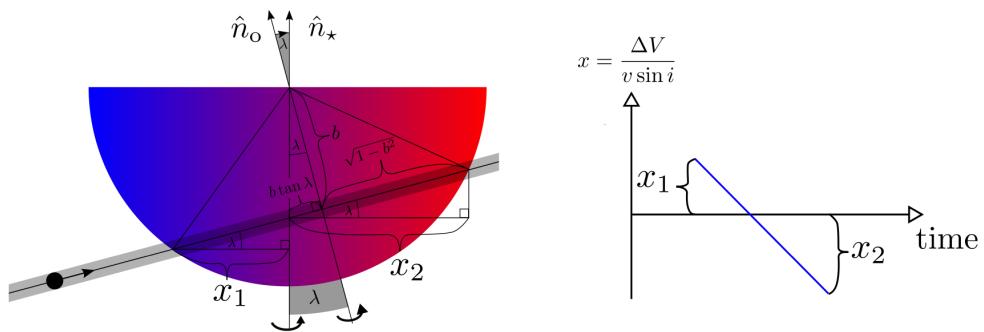


Figure 3.21: Geometry of the Rossiter-McLaughlin effect and corresponding angles

with $\Delta V = v \sin i x(t)$, where $x(t)$ is the planets cord trajectory in units of the stellar radius along the coordinate axis running perpendicular to the stars projected rotation axis. The extremes of the signal occur at ingress x_1 and egress x_2 , with amplitudes $\Delta V_1 = v \sin i x_1(t)$ and $\Delta V_2 = v \sin i x_2(t)$. Based on the transit geometry display in Fig. 3.21

$$x_1 = \sqrt{1 - b^2} \cos \lambda - b \sin \lambda \quad (3.94)$$

$$x_2 = \sqrt{1 - b^2} \cos \lambda + b \sin \lambda \quad (3.95)$$

allowing us to write

$$\Delta V_2 - \Delta V_1 = 2 v \sin i \sin \lambda b \quad (3.96)$$

$$\Delta V_2 + \Delta V_1 = 2 v \sin i \cos \lambda \sqrt{1 - b^2} \quad (3.97)$$

This system of equations makes clear that the asymmetry of the signal depends on $\sin \lambda$, while the amplitude depends on $\cos \lambda$. It also indicates that measurements of both the amplitude and the asymmetry are formally sufficient to compute $v \sin i$ and λ , as long as b is not too close to 0 or 1.

In term of geometry of the crossing one should bear in mind than λ is a projected angle and an addition angle ψ - the obliquity- is needed to obtain a full characteristics of the system. Using the Fig. 3.22 as a reference many authors define a Cartesian coordinate system with $\hat{z} = \hat{n}_{obs}$ and y aligned with the sky projection of \hat{n}_0 , although it is sometimes more convenient to align y with the sky projection of \hat{n}_* .

$$\hat{n}_* \cdot \hat{n}_0 = \cos \psi = \cos i \cos i_0 + \sin i \sin i_0 \lambda \quad (3.98)$$

$$(\hat{n}_* \times \hat{n}_0) = \sin i \sin i_0 \sin \lambda \quad (3.99)$$

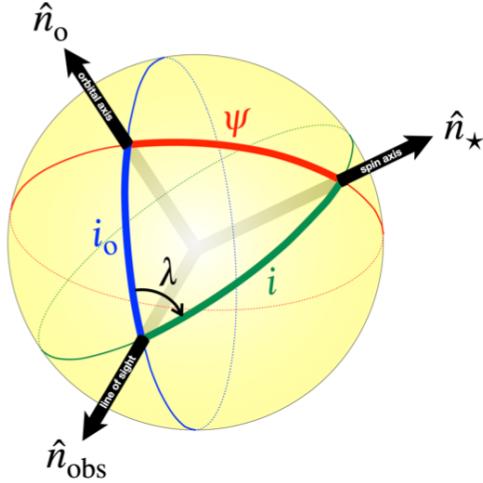


Figure 3.22: Angles that specify the orientation of the spin and orbital angular momentum vectors. The obliquity is ψ , the orbital inclination is i_0 , and the inclination of the stellar rotation axis is i

The effect on spectral lines is a distortion, not an overall Doppler shift. Nevertheless, a radial-velocity extraction algorithm will respond to the distortion by reporting an

A decent approximation for the maximum amplitude of the anomalous RV is

$$\Delta v_{RM} \approx 0.7k^2\sqrt{1 - b^2}(v_* \sin i_*). \quad (3.100)$$

The factor of 0.7 accounts for limb darkening. Fig.

With this technics it has been found that planets are not necessarily aligned. Some short-period giant planets have been found on prograde, polar, and retrograde orbits. It seems likely that dynamical processes such as planet-planet scattering and secular perturbations are responsible for tilting the orbits of close-in giant planets. These results suggest it is unsafe to assume that stars and their protoplanetary disks are aligned. Primordial misalignments might even be considered by neighboring stars or more complex events that occur during the epoch of planet formation.

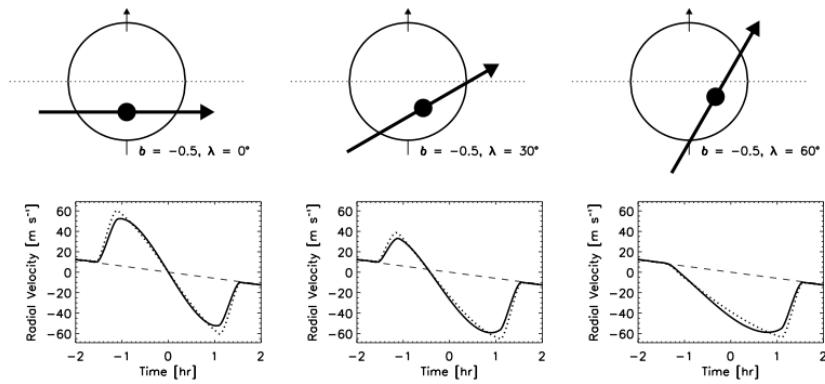


Figure 3.23: Effect on spectral line profil of a transiting planet