Planet Formation
Is the Solar System Misleading?

Günther Wuchterl
Max-Planck-Institute for extraterrestrial Physics
Giessenbachstraße, D–85748 Garching
wuchterl@mpe.mpg.de, http://www.xray.mpe.mpg.de/~wuchterl/

Abstract
The discovery of more than hundred extrasolar planet candidates challenges our understanding of planet formation. Do we have to modify the theories that were mostly developed for the solar system in order to understand giant planets orbiting their host stars with periods of a few days? Or do we have to assume particular circumstances for the formation of the Sun to understand the special properties of the solar system planets? I will review the theories of planet formation and outline processes that may be responsible for the diversity of planetary systems in general. Finally I will discuss two questions raised by extrasolar planets: (1) the formation of Pegasi-planets and (2) the relation between discovered extrasolar planets and the metallicity of their host stars.

1 Introduction
Following the discovery of the first dozens of extrasolar planets, voices have been raised that the theories developed for the formation of the solar system were not relevant for planet formation in general. New processes were invoked to explain the unexpected diversity in the observed exoplanet population. Some of the new processes apparently played no or only a minor role in the solar system. The unpredicted properties of extrasolar planets — giant planets in orbits with periods of a few days, eccentricities much larger then any of the Sun’s planets, and many giant planets with orbital radii much smaller than Jupiter’s — prompted strong comments:

Upon the discoveries, theorists have lost the understanding of the formation of the solar system,

as Pavel Artymovic put it at the IAU symposium 202 at the Manchester general assembly or more drastically:

Forget the solar system!,

as a well known astrophysicist recommended off the records. Are we witnessing that the copernican principle is finally failing in cosmogony? Is the formation of our own home-system fundamentally different from the typical
planet formation process in the galaxy? Are special processes or unlikely circumstances required to explain the properties of the Solar System? I will review the present state of the theory of planet formation and confront it to observations in two cases: (1) the planet-metallicity correlation and (2) the formation of Pegasi-planets, i.e. giant planets in short period orbits. I will focus on the more detectable giant planets and refer to Kokubo’s (2001) recent review for a beautiful, more detailed discussion of the terrestrials.

2 The Solar System

More planetary candidates are now known in our galactic neighborhood than planets orbit our star. But the Solar System is by far the best studied and most completely known system. Masses, radii and composition of the major and some minor bodies are known. Almost a million orbital elements allow a detailed study of the dynamical and stability properties of the system (Lecar et al. 2000). The information about the Sun, the planets and their satellite systems provides a clear picture of the angular momentum distribution. Interior structures and surface properties contain a record of the formation history. The heat budgets of the giant planets show an excess of emitted radiation over absorbed sunlight (except for Uranus where only an upper limit is available). These faint intrinsic luminosities, in the $nL \odot$ range ($1 nL \odot = 10^{-9}$ solar luminosities) are directly related to the formation process that stored the heat 4.5 Ga ago ($1 Ga = 10^9$ years). The solid, icy and rocky surfaces of terrestrial planets and many satellites show a record of impacts that samples a time-range reaching from the present, e.g. the surface of Io, back to the formation epoch, e.g. in the lunar highlands. Radioactive dating provides absolute ages for the Earth, Moon and meteorites. The latter provide an accurate age of 4.565±1 Ga for the oldest nebula condensates, the Calcium-Aluminium rich inclusions in primitive meteorites, most prominently, Allende (Allègre et al. 1995). Ongoing processes as orbital evolution, asteroidal collisions and even large impacts, as those of the fragments of comet Shoemaker-Levy 9 onto Jupiter can be studied directly and in great detail. Finally there is abundant information gathered by fly-by space-craft missions as the Voyagers, orbiters as Galileo, Mars-Odyssey and Mars-Express, as well as from in-situ exploration by landers and atmospheric entry probes for the Moon, Venus, Mars and Jupiter. An atmospheric entry of the Huyghens probe, with a possible successful landing on Saturn’s moon Titan, is scheduled for 2005.

Our four giant planets contain 99.5% of the angular momentum of the Solar System, but only 0.13% of its mass. Terrestrial planets contribute another 0.16 %, to the angular momentum. On the other hand, more than 99.5% of the mass and thermal energy of the planetary system is in the four largest bodies, with the remaining 0.5 % mostly in the second group of four, the terrestrial planets. Modern models of the interiors and evolution of giant planets in our Solar System account for the high pressure properties of hydrogen, helium and the heavier elements as well as energy transfer by radiation and convection.
When fit to the observed global properties of Jupiter at an age of 4.5 Ga they show that 10–42 of the 318 Earth-masses are due to heavy elements. That corresponds to 3 to 13 %. The respective mass fractions implied for the other solar system planets are all higher. That points to a bulk enrichment of heavy elements more than a factor of two above solar composition and implies heavy element cores ranging from greater than one Earth mass to a considerable fraction of the total mass (Guillot 1999, Wuchterl, Guillot and Lissauer 2000). The heavy element enrichment is even more obvious in terrestrial planets that may be also viewed as failed cores of giant planets. Hydrogen and helium that constitute for 98% of the Sun’s mass and between 87% and 15% of the giants is a minor constituent of terrestrial planets. Forming planets enriches heavy elements relative to the central star, that is formed from the same protostellar cloud. Such an extensive enrichment is not predicted by any mechanism proposed for the formation of stars and brown dwarfs. That alone is already indicating that planet formation is fundamentally different from star formation.

3 Solar System Formation

The distribution of mass and angular momentum in the Solar System can be understood on the basis of the nebula hypothesis (Kant 1755). The nebula hypothesis assumes concurrent formation of a planetary system and a star from a centrifugally supported flattened disk of gas and dust with a pressure supported, central condensation (Laplace 1796, Safronov 1969, Lissauer 1993). Flattened preplanetary nebula disks explain the coplanarity and circularity of planetary orbits by the respective properties of the parent disk. Theoretical models of the collapse of slowly rotating molecular cloud cores have demonstrated that such preplanetary nebulae are the consequence of the observed cloud core conditions and the dynamics of radiating fluids, provided there is a macroscopic angular momentum transfer process (Cassen and Moosman 1981, Morfill et al. 1985). Assuming turbulent viscosity to be that process, dynamical models have shown how mass and angular momentum separate by accretion through a viscous disk onto a growing central protostar (Tscharnuter 1987, Tscharnuter and Boss 1993). Such cloud collapse calculations, however, still do not reach to the evolutionary state of the nebula where planet formation is expected. Observationally inferred disk sizes and masses are overlapping theoretical expectations and confirm the nebula hypothesis.

High resolution observations at millimeter-wavelengths are sensitive to disk conditions at orbital distances > 50 AU. However, observations thus far provide little information about the physical conditions in the respective nebulae on scales of 1 to 40 AU, where planet formation is expected to occur. Planet formation studies therefore obtain plausible values for disk conditions from nebulae that are reconstructed from the present planetary system and disk physics. The so obtained minimum reconstituted nebula masses defined as the total mass of solar composition material needed to provide the
observed planetary/satellite masses and compositions by condensation and accumulation,

are a few percent of the central body for the solar nebula and the circumplanetary protosatellite nebulae (Kusaka et al. 1970, Hayashi 1980, Stevenson 1982a). See Lunine et al. (2004) for a more detailed discussion of reconstructing the preplanetary nebula from observational constraints provided by solar system data.

4 Planet Formation — the Problem

Giant planet formation requires (1) a compression of the solar nebula gas by about 10 orders of magnitude to form a gaseous condensation held together by its own gravity, at Jupiter’s present mean density of \(1.33 \times 10^3 \text{ kg/m}^3\), and (2) an enrichment of the heavy elements — that are condensable in the nebula — by typically a factor of at least three above the nebula value, most likely with a substantial fraction contained in a core. Gas in the midplane of a minimum mass solar nebula typically has a density of \(10^{-8} \text{ kg/m}^3\) at Jupiter’s present orbital radius (Hayashi et al. 1985) and a temperature around 100 K. The nebula gas pressure, the young Sun’s tides and the radially decreasing orbital velocities in a circumstellar disk, that shows an almost ‘keplerian’ shear, counteract the compressing force of nebula-gas self-gravity. Accordingly most circumstellar nebulae — modelled and observed — are gravitationally stable. Unlike interstellar clouds, larger mass fragments in a circumstellar disk are generally not more unstable, because larger mass fragments at given nebula densities require larger scales that are subject to stabilisation by the stellar tidal pull and the keplerian shear. In addition to the above mechanical barriers against gravitational self-compression of nebula gas, there is a thermal barrier. A Jupiter mass fragment is optically thick, i.e. optical depth, \(\tau > 1\), for a blob of mass \(M\), even at unperturbed nebula densities, \(\rho_{\text{Neb}}\), and for typical (dust) opacity, \(\kappa\):

\[
\tau = 36 \left( \frac{M}{M_{\text{Jupiter}}} \right)^{1/3} \left( \frac{\kappa}{0.01 \text{ m}^2/\text{kg}} \right) \left( \frac{\rho_{\text{Neb}}}{10^{-8} \text{ kg/m}^3} \right)^{2/3}.
\]

Any rapid, i.e. dynamical compression under such conditions will result in a temperature-increase determined by the efficiency of transfer processes and a much stronger counteracting pressure than in a simple isothermal analogue of protostellar collapse scaled down to planetary masses. In the stellar case compressional heat would leak out as fast as it is produced keeping the parent cloud isothermal during many orders of magnitude in compression (\(\sim 10\) for the collapse of a Jeans-critical solar mass). The rapid compression under optically thick conditions in the planetary case produces a thermal pressure increase, that typically leads to a slow down of compression from the dynamical time-scale of the fragment, a few years, to the thermal (cooling) timescale which is found to be of the order of million years in detailed models. Compression of the nebula gas — and therefore giant planet-formation,
their mass-growth and evolution — is then controlled by the heat loss of the fragment or protoplanet (e.g. Safronov and Ruskol 1982). Collapse, i.e. fast, gravity-driven compression in an essentially free-falling manner, with pressure playing a negligible role is a very unlikely event under such conditions.

The physical nature of giant planet formation — collapse, thermally controlled quasi-hydrostatic contraction or static accumulation — is decided by the dynamical stability of the nebula and the pressure build-up inside the protoplanets regulated by the thermal budget of the protoplanetary envelopes. The thermal budget is determined by heating due to contraction of the gaseous envelopes and dissipation of kinetic energy at planetesimal impacts and ‘cooling’ due to energy transfer to the ambient nebula by radiation and convection.

5 How to Compress by $10^{10}$?

The transition from dilute, weakly gravitating nebula conditions to compact planets with a spherical shape rounded by self-gravitation involves a compression of ten orders of magnitude. The nebula gas apparently had to be compressed by a macroscopic process from the earliest stages of planetary growth to the final planetary densities of $\sim 1000 \text{ kg/m}^3$. For stars, the runaway of the Jeans-instability easily multiplied the factor $10^{10}$ to the original cloud density, only inhibited to proceed further by the centrifugal force forming the very nebula. But unlike the protostellar collapse the planetary compression process cannot be analogous for the remaining $10^{10}$ from the nebula to the final planetary densities because it has to enrich the condensible material at the same time.

Since the 1970s two hypothesis have been discussed that try to account for nebula gas compression and condensible element enrichment. The gravitational disk instability hypothesis tries to find a nebula-analogue of the gravitational ‘Jeans’-instability of star formation and the nucleated instability hypothesis aims to explain giant planet formation as the consequence of the formation of solid, condensable element planetary embryos that act as gravitational seeds for nebula gas capture/condensation. The disk-instability hypothesis requires nebulae that undergo self-compression in a dynamically unstable situation and lead to a transition from a smooth regular disk to an ensemble of clumps in orbit around the Sun. Such clumps may be regarded as candidate precursors of protoplanets. The nucleated instability model looks at giant planet formation as a second step of rocky planet formation. A terrestrial planet embryo acts as a gravitating seed to permanently bind nebula gas thus forming a massive gaseous envelope around a condensible element core.

The key problem for both compression processes is that preplanetary disks are only weakly self-gravitating equilibrium structures supported by centrifugal forces augmented by gas pressure (Hollenbach et al. 2000, Calvet et al. 2000, and Beckwith et al. 2000 in PPIV). Any isolated, orbiting object below the Roche density is pulled apart by the stellar tides. Typical nebula densities are more than two orders of magnitude below the Roche density, so
compression is needed to confine a condensation of mass $M$ inside its tidal or Hill-radius at orbital distance $a$:

$$R_T = a \left( \frac{M}{3M_{\odot}} \right)^{1/3}.$$  \hspace{1cm} (2)

Mature planets are dense enough so that their radii are much smaller than the Hill-radius. Hence their high densities and as a consequence their high surface gravities usually protect them from tidal disruption or noticeable mass-loss. Stellar companions for comparison reduce their densities due to evolutionary effects when they become giants. Consequently their radii may approach the Hill values depending on their orbital radius. The consequence is Roche-lobe overflow. Planet formation requires a somewhat inverse process, where an extra force compresses the nebula material into the Hill-sphere, allowing more material to flow into the Roche lobe and to increase the planetary mass inside the lobe. All theories of planet formation rely on an extra gravity field to perform this compression.

6 How to Provide the Extra Gravity Field?

Giant planet formation theories may be classified by how they provide the gravity enhancement:

1. the nucleated instability model relies on the extra gravity field of a sufficiently large solid core (condensed material represents a gain of ten orders of magnitude in density and therefore self-gravity compared to the nebula gas),

2. a disk instability may operate on length scales between short scale pressure support and long scale tidal support, or

3. an external perturber could compress an otherwise stable disk on its local dynamical time scales, e.g., by accretion of a clump onto the disk or rendezvous with a stellar companion.

7 From Dust to Planets

Dust growth in the nebula via pairwise collision to cm sizes is now fairly well understood theoretically and experimentally (see Poppe this volume). A key open question is how the transition from $\sim$ 0.1 m dust-agglomerates to km sized planetesimals can be accomplished. Planetesimals, that are the building blocks of gravitationally controlled planetary accumulation, may form by a gravitational instability of a dust subdisk or by continued growth via pairwise collisions provided growth is sufficiently large to dominate over losses due to a radial inward drift (cf. e.g. Wuchterl et al. 2000 for more details). I will follow the planetesimal hypothesis here without further discussion and assume
that protoplanetary nebulae are forming km sized bodies made of condensible elements within a time-frame of about 10,000 a (see e.g. Hueso and Guillot 2003).

The next step of runaway planetesimal accretion proceeds from \( \sim 10^{-9} \, M_{\text{Earth}} \) planetesimals to \( \sim 0.1 \, M_{\text{Earth}} \) embryos. Based on many-body-simulations to simultaneously determine the distribution of orbital elements of planetesimals and the growth to protoplanetary embryos Tanaka and Ida (1999) estimate accretion times for protoplanets of mass \( M_p \) by runaway accretion at orbital radius \( a \):

\[
\tau_{\text{grow}}/[a] = 8 \times 10^5 (M_p/M_{\text{Earth}})^{1/3} (a/\text{AU})^{12/13}.
\]

Runaway accretion stops at the isolation mass. The isolation mass is reached when a planetary embryo has accreted all planetesimals within its gravitational range — the so called feeding-zone —, of a few, typically \( \sim 5 \) Hill-radii around its orbit. The values of the isolation mass depend on the nebula solid surface density and the orbital radius of the embryo. Values are typically, 1 \( M_{\text{Earth}} \) in the outer solar system and a Mars-mass, 0.1 \( M_{\text{Earth}} \), at 1 AU. Protoplanets with masses larger than \( \sim M_{\text{Earth}} \) enter the oligarchic growth stage and have much larger growth times. Kokubo and Ida (2002) estimated the total accretion times of planetary cores through runaway accretion and the late phases of oligarchic growth in the jovian planet region to be:

\[
T_{\text{grow}}/[a] \sim 9 \times 10^4 (e/h_M)^2 (M/10^{26} \, \text{g})^{1/3} (\Sigma/[4 \, \text{g}/\text{cm}^2])^{-1} (a/[5 \, \text{AU}]^{-1/2}.
\]

For an eccentricity in Hill-units, \( e/h_M \), solid surface density, \( \Sigma \), final protoplanetary mass, \( M \) and semi-major axis, \( a \). They estimate that at 5 AU the final mass of a protoplanet would be 5 \( M_{\text{Earth}} \) that would be accreted in 40 Ma. A 9 \( M_{\text{Earth}} \) core at Saturn’s position would require 300 Ma.

8 Solar System Formation Modelling

Understanding the formation of the solar system presently means the reconstruction of a history. That approach is necessary because due to incomplete knowledge of important physical processes it is necessary to include parameterized descriptions of uncertainties. The most famous parameterization is the one of anomalous, turbulent \( \alpha \)-viscosity that is assumed to allow angular momentum redistribution and accretion of mass onto the star. I will not count it here because a structure model of the pre-planetary nebula has to be assumed anyway. Two mayor classes of nebula models may be distinguished: (1) \textit{active} viscous \( \alpha \)-disks (e.g. Ruden and Pollack 1991, Drouart et al. 1999, Hueso and Guillot 2003) and (2) \textit{passive} disks that are heated by absorbed stellar radiation (e.g. the Kyoto minimum mass nebula, Hayashi et al. 1985). Once the class and parameterization of the nebula model is chosen (passive or active) the planet formation processes have to be specified and parameterized. The key nebula processes and parameters are:
1. distribution of temperature and density as a function of orbital radius. That follows for a given class of nebula models from a chosen mass and mass distribution. In practice the discussion is parameterized by the local surface densities $\Sigma_{\text{dust}}$ and $\Sigma_{\text{gas}}$ of nebula condensates and nebula gas, respectively.

2. planetesimal properties and size distribution;

3. planetesimal collision properties, i.e. coefficients of restitution and outcome of collisions — merging into a larger planetesimal or fragmentation into smaller pieces;

4. energy transfer properties of the nebula gas:
   
   (a) dust properties (size distribution, composition, mineralogy) to determine the dust opacities and the efficiency of radiative transfer. Nebula dust differs considerably in size and composition from the interstellar dust due to growth and condensation processes in the nebula;

   (b) a prescription and parametrization for convective energy transfer.

Solar system data are used at two stages: (1) in the construction of the nebula surface densities and (2) in the adjustment of parameters by comparing the final outcome of planets to the empirical data from solar system planets. Because the uncertainties in the initial nebula structure are very large they are the prime parameters that are adjusted. A typical procedure is as follows: In a first step a nebula is constructed e.g. by assigning a volume to every solar system planet. Then hydrogen and helium is added until the presolar abundances are reached. The resulting mass of solids and gas is smeared out across the volume and fitted to the chosen class of nebula models. The result is e.g. a minimum reconstituted nebula with solid and gas surface densities described by parameterized power laws (e.g. Hayashi et al 1985). For the so constructed nebula the outcome of planet formation is deduced in a multi-step process: (1) planetesimal formation, (2) planetesimal accretion, (3) formation of planetary envelopes, (4) nebula gas capture by large envelopes (5) termination of planetary accretion and dissipation of remnant nebula gas. A typical result for the minimum mass nebula is that predicted accretion times turn out to be much larger than plausible nebula life-times. In consequence the original assumptions going into the construction of the nebula are reconsidered. Lissauer 1987, e.g. described how a solid surface density increased by a factor less than ten could account for a Jovian planet within the time constraints. Wuchterl (1993) showed how an increase of the gas surface density by less than a factor of ten would lead to a new class of protoplanets with massive envelopes, that dynamically could grow to a few hundred Earth masses (Wuchterl 1995b). Pollack et al. (1996) adjusted nebula and planetesimal parameters to account for the accretion of Jupiter and Saturn with detailed models of planetesimal accretion and gaseous envelope capture. When coupled to evolutionary models (Guillot 1999, reviewed in Wuchterl et al. 2000) the observed properties
of gravitational fields, the radii and present excess luminosities can be reproduced when interior structures are fitted by detailed planetary structure and evolution models with three compositional layers.

9 Specific Models for Solar System Planets

9.1 Growth of Rocky Terrestrial Planets

A kilometer-sized planetesimal moving around the young Sun has essentially dynamically decoupled from the gas and is approximately following a keplerian orbit. The primary perturbations for this and larger bodies in protoplanetary disks are mutual gravitational interactions and physical collisions (Safronov 1969). These interactions lead to accretion (and in some cases erosion and fragmentation) of planetesimals. Gravitational encounters are able to stir planetesimal random velocities up to the escape speed from the largest common planetesimals in the swarm (Safronov 1969). The most massive planets have the largest gravitationally-enhanced collision cross-sections, and accrete almost everything with which they collide. If the random velocities of most planetesimals remain much smaller than the escape speed from the largest bodies, then these large ‘planetary embryos’ grow extremely rapidly (Safronov 1969). The size distribution of solid bodies becomes quite skewed, with a few large bodies growing much faster than the rest of the swarm in a process known as runaway accretion (Wetherill and Stewart 1989, Kokubo and Ida 1996). Eventually, planetary embryos accrete most of the (slowly moving) solids within their gravitational reach, and the runaway growth phase ends. Planetary embryos can continue to accumulate solids rapidly beyond this limit if they migrate radially relative to planetesimals as a result of interactions with the gaseous component of the disk (Tanaka and Ida 1999). The eccentricities of planetary embryos in the inner Solar System were subsequently pumped up by long-range mutual gravitational perturbations; collisions between these embryos eventually formed the terrestrial planets (Wetherill 1990, Chambers and Wetherill 1998).

9.2 Gaseous Envelopes – Giant Planets

Planetesimals in the solar nebula are small bodies surrounded by gas. A rarified equilibrium atmosphere forms around such objects. Early work in the nucleated instability hypothesis that assumes that such solid ‘cores’ trigger giant planet formation was motivated by the idea that at a certain critical core-mass the atmosphere could not be sustained and isothermal, shock-free accretion (Bondi and Hoyle 1944, Bondi 1952) would set in. Determinations of this critical mass were made for increasingly detailed description of the envelopes: adiabatic (Perri and Cameron 1974), isothermal (Sasaki 1989), isothermal-adiabatic (Harris 1978, Mizuno et al. 1978), radiative and convective energy transfer (Mizuno 1980). By then, modeling the formation and evolution of a proto-giant planet had become essentially a miniature stellar
structure calculation with energy dissipation of impacting planetesimals replacing the nuclear reactions as the energy source. Already, Safronov and Ruskol (1982) pointed out that the rate of gas accretion following instability [at the critical mass] is determined not by the rate of delivery of mass to the planet [as in Bondi-accretion] but by the energy losses from the contracting envelope. Consequently, the energy budget of the envelope has been modeled more carefully taking into account the heat generated by gravitational contraction (quasi-hydrostatic models by Bodenheimer an Pollack 1986).

Most aspects of early envelope growth, up to \( \sim 10 \, M_{\text{Earth}} \), can be understood on the basis of a simplified analytical model given by Stevenson (1982a) for a protoplanet with constant opacity \( \kappa_0 \), core-mass accretion-rate \( \dot{M}_{\text{core}} \), core-density \( \rho_{\text{core}} \), inside the tidal radius \( r_T \). The key properties of Stevenson’s model come from the ‘radiative zero solution’ for spherical protoplanets with static, fully radiative envelopes, i.e., in hydrostatic and thermal equilibrium. Wuchterl et al. (2000) presented a solution relevant to the structure of an envelope in the gravitational potential of a constant mass, for zero external temperature and pressure and using a generalized opacity law of the form \( \kappa = \kappa_0 P^{a} T^{b} \).

The critical mass, defined as the largest mass a core can grow to with the envelope kept static is then given by

\[
M_{\text{crit}}^{\text{core}} = \left[ \frac{3^3}{4^4} \frac{R}{\mu} \frac{4}{4\pi G} \frac{1}{1 + a\pi\sigma} \left( \frac{4\pi}{3} \rho_{\text{core}} \right)^{\frac{3}{4}} \frac{\dot{M}_{\text{core}}}{\ln(r_T/r_{\text{core}})} \right]^{\frac{2}{7}},
\]

and \( M_{\text{crit}}^{\text{core}} / M_{\text{crit}}^{\text{tot}} = 3/4 \); \( R \), \( G \), \( \sigma \) denote the gas constant, the gravitational constant, and the Stefan-Boltzmann constant respectively. The critical mass does neither depend on the midplane density \( \varrho_{\text{neb}} \), nor on the temperature \( T_{\text{Neb}} \) of the nebula in which the core is embedded. The outer radius, \( r_T \), enters only logarithmically weak. The strong dependence of the analytic solution on molecular weight \( \mu \), led Stevenson (1984) to propose ‘superganymedean puffballs’ with atmospheres assumed to be enriched in heavy elements and a resulting low critical mass as a way to form giant planets rapidly (see also Lissauer et al. 1995). Except for the weak dependences discussed above, a proto-giant planet essentially has the same global properties for a given core wherever it is embedded in a nebula. Even the dependence on \( \dot{M}_{\text{core}} \) is relatively weak: Detailed radiative/convective envelope models show that a variation of a factor of 100 in \( \dot{M}_{\text{core}} \) leads only to a 2.6 variation in the critical core mass (Wuchterl 1995b).

However, other static solutions are found for protoplanets with convective outer envelope, which occur for somewhat larger midplane densities than in minimum mass nebulae (Wuchterl 1993, Ikoma et al. 2001). These largely convective proto-giant planets have larger envelopes for a given core and a reduced critical core mass. Their properties can be illustrated by a simplified analytical solution for fully convective, adiabatic envelopes with constant first
adiabatic exponent, $\Gamma_1$:

$$M_{\text{crit}} = \sqrt[3]{\frac{1}{4\pi} \left( \frac{\Gamma_1 - 4}{3} \right)^2 \left( \frac{\Gamma_1 R}{GM} \right)^{\frac{2}{3}} T_{\text{neb}}^{\frac{2}{3}} \rho_{\text{Neb}}^{-\frac{1}{3}}}$$

and $M_{\text{crit}}/M_{\text{crit}}^{\text{tot}} = 2/3$. In this case, the critical mass depends on the nebula gas properties and therefore the location in the nebula, but it is independent of the core accretion-rate. Of course, both the radiative zero and fully convective solutions are approximate because they only roughly estimate envelope gravity and all detailed calculations show radiative and convective regions in proto-giant planets. The critical mass can be as low as $1M_{\text{Earth}}$, and subcritical static envelopes can grow to $48M_{\text{Earth}}$. See Wuchterl et al. (2000) and Wuchterl (1993) for more details. Ikoma et al. (2001) study largely convective protoplanets for a wide range of nebula conditions and show the limiting role of gravitational instability.

The early phases of giant planet formation discussed above are dominated by the growth of the core. The envelopes adjust much faster to the changing size and gravity of the core than the core grows. As a result the envelopes of proto-giant planets remain very close to static and in equilibrium below the critical mass (Mizuno 1980, Wuchterl 1993). This has to change when the envelopes become more massive and cannot re-equilibrate as fast as the cores grow. The nucleated instability was assumed to set in at the critical mass, originally as a hydrodynamic instability analogously to the Jeans instability. With the recognition that energy losses from the proto-giant planet envelopes control the further accretion of gas, it followed that quasi-hydrostatic contraction of the envelopes would play a key role.

9.3 Hydrodynamic Accretion Beyond the Critical Mass

Static and quasi-hydrostatic models rely on the assumption that gas accretion from the nebula onto the core is very subsonic, and the inertia of the gas and dynamical effects as dissipation of kinetic energy do not play a role. To check whether hydrostatic equilibrium is achieved and whether it holds, especially beyond the critical mass, hydrodynamical investigations are necessary. Two types of hydrodynamical investigations of protoplanetary structure have been undertaken in the last decade: (1) linear adiabatic dynamical stability analysis of envelopes evolving quasi-hydrostatically (Tajima and Nakagawa 1997) and (2) nonlinear, convective radiation hydrodynamical calculations of core-envelope proto-giant planets (see Wuchterl et al. 2000). In the linear studies it was found that the hydrostatic equilibrium is stable in the case they investigated. The nonlinear dynamical studies follow the evolution of a proto-giant planet without a priori assuming hydrostatic equilibrium and they determine whether envelopes are hydrostatic, pulsate or collapse, and at which rates mass flows onto the planet assuming the mass is available in the planet’s feeding zone. Hydrodynamical calculations that determine the flow from the nebula into the protoplanet’s feeding zone are discussed in Sect. 17.
Figure 1: Hydrostatic envelope accretion due to a core growing by accretion of planetesimals (left) and hydrodynamical ejection of protoplanetary envelope gas due to pulsation-driven mass-loss (right). The evolution is shown for Mizuno’s (1980) Neptune conditions and a planetesimal accretion-rate of the core \( \dot{M}_{\text{core}} = 10^{-6} M_{\text{Earth}} \, \text{a}^{-1} \) (full line). The dashed line is the same but with time-dependent MLT-convection, updated molecular opacities, and a particle-in-a-box core-accretion-rate (Wuchterl 1997) with a planetesimal surface density of 10 kg m\(^{-2}\) and a gravitational focusing factor of 2000.
The first hydrodynamical calculation of the nucleated instability (Wuchterl 1989, Wuchterl 1991a,b) started at the static critical mass and brought a surprise: Instead of collapsing, the proto-giant planet envelope started to pulsate after a very short contraction phase (see Wuchterl 1990 for a simple discussion of the driving \( \kappa \)-mechanism). The pulsations of the inner protoplanetary envelope expanded the outer envelope and the outward traveling waves caused by the pulsations resulted in a mass-loss from the envelope into the nebula. The process can be described as a pulsation-driven wind. After a large fraction of the envelope mass has been pushed back into the nebula, the dynamical activity fades and a new quasi-equilibrium state is found that resembles Uranus and Neptune in core and envelope mass (cf. Fig. 1, full line).

The main question concerning the hydrodynamics was then to ask for conditions that allow gas accretion, i.e., damp envelope pulsations. Wuchterl (1993) derived conditions for the breakdown of the radiative zero solution by determining nebulae conditions that would make the outer envelope of a ‘radiative’ critical mass proto-giant planet convectively unstable. The resulting criterion gives a minimum nebula density that is necessary for a convective outer envelope:

\[
\frac{\rho_{\text{crit}}^{\text{neb}}}{[10^{-10} \text{g cm}^{-3}]} = 2.2 \left( \frac{T}{[100 \text{K}]} \right)^3 \left( \frac{\nabla_s}{[2/7]} \right) \left( \frac{\mu}{[2.2]} \right) \left( \frac{\kappa}{[\text{cm}^2 \text{g}^{-1}]} \right)^{-1} \left( \frac{M_{\text{core}}}{[10^{-6} \text{M}_{\text{Earth}} \text{a}^{-1}]} \right)^{-1/2} \left( \frac{M_{\text{core}}}{[10 \text{M}_{\text{Earth}}]} \right)^{1/4} \left( \frac{\rho_{\text{core}}}{[5.5]} \right)^{-1/4}
\]

Protoplanets that grow under nebula conditions above that density have larger envelopes for a given core and a reduced critical mass. For sufficiently large nebula densities Wuchterl (1995b) found that the pulsations were damped and rapid accretion of gas set in and proceeded to 300 \( \text{M}_{\text{Earth}} \). The critical core masses required for the formation of this class of proto-giant planets are significantly smaller than for the Uranus/Neptune-type (see Wuchterl 1993, 1995b, Ikoma et al. 2001).

### 10 Importance of Convection

Convection plays an important role in determining the mass of protoplanets by controlling energy transfer in the outer layers under specific nebula conditions and the dynamical behaviour of their envelopes beyond the critical mass as described in the last section. Most giant planet formation studies use zero entropy gradient convection, i.e., set the temperature gradient to the adiabatic value in convectively unstable layers of the envelope or use the time-independent mixing length theory. That is done for simplicity but can be inaccurate, especially when the evolution is rapid and hydrodynamical waves are present (see Wuchterl 1991b). Furthermore convection in the outer layers of a protoplanet occurs at weak gravities and relatively low optical depths, hence departures form an adiabatic behavior might be expected. It
was, therefore, important to develop a time-dependent theory of convection that can be solved together with the equations of radiation hydrodynamics in the entire protoplanetary flow-regime. Such a time-dependent convection model (Kuhfuß 1987) has been reformulated for self-adaptive grid radiation hydrodynamics (Wuchterl 1995a) and applied to giant planet formation (Götz 1993, Wuchterl 1996, 1997). In a reformulation by Wuchterl and Feuchtinger (1998), it closely approximates standard mixing length theory in a static local limit and accurately describes the solar convection zone and RR-Lyrae lightcurves.

The heart of this convection model is a dynamical equation for the specific kinetic energy density, $\omega$, of convective elements. The equation accounts for creation of eddies by buoyancy, dissipation of eddies due to viscous effects as well as eddy advection and radiative losses:

$$\frac{d}{dt} \left[ \int_{V(t)} \varrho \omega \, d\tau \right] + \int_{\partial V} \varrho \omega u_{rel} \cdot dS = \int_{V(t)} \left( S_\omega - \tilde{S}_\omega - D_{\text{rad}} \right) \, d\tau$$  

where the eddy kinetic energy generation-rate, the eddy dissipation-rate, the convective enthalpy flux, the reciprocal value of the mixing-length, $\Lambda$ and the time-scale for radiative eddy losses, respectively, are:

$$S_\omega = -\nabla_s T \frac{\partial P}{\partial r} \Pi$$  

$$\tilde{S}_\omega = \frac{c_D}{\Lambda} u^{3/2}$$  

$$j_w = \varrho T \Pi, \quad \Pi = \frac{w}{T} u_c F_L \left[ -\sqrt{\frac{3}{2}} \alpha_s \frac{\lambda T}{w} \frac{\partial s}{\partial r} \right]$$  

$$\frac{1}{\Lambda} = \frac{1}{\alpha_{\text{ML}} H_p^{\text{stat}}} + \frac{1}{\beta_{r,T}}, \quad H_p^{\text{stat}} = \frac{p}{\varrho G M_c}$$  

$$\tau_{\text{rad}} = \frac{c_p \kappa \rho^2 \Lambda^2}{4\sigma T^4 \gamma R_c^2}, \quad D_{\text{rad}} = \frac{\omega}{\tau_{\text{rad}}}$$

In the time-independent and static limit this is essentially mixing-length theory and the accuracy is assured by fitting the prescription to the Sun via a solar model. The difference is that the parametrization is now brought into a fluid-dynamical framework and basic physical plausibility constraints that are required in the time-dependent regime are fulfilled (cf. Wuchterl and Feuchtinger 1998). The Schwarzschild-Ledoux criterion is contained in the formulation via $-\partial s/\partial r = c_p/H_p(\nabla - \nabla_s)$ and $\nabla_s = \nabla_{\text{ad}}$ in the absence of energy sources and sinks inside eddies. Convectively unstable stratifications occur in this model when pressure and temperature gradients have the same sign and produce a positive value of $S_\omega$ that then contributes a source of turbulent kinetic energy, $\omega = 3/2 u_c^2$ to the balance equation of turbulent kinetic energy. $u_c$ being the convective velocity corresponding to mixing length theory. A general problem of mixing length theory — the violation of a convective flux-limit — has been corrected as described by Wuchterl and Feuchtinger by
introducing a flux-limiting function (cf. Wuchterl and Tscharnuter 2003). The
great advantage of that approach is that a general prescription can be used
for the Sun, stellar evolution, pulsating stars, brown dwarfs, planets and pro-
toplanets. Any calibration of parts of the convection model obtained in one
astrophysical system — the mixing length parameter calibrated by the Sun,
the time-dependent behavior tested by RR-Lyrae stars — will decrease the
uncertainties in applications to not-so-easy-to-observe systems as protoplan-
etas.

11 Fluid-Dynamics of Protoplanets

The time-dependent convection model allows the formulation of a fully time-
dependent set of equations that describe the radiative and convective en-
velopes of protoplanets as well es the protostellar collapse and pre-main se-

\[ \Delta M_r = \int_{V(t)} \varrho \, dr, \]

\[ \frac{d}{dt} \left[ \int_{V(t)} \varrho \, dr \right] + \int_{\partial V} \varrho (u_{\text{rel}} \cdot dS) = 0, \]

\[ \frac{d}{dt} \left[ \int_{V(t)} \varrho u \, dr \right] + \int_{\partial V} \varrho u (u_{\text{rel}} \cdot dS) + \int_{V(t)} \left( \frac{\partial \varrho}{\partial r} + \varrho \frac{GM_r}{r^2} \right) \, dr = C_M, \]

\[ \frac{d}{dt} \left[ \int_{V(t)} (\varrho(e + \omega)) \, dr \right] + \int_{\partial V} [\varrho(e + \omega) u_{\text{rel}} + j_{\omega}] \cdot dS + \int_{V(t)} p \text{div} u \, dr = -C_E, \]

\[ \frac{d}{dt} \left[ \int_{V(t)} E \, dr \right] + \int_{\partial V} [E u_{\text{rel}} + F] \cdot dS + \int_{V(t)} P \text{div} u \, dr = C_E, \]

\[ \frac{d}{dt} \left[ \int_{V(t)} \frac{F}{c^2} \, dr \right] + \int_{\partial V} \frac{F}{c^2} (u_{\text{rel}} \cdot dS) + \int_{V(t)} \left( \frac{\partial P}{\partial r} + \frac{F}{c^2} \frac{\partial u}{\partial r} \right) \, dr = -C_M, \]

\[ \frac{d}{dt} \left[ \int_{V(t)} \varrho \omega \, dr \right] + \int_{\partial V} \varrho \omega u_{\text{rel}} \cdot dS = \int_{V(t)} \left( S_\omega - S_{\omega} - D_{\text{rad}} \right) \, dr, \]

\[ C_M = \int_{V} \kappa \varrho \frac{F}{c} \, d\tau, \quad C_E = \int_{V} \kappa \varrho (4\pi S - cE) \, d\tau, \quad P = \frac{1}{3} E. \]

The equations are applied to the volume taken by the protoplanetary enve-
lope, and assuming spherical symmetry. They determine the motion of gas in
the protoplanetary envelope or determine the hydrostatic equilibrium if forces
balance out. As a consequence of the structure and motion in the envelope
the mass exchange with the nebula results and determines whether the planet
gains or loses mass. Material belongs to the planet when it in inside the
planet’s gravitational sphere of influence. The sphere of influence is approxi-
mated by a spherical volume of radius, \( R_{\text{Hill}} \) around the condensible element
planetary embryo at the center. At the outer boundary of the volume, i.e.
at the Hill-sphere the protoplanet radiates into the ambient nebula and may
exchange mass with it. Planetesimals enter the sphere of influence and collide with the core. The core’s surface is the inner boundary. The surface changes its radius as the core grows due to planetesimal accretion. The planetesimals add mass to the core and dissipate their kinetic energy at the core surface. The core radiates the planetesimal’s energy into the adjacent planetary envelope gas. That heats the inner parts of the protoplanetary envelope. The resulting temperature increase relative to the nebula induces temperature gradients that drive energy transport through the envelope towards the nebula. In general transfer occurs by both, radiation and convection.

12 Dynamic Diversity

Quasi-hydrostatic models of giant planet formation always find envelope-growth once the critical mass has been reached. When departures from hydrostatic equilibrium are allowed and the dynamics of the envelopes are calculated the situation is more diverse: the occurrence of accretion depends on the nebula properties and the properties of the protoplanet at the critical mass (see Wuchterl et al. 2000). Furthermore the onset of planetary envelope mass-loss depends on the planetesimal accretion-rate of the core and the treatment of energy transfer. The dependence is quantitatively significant on the scale of a few Earth-masses that is comparable to the masses of terrestrial planets, the cores of giant planets and the envelopes of planets like Uranus and Neptune. In Fig. 1 two calculations are compared for the Neptune position in the Kyoto-nebula\(^1\). Hence nebula-properties and orbital-radius-effects (orbital dynamic time-scale, solar tides, size of the Hill-sphere) are identical for both calculations. The difference is in the energy input and content of the envelopes, i.e. the thermal structure. The first calculation (full line in Fig. 1) is for a constant mass accretion-rate and simple instantaneous zero entropy gradient convection. The second calculation (dashed line) is physically more refined, with a particle-in-box planetesimal accretion-rate and time-dependent convection as described above. The gravitational focusing factor is chosen appropriate for the runaway phase up to the isolation mass. The outcome is qualitatively very similar to the one for simpler physics: with growing core mass the envelope mass increases until the slope becomes almost vertical in the vicinity of the critical mass (Fig. 1, left panel). But the values of the critical core mass and the envelope mass at given core mass are significantly different (critical core masses, 13 and 4 M\(_{\text{Earth}}\), with envelopes of 5 and 0.5 M\(_{\text{Earth}}\), respectively). The evolution beyond the critical core mass is shown as a function of time in the right panel. Note that the evolution is now on the short dynamical time-scale (a few years) of the envelopes. Mass-loss is driven in both cases, and both calculations approach a new quasi-equilibrium state.

\(^1\)Mizuno’s Neptune (17.2 AU, 45 K, \(3.0 \times 10^{-13}\) g/cm\(^3\)), for orbital, radius, nebula temperature and midplane nebula density, resp., is located inside of Neptune’s present orbital radius (semi-major axis 30.06 AU) to allow for outward migration after formation (cf. Hayashi et al. 1985).
with smaller envelope mass. But the envelope masses ultimately differ by approximately a factor of ten. Even with the relatively well known properties of giant planets at the critical mass, no general conclusion is possible about the dynamical processes that happen thereafter and the expected envelope mass of e.g. a Uranus-type planet. It is obvious that a more general understanding is needed to predict the outcome of planet formation when realistic physics, as runaway planetesimal accretion, dynamical effects and plausible convection are included.

Following the usual approach for solar system planet formation we might try adjust the parameters of planetesimal accretion to account for the observed properties of Neptune, say, but that will not lead to a predictive theory or a general understanding of planet formation. I will outline an alternative approach below.

13 A Few Problems of Solar System Theory

To conclude the discussion of solar system planet formation theory I will describe open problems that were known before the discovery of the first extrasolar planet. These problems might help to understand what parts of the theory might need modifications for the general application to planet formation in the galactic neighborhood. With dust growth to cm size now increasingly well understood by theoretical and experimental work (Blum and Wurm 2000, see also Poppe, this volume) the most important remaining problems are:

1. planetesimal formation,
2. the total growth times in the outermost solar system, and
3. the final planetary eccentricities.

13.1 Planetesimal Formation

Planetesimal formation by coagulation and agglomeration of dust grains may stall at dm to m size where loss processes by radial drift may halt the planet formation process. Planetesimal formation by a gravitational instability of a dust-subdisk may require special nebula conditions that are incompletely explored to decide under how wide a range the instability will operate and whether the non-linear outcome are the consolidated condensible element bodies that are envisaged and assumed in the planetesimal hypothesis. The related key question is how wide a diversity of nebulae will lead to instabilities that produce appropriate planetesimals. Being appropriate mostly means a size large enough to decouple from the head-wind of the nebula gas. An event that typically occurs at km size. Production of non-standard planetesimals does not automatically mean that planet formation will not proceed as presently imagined but new pathways in a theoretically essentially unexplored regime have to be worked out in that case.
13.2 Late Accretion: Total Planetary Growth-Times

The standard model is centered around the planetesimal hypothesis that has been successful to understand a wide range of solar system bodies, to a large extent in a quantitative way. But observational results obtained for nearby star-forming and young star regions quantitatively challenge the standard model because indicators of the presence of circumstellar disks (Haisch et al. 2001) suggest disk depletion time-scales that are comparable or shorter than calculated formation times for solar system giant planets of at least $10^8$ years (Safronov 1969). Moreover, unless the eccentricities of the growing embryos are damped substantially, embryos will eject one another from the star’s orbit (Levison et al. 1998). Runaway growth, possibly aided by migration (Tanaka and Ida 1999), appears to be the way by which solid planets can become sufficiently massive to accumulate substantial amounts of gas while the gaseous component of the protoplanetary disk is still present (Lissauer 1987, Kokubo and Ida 2002).

The theoretical estimates for planetary growth times have been known to be idealized because the size distribution of planets, embryos and planetesimals and the interaction with the residual nebula gas can only be incompletely accounted for in then n-body calculations that are necessary to reliably calculate the final orbital outcome at least for an idealized situation to allow a quantitative discussion and theoretical progress. Recently Inaba and Ikoma (2003) and Inaba et al. (2003) have looked at the collisional cross sections of planetesimals with gaseous envelopes and found a significant increase for their accretion-rates reducing the planet-growth times considerably. This is especially important for the giant planet regime where the envelopes may become comparable in mass to the condensible element cores during the runaway phase. Hence total growth times can be expected to decrease further when the nebula gas is not neglected in determining the collision cross sections of planetary embryos.

13.3 Late Accretion and Final Eccentricities

Late accretion and hence the evolution to the final orbital parameters of a planet is governed by interactions with other planets, the remaining planetary embryos and planetesimals (Levison et al. 1998, Thommes and Lissauer 2003, Levison and Agnor 2003). The relevant overall masses in all components may or may not be comparable to the mass of the largest planet. There is probably still a large and locally dominant number of bodies around, that in case of the solar system are responsible, e.g. for the formation of the moon and the late heavy bombardment. Late accretion effects are apparently important in the asteroidal region of the solar system where it is possible that Jupiter’s perturbations precluded the accretion of embryos into a planet. The important remaining dynamical process is then the orbital evolution of planets and embryos due to secular mutual perturbations. In models of final planetary growth they typically lead to eccentricities larger than observed in the
solar system (Wetherill 1990, Chambers and Wetherill 1998) and found in long-range backward integrations of the planetary system (Lecar et al. 2000). Eccentricity excitation is possibly damped by a remaining small-size planetesimal population via dynamical friction. Studies with an increased number of planetary embryos reduce the discrepancy (Chambers 2001) but still do not reproduce the low time-averages of the planetary eccentricities in the solar system. Eccentricity damping by residual nebula gas or by a remnant population of planetesimals or small planetary embryos might resolve that problem.

14 The Era of Discovery

In the last decade of the 20th century the precision to detect extrasolar planets was achieved by measurements of radial velocities and arrival times of pulsar signals. In 1992 three planetary mass bodies were detected around the millisecond pulsar PSR 1257+12 (Wolszczan and Frail 1992). The planetary interpretation of the measurement was confirmed when predicted changes in the orbital elements due to a mean motion resonance of the two larger, few, $M_{\text{Earth}}$ planets were detected (Wolszczan 1994). A 12 year long survey of the radial velocity of 27 solar-like stars at the University of British Columbia (Walker et al. 1995) yielded the stunning result that less than 1 of 10 stars had a detectable planet with properties of Jupiter. That immediately raised the question whether the solar system was typical and planet formation ubiquitous, at least for stars similar to the Sun. Two of the surveyed stars do now have a detected planet: $\epsilon$ Eri (Hatzes et al. 2000) and $\gamma$ Cep (Hatzes et al. 2003), — the 12 years of the Walker study were just not sufficient for the signal retrieval. The unexpected discoveries reached a first climax when a giant planet with a period of about 4 days was discovered to orbit the star 51 Pegasi, in October 1995 (Mayor and Queloz 1995). Within months planets around 70 Vir and 47 UMa were announced (Marcy and Butler 1996a,b) — the former in a very eccentric orbit, the latter finally at least closer to orbital properties that are familiar from the solar system. More than a hundred extrasolar planet candidates have been detected so far (Jan 2004) (e.g. Mayor, Udry and Santos 2003). More than a dozen of them are already known to be members of a multiple planet system. Some candidates of extrasolar planets orbit binary stars and one is located it the M4 globular cluster in a system together with a white dwarf and a pulsar.

The planetary properties of this first exoplanet harvest were unpredicted by theory and surprising because of the detection of:

1. giant planets with orbital periods of a few days, corresponding to 0.01 of Jupiter’s orbit-radius,
2. planet candidates with $M \sin i$ up to $13M_{\text{Jupiter}}$ \(^2\),

\(^2\) $M_{\text{Jupiter}} = 317.71 M_{\text{Earth}} = 1.898 \times 10^{27} \text{kg} = 0.95 \times 10^{-3} M_\odot$
3. a broad range of eccentricities larger than known for the solar system planets\(^3\),

4. planets in binaries\(^4\).

15 Theory-PlanetMania

The 1995 discovery of a planetary companion to 51 Peg electrified theorists. Very rapidly Guillot et al. (1996) showed that planets like 51 Peg b could indeed survive for the estimated ages of their host stars. Within a year it was shown that 51 Peg b could form at its present location when existing fluid-dynamical models of giant planet formation were applied to orbital distances of 0.05 AU, provided sufficient building material was within the planet’s feeding zone (Wuchterl 1996).

But very rapidly alternative theories emerged. They kept the often communicated view that giant planets would only be able to form beyond the ice-line, typically beyond a few AU from their parent star. If that remained true the planets had to move from their formation place to a position much closer to the star, like 51 Peg b’s. How could a massive planet like Jupiter move from 5 to 0.05 AU, say? Mechanisms to change the orbital elements were proposed:

1. violent dynamical relaxation of multi giant planet systems the so called jumping Jupiters, and

2. a gradual decrease of the planetary orbital radius due to interaction with the disk of gas and planetesimals: orbital migration.

Violent dynamical relaxation (Weidenschilling and Mazari 1996) needs synchronizing of planet formation to provide a number of giant planets within a narrow time-span. They would subsequently very rapidly interact via mutual perturbations that typically destroy the system on a dynamical time-scale leaving a close-in giant planet in some of the cases. While it is unlikely that the assumed very unstable initial state would be reached as the final state of the preceding planet formation process, there are additional problems. The close orbits produced typically would not be as small as observed, hence requiring further orbital evolution and the final systems would be rather disturbed with one close in and one planet far out. That is unlike a system like \(\upsilon\) Andromedae with relatively close orbiting giant planets in addition to the Pegasi-planet at 0.05 AU.

The other alternative, involving migration caused by disk-planet-interactions, is favored by many researchers (e.g., Lin et al. 1996; Trilling et al. 1998). It

\(^3\)that might have been expected because of the difficulties to explain the low eccentricities but have not, most likely because missing elements in late accretion were obvious (see above)

\(^4\)in spite of the fact that dynamicists had shown planetary orbits to be stable in binary systems as well as already classified them as F- and S-type (Dvorak 1984, Dvorak et al. 1989) in analogy to planets and satellites in the solar system
starts out with a standard situation of planet formation: a planetary embryo or proto-giant planet orbiting at a conventional giant planet orbital distance. The change of orbital elements is continuous to very small values and there is now requirement of other planets to be present at the same time. I will discuss migration as the dominating theory of Pegasi-planet formation below.

Even more radical rethinking of planet formation has been proposed. It was reconsidered that giant planets might form directly via a disk instability (see Wuchterl et al. 2000 for a review). While this was more directed towards the time-scale problem of planet formation in the outer solar system it also might offer a way to explain the diversity in the detected extrasolar planets. Maybe some of the systems, in particular the very massive planetary candidates, with minimum masses $M \sin i \sim 10 M_{\text{Jupiter}}$, were formed by a disk instability and others by the nucleated instability?

With the formation process reconsidered, the relatively large minimum masses of many of the early exoplanet discoveries, the large eccentricities, that are hardly distinguishable from those of binary stars, and the Pegasi-planets the old question was re-posed: What is a planet?

15.1 What is a Planet?

Given the unexpected properties of extrasolar planet candidates and claims of discoveries of so called free floating planets the IAU's working group on extrasolar planets issued a preliminary working definition based on the following principles:

- objects with true masses below the limiting mass for thermonuclear fusion of Deuterium that orbit stars or stellar remnants are planets,
- substellar objects above that mass are brown dwarfs,
- free floating objects in young star clusters with masses below the Deuterium limit are no planets.

With these guidelines for a definition, most of the discovered extrasolar planets are candidates because their true masses are not yet sufficiently well known. Due to the unknown orbital inclination only the minimum masses $M \sin i$ are known. These values also probably approximate the true masses well, but about 1% may turn out not to be below the 13 $M_{\text{Jupiter}}$ mass limit. Only the masses of the pulsar planets (masses from mutual perturbations) and the transiting Pegasi-planet, HD 209 458 b (orbital inclination determined from transit light curve) are true masses in the sense of the above definition.

The IAU definition also repeatedly and explicitly excludes the way of formation from the definition. But by referring to the Deuterium-limit the formation history implicitly enters the definition through a back-door.

The minimum mass for thermonuclear fusion of Deuterium is a concept that is shaped in some analogy to the minimum mass for hydrogen burning that defines the lower end of the stellar main sequence. But the main sequence is defined by the stellar thermal equilibrium, where nuclear burning
fully balances the surface energy-losses. The stellar luminosity is balanced by nuclear energy production of the same magnitude. Objects with masses below the lower end of the main sequence also burn hydrogen, but insufficiently to fully balance the luminosity (Kumar 1963) — they are called brown dwarfs. Because of that fact stellar thermal equilibrium and a phase of constant radius is never reached. The objects have to contract forever to (at least partly) supply their luminosity need from contracting into their own gravitational field. Because they never reach an equilibrium state (as the main sequence) their evolution always depends on their history. Ultimately that means it depends on their formation. Recent calculations of spherical collapse of stars and brown dwarfs show that the young objects after the end of significant mass accretion (on the pre-main sequence for young stars) contain the thermal profile shaped by the collapse (Wuchterl and Tschamnter 2003). But it is exactly that thermal profile that controls the ignition of thermonuclear fusion processes. That is particularly true for Deuterium that is in all models (hydrostatic and dynamic) a very episodic event in the first few million years for stars and brown dwarfs.

That means that for masses below the main sequence, the question of whether Deuterium (or hydrogen) will just start to burn — and that defines the borderline in the IAU-definition — will depend on the history of the respective low mass object. For Deuterium, that burns early on, that means it will depend on the formation process.

In summary, hydrogen burning to define the lower end of the main sequence means using a major, physically dominating process that defines a long-lasting equilibrium state that contains no memory about the formation history. To the contrary using deuterium burning for characterizing a planet means using an essentially irrelevant process for the evolution of low mass objects, that is history-dependent in a way that will be very hard to predict.

I think we should rethink the definition of a planet along the following mayor characteristics:

- heavy element enrichment,
- orbital stability properties,
- mass,

The first item is straightforward with the large enrichments (bulk and atmospheric) of planets relative to their parent star. A factor of 3 and more should also be a working basis, that is empirically very likely much less challenging in terms of future determinations in exoplanets than trying to observe the presence of Deuterium in a few Jupiter-mass companion in a 10 AU orbit, even around a nearby star.

The second point is still hard to characterize quantitatively but large progress has been made in the understanding of the stability-properties of the solar system (Lecar et al. 2000). The basis could be volume exclusion principles based on the non-overlap of planet-domains with a width of multiples of the Hill-radii. Laskar has recently shown that they are the consequence
of simple assumptions about planetary growth via pairwise embryos collisions. The low solar system planetary eccentricities could be a special case of that. Certainly low eccentricity planets can orbit closer together in terms of Hill-exclusion.

The third point is the most uncertain. Observationally the characteristic mass of the detected planetary population seems to decrease as more discoveries are made. Presently it may be around 3 Jupiter-masses, with the estimated true distribution still peaked towards the detection limit. Theoretically the planetary masses are presently essentially unconstrained at the upper end. The best hypothesis for the moment is that planetary masses are limited by the amount of material that is within the respective feeding zone in the nebula. This definition contains a considerable degree of circularity in general but has been consistently worked out at least for planetary embryos.

In summary, I think we will see the definition of a planet remaining a morphological type, i.e. without an explicit, constructive definition for some time. But I think that condensible elements should play the major role, not a hydrogen trace-isotope.

16 Why Not Abandon Solar System Formation Theory?

If the extrasolar planet properties are so different, and the theories developed for the solar system did not predict their properties, why not look for a completely new theory of extrasolar planet formation? Should this new theory be more along the lines of binary star formation? As one might conclude given the fact that the period-eccentricity distribution of extrasolar planets is indistinguishable from the one of binary stars.

I think we should not throw away the solar system formation theory as a general theory of planet formation too early. It does not only provide a fairly consistent picture of solar system bodies ranging in size from interplanetary dust particles to Jupiter, but it has also led to the development of predictive elements that led e.g. to the correct prediction of many orbital properties of trans-neptunian objects.

On the other hand scenarios like the jumping Jupiters to explain Pegasi planets raise more questions than they answer. Instead of one planet at 0.05 AU the simultaneous formation of many massive planets is required as a presumption.

Moving a planet in place by migration, requires in addition to the planet a mechanism that counteracts the migration process to park the planet once it has a arrived at the intended final orbit. That is by no means trivial because of the large migration-rates (on the local disk evolution time-scale and shorter) that increase as the star is approached resulting in acceleration rather than slow-down for small orbital radii. Numerous parking-processes have been suggested, but ultimately the way out of the dilemma might be only the dissipation of the nebula. Planets would then continue to form, drift
inward, and disappear into the star until the exhaustion of nebular material finally ends this road of destruction. Observationally no metal trend versus effective temperature is found on the main-sequence, reflecting the different sizes of convection zones that would play the role as a planetary graveyard and hence might be expected to be heavy-element enhanced (Santos et al. 2003).

The disk instability model, if it works and does indeed form planets, as a general alternative would have to be augmented by a separate way of terrestrial planet formation. For the giant planets the disk instability would probably require a separate heavy element enrichment process. Even if Jupiter formed by a disk instability, the craters on a Galilean moon would recall the planetesimal picture.

The basic feature of the non-standard planet formation theories is that they quickly provided scenarios for newly discovered objects. But typically they would fail the solar system test. Let us look at planetary migration as an example of the new pathways of planet formation, and discuss it in more detail.

17 Planet-Disk Interaction

17.1 What is Planetary Migration?

Planetary migration presently seems to denote any systematic change of the orbital semi-major axis of a planet that does not change direction. Historically outward migration of Uranus and Neptune as a consequence of ‘passing comets down to Jupiter’ seems to be the first large scale post-formation reshaping process of planetary orbits that was considered. It had been noticed by Fernandez and Ip (1984) and has been considered by Hayashi et al. (1985) as a process that would allow shorter growth times for Uranus and Neptune. In the late stages of outer solar system formation these planets would move outward as a consequence of angular momentum exchange when perturbing comets into a Jupiter controlled orbit with subsequent ejection to the Oort cloud.

After the discovery of 51 Peg b it has become a custom in planet formation theory to denote many kinds of changes in the planetary semi-major axis or orbital distance as planetary migration. That is usually independent of the physical process underlying the respective orbital change. With processes proposed and a terminology of types I and II, suggested by Ward (1997), and a type III added later by analogy we have in particular:

**type I migration:** an embedded planetesimal or planetary embryo that interacts with its own disk-density-waves;

**type II migration:** a protoplanet that has opened a gap in the nebula — i.e. produced a region of reduced nebula density in its feeding-zone — is locked in that gap and follows the gradual inward motion of viscous disk gas together with the gap;
type III migration: an instability of the planet-disk-interaction that leads to orbital decay within a few orbital periods.

We distinguish here between migration processes that modify the orbit by less than a factor of $e^2$ (or $\sim 10$) and those that may lead to larger changes up to orders of magnitude in the orbital radius, and may ultimately result in the loss of the planet. The latter processes we will call violent migration in the following. They may dominate the planet formation processes if they operate in many and diverse nebulae.

After planetesimal formation violent migration is the second key problem of planet formation. Like an inefficient planetesimal formation mechanism it has the potential to make the formation of systems similar to the solar one, very unlikely. It is expected by many investigators to become important in the mass range resulting from the early fast runaway-mode of planetesimal growth. The runaway-phase ends when all planetesimals that are within the gravitational range of the locally largest body have been accreted and hence its feeding zone has been emptied. Planetary embryos gravitationally interact with the ambient gas disk, planetesimal disk and other planetary embryos or planets. As a result planetary migration can come about (see Thommes and Lissauer (2004), for a review).

The migration effects become severe at larger sizes because they are proportional to the planetesimal mass, for type I (after Thommes and Lissauer 2004):

$$v_I = k_1 \frac{M}{M_*} r \frac{\Sigma_d \Omega^2}{M_*} \left(\frac{r \Omega}{c_T}\right)^3$$

where $k_1$ is a measure of the torque asymmetry, $M$, $M_*$ the masses of the planet and the primary resp., $r$ is the orbital radius, $\Omega$ is the disk angular velocity, that is approximately keplerian with $\Omega_{\text{Kepler}} = \sqrt{GM_*/r^3}$, $\Sigma_d$ the disk surface density and $c_T$ the isothermal sound speed.

### 17.2 Violent Migration

Violent migration is a back-reaction of the planetary embryos ‘bow wave’ in the nebula onto the embryo itself. As the embryo orbits the star, its gravitational potential adds a bump to the stellar one. At the embryo’s orbit — at the corotation resonance, in the linear terminology of migration theory — the embryo and its potential move almost at the same, keplerian, velocity. That is co- orbital motion as in the case of Jupiter and the Trojan asteroids. Inside the embryo’s orbit, the gas in a quasi-keplerian disk orbits faster and hence the embryos potential and gravitational acceleration travels at a different speed relative to the gas. This accelerates the disk gas and excites a pressure- and density-wave that travels with the embryo. Because matter deeper in the primaries potential must orbit faster the waves are dragged forward inside, and backwards, relative to the embryo, outside the embryos orbit. These rather particular protoplanetary bow-waves include density enhancements that gravitationally back-react onto the planet. Due to
the inherently asymmetric nature of the situation (keplerian orbital velocities changing $\propto r^{-1/2}$) and the particular wave-pattern the forces (and in particular the torques) on the planet may not cancel out. They do not cancel out for simple disk models like plausible radial power-laws for the nebula surface density. That leads to a net exchange of angular momentum between the planetary embryo and the gaseous disk if the waves dissipate or break in the disk, neighboring the planet’s gravitational sphere of influence. The result is the familiar reaction of orbiting matter if angular momentum exchange is allowed: most matter (the embryo) moves in and a small amount (some gas) moves out carrying away the angular momentum. The very growing of the embryo would lead to orbital decay and gradual movement towards the star on time scales of disk evolution or much smaller. Many studies are presently devoted to determine the strength of the effect and evaluate the rates of orbital decay and hence the possible survival times for planets of given mass in a given disk. If migration dominates planet formation, it has the potential to wipe out any and many generations of planets. In that case and because the basic effect originates from a relatively small difference in a delicate torque balance, in a significantly perturbed non-keplerian disk, I doubt that we will be able to reliably predict much about planet formation any time soon.

17.3 A Closer Look

Modern planetary migration theory originated from the study of planetary rings (cf. Thommes and Lissauer 2004). While the basic physical processes, density waves in quasi-keplerian disks are well studied, the application to the problem of forming planets in the nebula disk is not straightforward.

The basic problem that has to be solved to determine migration-rates for a proto giant planet orbiting in a nebula disk is the fluid dynamical analogue to the restricted three body problem of celestial mechanics. In the classical problem of celestial mechanics the motion of a test particle is considered in the combined gravitational field of the Sun and a planet. For a proto giant planet two modifications have to be made:

1. a protoplanet, unlike a mature planet is not well approximated by a point mass,

2. the test particles are replaced by a fluid with a finite pressure.

A protoplanet fills its Hill-sphere and a considerable fraction of its mass is located at significant fractions of the Hill-sphere (e.g. Mizuno 1980, Pečnik and Wuchterl 2004). Furthermore the protoplanet builds up a significant contribution to the gas pressure at the Hill-sphere. Typically planet and nebula are in a mechanical equilibrium. This may only change when and if the planet collapses into the Hill-sphere and does not rebounce. Fluid dynamical calculations show that this is a non-trivial question that depends on the structure of the outer protoplanetary layers, near the Hill-sphere (Wuchterl 1995b). In consequence, the problem of a protoplanet in a nebula disk is
not only a problem of gas-motion in the gravitational potential of two centers, but it is controlled by the nebula gas-flow and the largely hydrostatic equilibrium of the protoplanets themselves. The Hill-spheres are filled by hydrostatic protoplanets at least up to the critical mass, $20 \, M_{\text{Earth}}$, say, and by quasi-hydrostatic structures, typically until 50-100 $M_{\text{Earth}}$. In fact strictly static solutions for protoplanets are published up to masses that closely approach that of Saturn (Wuchterl 1993). Static isothermal protoplanets may be found with masses comparable to Jupiter's (Pečnik and Wuchterl 2004). As a consequence the protoplanetary migration problem is very far from the idealizations of essentially free gas motion in the potential of two point masses.

Because the problem is basically three-dimensional and the density structure of a protoplanet covers many orders of magnitude, additional approximations have to be made to solve the problem — either numerically or analytically. The basic analytical results (see Ward 1997) stem from solving the linearized fluid dynamical equations for power-law nebula-surface densities and an approximate gravitational potential of the problem. The starting point is an unperturbed, quasi-keplerian disk. The planet is approximated by an expansion of the perturbations induced by point mass. The linear effect (spiral density waves launched in the disk) is deduced and the resulting torques of the waves on the planet are calculated, assuming how the waves dissipate (break in the disk). If they waves (and the angular momentum carried) would be reflected and return there would be no effect. This approach has at least two potential problems:

1. the dense parts of the protoplanets, in the inner half of the Hill-sphere, say, that potentially carry a large fraction of momentum are treated as if there were no protoplanet — the density structure of the disk is assumed to be unperturbed by the protoplanetary structure even at the position of the planet’s core, certainly throughout the Hill-sphere. In that way the pressure inside the Hill-sphere is dramatically underestimated. The Hill-sphere effectively behaves like a hole in the idealized studies of the problem: the gravity of the protoplanet is introduced into the calculations, but the counteracting gas pressure of the static envelopes is omitted.

2. the unperturbed state, that is needed for the linear analysis is an unperturbed keplerian disk. But if a planet with finite mass is present, the unperturbed state is certainly not an axially symmetric disk and corrections have to be made at all azimuthal angles along the planets orbit. The keplerian disk in the presence of a protoplanet or an embryo is an artificial state that is found to decay in any non-linear calculation. Certainly it is not a steady state as would be required for a rigorous linear analysis. Therefore the approach is not mathematically correct. It may turn out that the corrections are minor, as in the case of the Jeans-swindle, and the basic results hold despite considerable mathematical violence. But unlike in the Jeans case, where Bonnor-Ebert-spheres show that there are indeed nearby static solutions, nothing similar is
available for the planet-in-disk problem. In fact the respective steady flows are essentially unknown and it is questionable whether they exist at all in the fluid dynamical problem — their might always be a non-steady planetary wake trailing the planet. High resolution calculations (Koller and Li 2003, Koller et al. 2003) indeed show considerably vorticity and important effects on the torques at the corotation resonances with potentially important consequences for the migration-rates.

Nonlinear 3D (and 2D) hydrodynamic calculations of the problem are very challenging, both in terms of time-scales and spatial scales. The ‘atmospheric’ structure of the protoplanet inside the Hill-sphere can barely be resolved even in the highest resolution calculations (D’Angelo et al. 2002, 2003) and the dynamics has to be done for simplified assumptions about the thermal structure and dynamical response of the nebula (usually locally isothermal or locally isentropic). The great value of this calculations it that they provide information about the complicated interaction of the planet with a nebula disk that can only be incompletely addressed by models with spherical symmetry that calculate the structure and energy-budget of the protoplanet with great detail. The interaction regime between the outer protoplanetary envelope, inside half a Hill-radius and the unperturbed nebula disk, at five Hill-radii, say, is only accessible by 2D or 3D calculations. It is unknown due to the lack of any reference flows, be it experimental or theoretical. This situation in my opinion is similar to the one of the restricted three body problem, at the time when numerical integrations just started.

Hence migration-rates calculated numerically or analytically have to be considered preliminary and await confirmation by more complete studies of the problem. Agreement that has been found between different investigations is within the very considerable assumptions outlined above and does not preclude considerable uncertainties in the migration-rates by many orders of magnitude.

17.4 Is a Planet a Hole or Not a Hole?

To illustrate the progress, that has been made by high-effort, state-of-the-art high resolution calculations, I briefly want to discuss a set of new calculations (D’Angelo et al. 2002, 2003) that have brought considerable insight into the problem of how accretion into the Hill-sphere of a protoplanet may occur if the protoplanet is assumed to accrete the gas into a small area, essentially onto a mature planet (D’Angelo et al. 2003). For the first time the planet was not assumed to be a point mass that accretes everything that approaches the limiting resolution of the calculation but alternative assumptions were made about central smoothing of the protoplanetary gravitational potential guided by analytical structure models (Stevenson 1982a and Wuchterl 1993) in order to look at the pressure feedback of the growing protoplanet onto the accretion flow. The flow in the Hill-sphere turned out to be qualitatively and quantitatively very different for the two assumptions: with and without
the pressure build up by the protoplanetary envelope — or in short with, or without a hole. The results demonstrate that the planetary structure feeds back on the flow, inside and also outside the Hill-sphere. Migration-rates derived from the 3D calculations were, depending on planetary mass, reduced down to 1/30 of the respective Ward (1997) analytical values.

It is important to add a note of caution to the interpretation of the planetary masses or mass-scaling used in the 2D and 3D calculations of planet-disk interaction and planetary migration. The scale of the critical mass for isothermal protoplanets with typical nebula temperatures (100 K) is \( \sim 0.1 \, M_{\text{Earth}} \), i.e. about a factor 100 below the ‘realistic’ values typically found in detailed planetary structure calculations \( 7 - 10 \, M_{\text{Earth}} \). Hence the typical regimes calculated in higher dimensional isothermal studies are a factor 100 supercritical! An isothermal dynamical calculation of planet-disk interaction, accretion and migration, for 10 \( M_{\text{Earth}} \) hence roughly corresponds to the accretion of a protoplanet of 3 Jupiter-masses, a Jupiter-mass isothermal case to a realistic one with 0.1 \( M_{\odot} \). The effective mass-scale of isothermal 2D/3D studies is \( 10 \) to \( 3000 \)\times supercritical. That scaling relates to all parts of the calculations within the gravitational range of the planet, a few Hill-radii, with the most severe effects located inside the Hill-sphere, in the protoplanetary envelope. Hence published studies of disk-planet interactions are presently in a much more violent regime than detailed 1D planet growth models require! An overlap is technically challenging but needed: studies at same effective physical scale that is set by the critical mass for the appropriate thermal planetary structure.

Calculations that treat the protoplanet \textit{and} the planet nebula interaction in detail, i.e. by accounting for the heating- and cooling-processes as well as realistic thermodynamics at the required resolution and over a significant fraction of the planetary growth time are still in the future. But a first study of the coupled problem seems to be within reach for the idealized isothermal case.

18 Towards a General Theory: In Search of the Planetary Main Sequence

The isothermal case is well studied in higher dimensional calculations of planet-nebula interaction but comparatively little attention has been payed to isothermal models of the structure of protoplanets (e.g. Sasaki 1989). The reason for that is the the isothermal assumption for protoplanets is physically unrealistic: the atmospheres of giant planetary embryos become optically thick very early in the growth (Mizuno 1980) and even for extreme assumptions about the metallicity (Wuchterl et al. 2000). The advantage of isothermal models is that they are comparatively easy to understand.

Motivated by unsolved problems of detailed statistical and dynamical models — the physical nature of the critical mass and unpredicted equilibrium structures found in dynamical models (Wuchterl 1991a,b, 1993, Wuchterl et

18.1 All Isothermal Protoplanets

The construction of all possible isothermal hydrostatic protoplanets is analogous to the construction of the main sequence for stars—both are defined by equilibria. Because of the nature of the equilibria they are independent of the history that leads to them. For the main sequence the stellar equilibria are long-lived and hence describe most of the observed stars. For planets, a similar survey to look for all equilibria has not been performed. Our knowledge presently is one of the end-states of planetary evolution—the compact cooled planets. Their best stellar analog may be white dwarfs. But in their youth, planets had rich and long lasting equilibria that have not been explored from a global point of view. First steps in the construction of a planetary ‘main-sequence’—the most probable planetary states in the nebula—have been started now. For the isothermal case, Pečnik and Wuchterl (2004) not only found the end-states (mature planets and the planetary embryo states) but a large number of previously unknown planetary equilibria. They found multiple solutions to exist in the same nebula and for the same protoplanetary embryo’s core mass. Those calculations can now be used to constrain isothermal 2D and 3D calculations such that consistent overall solutions of the planet formation problem may be found.

For the first time analogous to the stellar main sequence all protoplanets are known for the isothermal case and a statistical discussion in a diversity of nebula as well as a classification of the pathways of planet formation are now possible.

19 Formation of Pegasi-Planets

To finish the theoretical considerations I will give an example of a complete formation history of a planet from planetesimal size to its final mass. I will briefly discuss the formation and early evolution of a Pegasi-planet, from 0 to 100 Ma.

19.1 A Pegasi-Planet: Formation and Properties

The first extrasolar planet discovered in orbit around a main-sequence star was 51 Peg b (Mayor and Queloz 1995). With a minimum mass $M \sin i = 0.46 M_{\text{Jupiter}} \sim 146 M_{\text{Earth}}$, a semi-major axis of 0.0512 AU, a period of 4.23 days and an eccentricity of 0.013 it is the prototype of short period giant planets, the Pegasi-planets.

To model the formation of such an object I assume the midplane properties of a standard minimum reconstitutive mass nebula (Hayashi et al. 1985) at 0.052 AU, i.e. 0.01 of Jupiter’s semi-major axis and feeding zone masses of 15
Figure 2: Total mass (full line) and core mass (dashed line) of a Pegasi-planet forming at 0.05 AU from a solar mass star. The accumulation of a gaseous envelope surrounding a condensible element core that grows by planetesimal accretion is shown for the first 100 Mio years. The structure of the envelope is calculated dynamically including time-dependent theory of convection that is calibrated to the Sun and including detailed equations of state and opacities.
of solids (sufficient to easily reach the critical mass) and 150 $M_{\text{Earth}}$ of nebula gas, respectively. This is motivated by the fact that the $M_{\text{sin}}$ value is only a lower limit for the mass and accretion may not be 100 % efficient.

With that assumptions, the equations of self-gravitating radiation fluid dynamics for the gas (Eq. 13 to Eq. 18) with time dependent convection (Eq. 19), calibrated at the Sun, are solved for a Hill-sphere that is embedded in the standard nebula at 0.05 AU. The condensible element core at the center of the sphere grows by planetesimal accretion according to a the particle-in-box rate. The minimum-mass solid-surface-density (a safe lower bound) and a gravitational focusing factor of 3 (Safronov-number of 1, also an assumption towards slow growth) are used.

19.2 Mass Accretion History – The First 100 Mio Years

The resulting mass accretion history from zero to 100 Ma is shown in Fig. 2 with a logarithmic time-axis. Age zero is chosen, following Wuchterl and Tscharnuter (2003) at the moment when the envelope becomes optical thick for the first time and hence a thermal reservoir is formed. The calculation starts at a core of $\sim 10$ km size and a mass of $\sim 10^{15}$ kg. The displayed evolution starts 10 ka after the embryos envelope became optically thick, at roughly a tenth of a lunar mass. At that time the total mass (full-line) and the core mass (dashed) are essentially the same because the envelope mass is negligible. At 200 ka and somewhat below 10 $M_{\text{Earth}}$ the two curves separate, due to a gaseous envelope of significant mass developing. The planetesimal accretion-rate at that time has already dropped due to depletion of the solids, and the core-growth-curve $M_{\text{core}}(t)$ starts to flatten out, with the total mass following. As the critical mass is approached, the total mass curve turns upward with the core-mass flattening further. That shows the onset of efficient envelope accretion. The contraction of the envelope is still quasi-static and the gas is practically at rest. The step in the total mass reflects a period of efficient envelope accretion that rapidly increases the total mass until the feeding zone is essentially emptied. The Mach-numbers are finite but during this stage, but the hydrodynamical part of the flow is basically a transition flow from the nebula onto the contracting inner parts of the protoplanetary envelope that are quasi-hydrostatic. After the flow from the feeding zone onto the planet has faded, the masses remain constant - a Pegasi-planet is born.

19.3 Luminosity of a Young Pegasi-Planet

The luminosity of the Pegasi-planet corresponding to the mass-accretion history above, is shown in Fig. 3. The luminosity increases during mass-growth, passes through a double maximum and then decays roughly exponentially. The two luminosity maxima reflect the maximum accretion of solids and gas, respectively. Initially the growth, starting approximately at a $nL_{\odot}$, rises due to the planetary embryos increased, gravity enhanced cross section for planetesimal accretion. As the planetesimals are removed from the feeding zone
and incorporated into the embryo, the surface density of the remaining condensible population fades and the luminosity turns over. Planetesimal accretion passes thermal control to the contracting gaseous envelope. As the envelope mass becomes comparable to the core its contraction controls the luminosity of the planet. With approaching the critical mass the luminosity turns upward again due to the rapid growth of the envelope reaching the sharp peak at maximum accretion. With the arrival at the final mass no further material is added and the only luminosity supply is contraction of the envelope, that slows down as larger parts of the planet degenerate. Thereafter the planet cools into the present with its luminosity being inverse proportional to age.

Most of the planetary evolution turns out to be quasi-hydrostatic with a brief dynamical period around maximum accretion: the step-like increase in Fig. 2 and the narrow luminosity peak in Fig. 3. During this brief period, most of the mass is brought into its final position and acquires its initial temperature. The rapid, dynamical phase is so fast, that there is essentially no thermal evolution occurring. Hence it sets the initial thermal state, and determines the bulk starting properties of the planet’s evolution at its final mass.

The luminosity of the planet during this period lasting a few hundred years at an age of a few hundred thousand years is shown in Fig. 4. The entire evolution shown in this Fig. 4 is present in Fig 3 but unresolved in the luminosity spike. The peak accretion-phase starts at the turnover of the luminosity, cf. Fig. 4. Because the contraction is rapid the outer parts of the envelope are
adiabatically cooled and the nebula gas starts radiating into the protoplanetary envelope. The luminosity becomes increasingly negative as the inner, most massive parts of the protoplanet contract further. The trend is reversed when the contraction of the central parts are slowed down again and the heat produced in the process reaches the outer boundary of the planet for the first time (the brief spike at 180 years in Fig. 4). Then the overall contraction and accretion of the planet takes over again and the luminosity rises to positive values, reaching the peak at maximum gas accretion. As accretion fades the luminosity turns over and decreases with the vanishing amount of nebula gas remaining in the gas-feeding zone. At the end of significant gas accretion, the luminosity-slope sharply changes, as the luminosity becomes determined by contraction alone (beyond 220 a in Fig. 4). The initial luminosity of the planet at its final mass is \( \sim 1 \, \text{mL}_\odot \) but fades rapidly. The total width of the spike in Fig. 3, at a tenth of this peak value, i.e. at \( 10^{-4} \, \text{L}_\odot \) is appr. 100 000 a.

The final mass of the planet is approximately equal to the \( M \sin i \) of 51 Pegasi b. Being below a Jupiter mass (cf. Wuchterl et al. 2000) there is a chance that spherical symmetry will give fairly correct results not only up to 100 \( M_{\text{Earth}} \) but even for the gas accretion-rates towards peak-accretion and stagnation when the feeding zone is emptied. The rapid transition from 20 to 150 \( M_{\text{Earth}} \) also may provide a way to escape violent migration by quickly removing any relevant, migration-driving mass from the nebula disk before significant orbital decay occurs in type-II migration mode.

The critical assumption concerning the above calculations is whether feed-
ing zones with the assumed masses are plausible. In a minimum mass nebula, the mass integrated over plausible feeding zones is only a few earth masses. But to understand extrasolar planets and planet formation in general, the minimum mass nebula that is reconstructed from the solar system is a much too narrow constraint for the plausibility of the available mass. Most likely the nebula will be gravitationally stable during the planet formation epoch. That requirement allows to derive a more general constraint on the feeding zone. The gravitational stability of disks can be roughly estimated by the Toomre-criterium for axial symmetric gravitational stability. I suggest to use the marginally Toomre-stable nebula as the limit for the variety of nebula conditions that are allowed for consideration. The so obtained maximum feeding zone masses for planet forming nebula are a factor 200 above the minimum mass at 0.05 AU. In a very conservative feeding zone of one Hill-radius around the orbit, a Jupiter mass ($318 M_{\text{Earth}}$) can easily be fit into such a feeding zone. Typical feeding zones would have more then five times the radius and hence provide a volume around the orbit that is 25 times larger. Hence many Toomre-stable nebulae can provide sufficient mass in plausible feeding zones at 0.05 AU to form a giant planet. The above calculations hence shows, that giant planets form in less then a million years at 0.05 AU if their orbits are stable, i.e. migration-rates remain small.

Relying on a standard minimum mass nebula and planet formation fluid-dynamics that are physically improved but following a pre-discovery, simple model-setup, it is possible to explain the formation of a Pegasi-planet in-situ, provided there is sufficient mass of gas and solids in the feeding zone. That is not the case for a minimum mass nebula. But assuming the diversity in extrasolar planets originates from a diversity in disk properties, we may vary the global nebula parameters. Gravitationally stable nebula that have less total angular momentum and hence more mass closer in, may provide sufficient mass to dynamically grow giant planets at 0.05 AU.

20 \textit{Is it} Misleading or Not?

I have shown that one of the surprising discoveries, the Pegasi-planets can be reconciled with standards solar system formation theory, if a diversity of nebula is accepted. I will now return to the empirical bases for the doubts raised about or allegedly solar-centric understanding of planet formation: the exoplanet discoveries.

20.1 \textit{Is There No Bias \ldots ?}

One of the first questions that immediately occurs is whether the discoveries are biased, as is very usual in astronomy because faint things are harder to see. Possible sources of bias introduced by the dominating radial velocity technique are:
1. sensitivity of the radial velocity measurements (highest for large masses and short periods),

2. the planet hunting grounds and hunting tactics,

3. the discovery race,

4. the binary issue, i.e. that the binary fraction in the galaxy is much larger than in the typical samples of radial velocity planet searches,

5. the selection of suitable host-stars to avoid variability, activity, youth, giants etc. ?

6. the extraction of reliable planetary signals from the data — two massive planets with widely separated orbits are easier to identify than two relatively low mass planets with comparable masses and relatively close orbits. That is especially true when the orbital periods are close to multiples of each other as in case of Jupiter and Saturn.

That questions can only be answered by very well-defined and complete samples and to some extent by other discovery methods. It is interesting to look at the planetary ‘yield’ of the ongoing transit-searches. They have by far not detected as many planets as would be (naively) expected from an extrapolation of the radial-velocity discoveries. It has to be seen whether this is due to difficulties of the transit method, or just due to different biases of radial-velocity and transit-searches. Clearly an overlap of methods is important and seems to be possible for astrometry and direct imaging within a few years.

20.2 Towards Normality

With the radial velocity method providing by far most of the information about extrasolar planets it is interesting to look at how the typical properties of the discovered exoplanets change as more and more are discovered. It seems notable that:

- the ‘outskirts’ of the eccentricity distribution approaches the solar system planets with an overlap in all parameters expected soon,

- the periods of the discovered planets increase with time, now starting to overlap with the solar system giant planets. Waiting seems to make the solar system more typical,

- the median of the distribution of minimum masses seems to continue to decrease. It seems that the characteristic mass has changed from about \( M \sin i = 7 \) in 1996, to 4 in 2000 and I understand observer’s talks such that \( M \sin i = 2 \) or 1 may be possible for the final outcome.

The next important step in the discoveries is an extrasolar planet that overlaps with a solar system giant planet in all properties, i.e. mass and eccentricity less then Jupiter’s, orbital period larger than Jupiter’s.
20.3 Brave Hearted Searches — Icarus!

To close the gap, searches are necessary where they are most difficult:

1. avoiding ‘hunting bias’, i.e. without a-priori input to select stars for planetary yield or assigning higher observation priority to ‘good’ stars, i.e. stars with low radial-velocity-‘noise’;

2. volume complete samples (see next section),

3. searches for planets in binaries (e.g. α-Cen, Endl et al. 2001)

4. planets of stars with a type earlier than late F — most RV samples focus on stars later than that. But Setiawan et al. (2003) detected a planet around an K1III giant, a star that was an A-star on the main sequence, expanding the mass range of known stars with planets,

5. planets in clusters (as opposed to cluster planets), that have the advantage of a more homogeneous and coeval stellar populations,

6. searches for young planets to determine the earliest time at which planets exist,

7. the host star mass-range is also considerably widened by studying M-stars. Kürster et al. (2003) showed that planets with $M\sin i$ of a few $M_{\text{Earth}}$ could be detected around M-stars,

8. Guenther and Wuchterl (2003) went to the extreme and searched for planets around brown dwarfs. While significant RV-variation was detected only an upper limit could be set for the presence of Jupiter mass planets.

9. direct imaging searches to look for planets in long-period orbits and start of the direct characterization (Neuhäuser et al. 2002, for a young brown dwarf).

20.4 Looking at Stars Near You — Metallicity

To show the present state of the discussion and possible problems with biases I will briefly discuss the planet-metallicity relation (Santos et al. 2003, Fisher et al. 2004). The authors find the frequency of planets to increase with metallicity (i.e. the [Fe/H] metallicity indicator). Fuhrmann (2002, 2003) studied a volume complete sample of nearby F,G and K stars that overlaps with the planet hunting samples. When compared to the Fisher et al. result and assuming the use of the Nidever et al. (2002) volume limited sample the following is noticed (Fuhrmann 2003, pers. comm.) when comparing it to the volume complete Fuhrmann-sample: Of the 166 Fuhrmann stars only 90 (54%) are in Fisher et al. study despite the smaller volume but complete Fuhrmann-sample. Missing in the Fisher et al. study are:
1. a few subgiants,
2. fast rotating stars with $v \sin i > 10$ km/s,
3. a few young and chromospherically active stars,
4. stars without a precise luminosity class,
5. binaries and multiple systems,

These are all properties that make planet hunting more difficult. From these findings of Fuhrmann, I can only conclude that stars that are unfavorable for planet hunting are under-represented. The average $[\text{Fe/H}] = -0.02$, of the 90 Fisher et al. stars, that fall into the intersection with the Fuhrmann-sample is only 0.01 dex higher than Fuhrmanns respective value. However, for $[\text{Fe/H}] > 0.2$ only 5 stars are missing but for $[\text{Fe/H}] \leq -0.2$, 13 stars are missing. Hence there is a slight trend the metal-poor stars are preferentially missing. Considering the still small numbers that are available for comparison, and the average metallicity of $[\text{Fe/H}]= 0.00$, for thin disk stars, there could be a metallicity effect of $+0.10$ to $+0.15$ (contrary to the $+0.25$ that are favored by Santos et al.). With the Sun at the thin disk average of $[\text{Fe/H}]$ for our distance to the galactic center (Fuhrmann 2003) the role of metallicity may well be a slight increase overall due to a significant increase for the Pegasi-planet’s host stars. That would put the solar system into normality as far as metallicity and planet hosting are concerned.

It is such a kind of bias, that I think might still be present in the planetary discoveries. That should be considered before abandoning what we know about the solar system.

## 21 Was the Solar System Misleading?

All things considered, was the solar system misleading? The Sun lead to much of modern physics and astrophysics. It is the calibrator of stellar evolution and the age of the universe. The Solar System offers the best traces for planet formation studies. But there are stars other than the Sun and there are planetary systems other than ours. After the dust of has settled, I think the solar system will still provide the basis to understand planetary diversity. It is not good to blame the path when getting lost!

The dominance of strange planets and the increased metallicity of planet hosts may be a result of hunting biases in the exoplanet sample. The close-in giant planets were not explicitly excluded by most investigators of planet formation. To a large part, the question of such planets was not considered in sufficient detail. The high eccentricities in the presently known exoplanets may be due to two effects:

1. they could represent the high-mass end of planet formation. Dynamically larger masses result in larger planet-planet interactions that typically increase the eccentricities;
2. eccentric planets might be easier to detect, because an eccentric planet excludes a wider range of neighboring planets according to Hill-exclusion stability criteria. As a result, for a sufficiently eccentric planet, there is no neighboring, competing signal, that causes confusion, thus not adding ‘planetary noise’ due to the unidentified additional radial-velocity-signals of neighboring (and smaller) planets.

We need careful studies of the observation biases, most importantly we need the study of predefined complete samples and we need an analysis of what planets may be extracted first if every planet host has a planetary system that is as dynamically filled as ours.

22 Planetendämmerung

To close, let us imagine we would start astronomy and wait for the stars appearing in the evening for the first time. Located at Freiburg, say we might only be able to get to a visual limiting magnitude of $\sim 3$. Imagine, that our entire knowledge of stellar astronomy would have to be derived from, at first, the ten brightest stars, then the 100 brightest stars, that are the first to become visible. What would be the mean stellar radius, that we would deduce for our nascent stellar astronomy? The answers — taken from the *Catalogue of Apparent Diameters and Absolute Radii of Stars*, are:

- for the 10 brightest stars the average radius is $36 R_\odot$, and
- for the 100 brightest stars the average radius is $32 R_\odot$.

This is not the astrophysics that we know. Similarly, I think, we will see more planets ‘appear’ on the sky! Maybe they will be more like the solar system planets.

23 Towards a Broad View

The next steps towards developing a broader picture into which the solar system and the extrasolar planets fit are:

1. the development of a general theory of planet formation for a diversity of protoplanetary nebulae,

2. instruments that are sensitive to the entire mass spectrum of the giant planets and below. The french-lead european COROT-mission with german participation via the DLR, should reach that goal for the first time if the planned launch in spring 2006 is successful.

3. In autumn 2007, the first jovian year will have passed since the discovery of 51 Peg b and many of the searches should have the first complete orbits of Jupiter like planets in their data.
Finally we will have the required sensitivity for a full Jovian orbit and the theory significantly advanced. Theoretical work is ongoing in preparation for that events. Unlike with 51 Peg b, this time theory will be prepared.

24 Conclusion

Are the first 100 exoplanets misleading?

After more than 2500 years of astronomy we have crossed the ocean of space between us and the stars. New worlds become detectable and in reach of analysis. What has been discovered was unexpected in detail but has shown that planets are abundant in the galaxy as expected. Much can be understood on the basis of what has been learned about planet formation from the solar system, but many problems remain.

After crossing an ocean the first cliff leads to land, but don’t try to get to close to it.

References

Blum, J. and G. Wurm 2000, Icarus 143, 138
Cassen, P.M. and Moosman, A. 1981, Icarus 48, 353
Chambers, J. E. 2001, Icarus 152, 205
Dvorak, R., 1984, Celestial Mechanics, 34, 369
Dvorak, R., Froeschle, C., and Froeschle, Ch., 1989 A&A, 226, 335
Guillot, T. 1999, Science 286, 72
Harris, A. W. 1978, Lunar. Planet. Sci. 9, 459
Hayashi, C. 1980, Prog. Theor. Phys. Suppl. 70, 35
Inaba, S., G. W.Wetherill, and M. Ikoma 2003, Icarus 166, 46