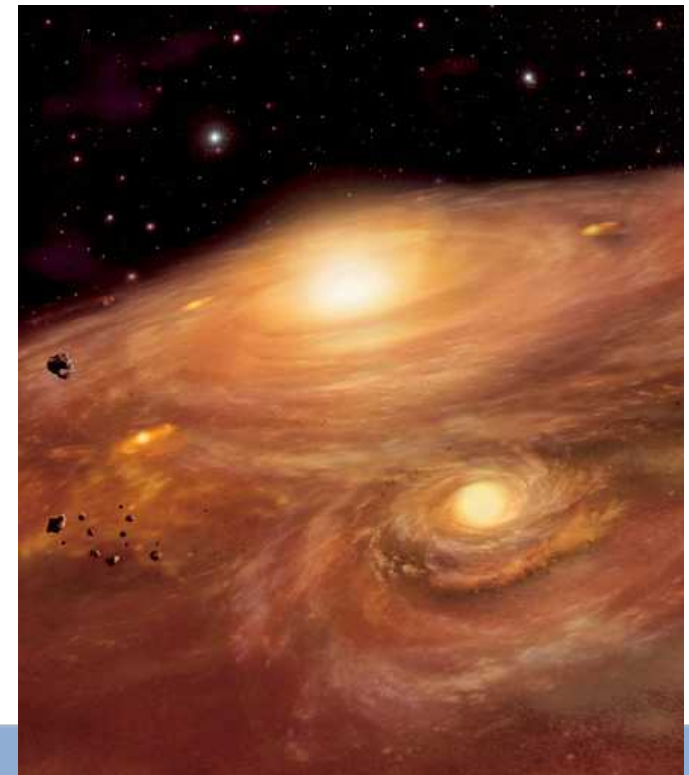


Planet formation and migration

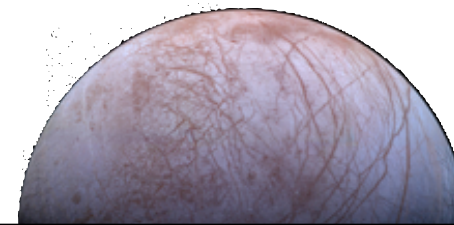
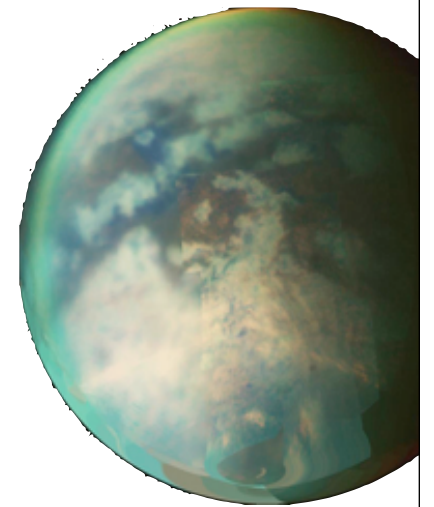
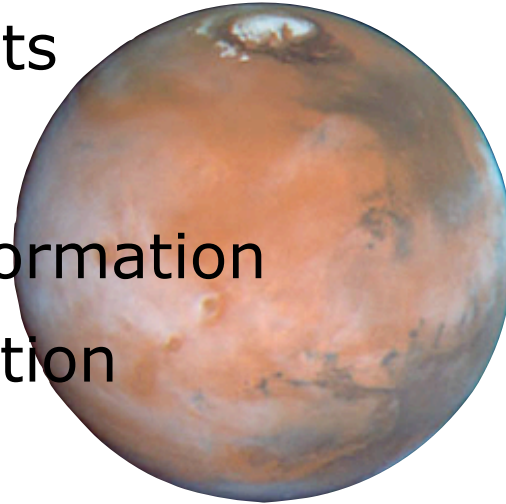
Planetary Sciences Chapter 12

Daphne Stam



Outline

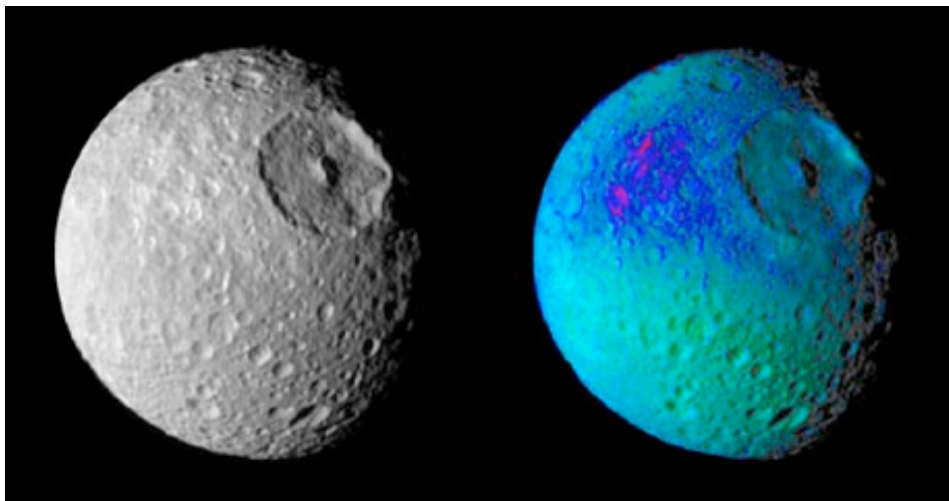
- Observational constraints
- Star formation
- Early stages of planet formation
- Terrestrial planet formation
- Giant planet formation
- Planetary migration
- Asteroids and comets
- Moons and rings



Observational constraints

Any theory on the formation and evolution of planetary systems must explain the following observed properties of the Solar System and exoplanetary systems:

- Orbits, spacings, and rotation of the planets
- Sizes, densities, compositions, and ages of the planets
- Properties of asteroids, comets, moons, rings
- The surface structure of terrestrial planets and moons
- The composition of planetary atmospheres

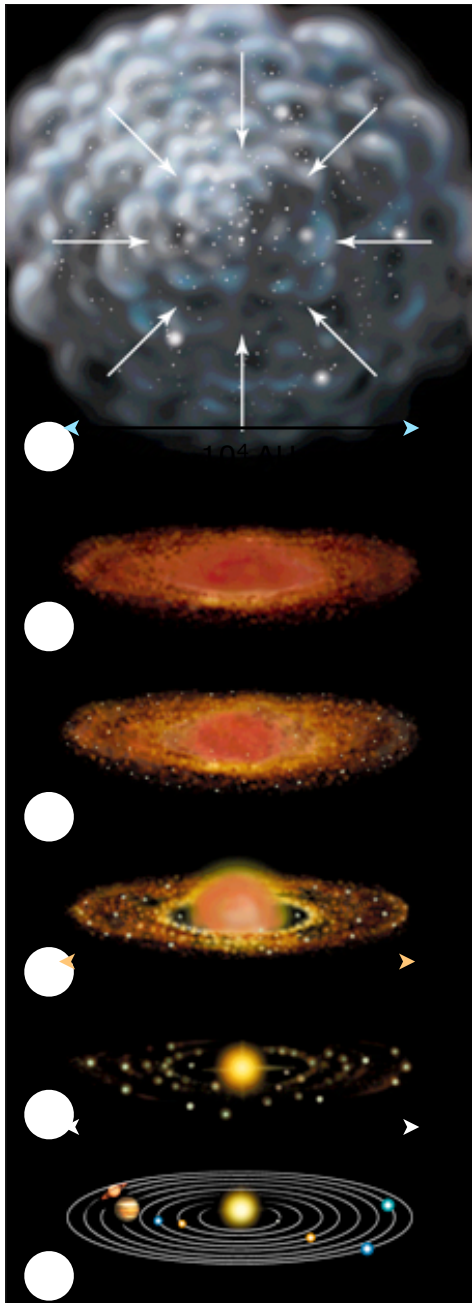


Mimas, a moon of Saturn, in total flux (left) and in false color to emphasize differences in surface composition (right).

The moon's diameter is about 400 km. The huge crater (Herschel) is 140 km across, its walls are 5 km high, and the central peak 6 km [credit: Cassini/NASA/ESA].

The birth of the Solar System in a nutshell

(the Nebular hypothesis was originally proposed by Kant and Laplace in the 18th century)



1. Collapse of an interstellar molecular cloud $t=0$
2. Formation of a circumstellar disk
3. Birth of the protostar $t=10^5 - 10^6$ yr
4. Development into main-sequence star $t=10^6 - 10^7$ yr
5. Formation of planetesimals
6. A mature planetary system has been formed $t>10^7$ yr

Interstellar molecular clouds



A giant molecular cloud in Ophiuchus

Credit: Loke Kun Tan/starryscapes.com

Properties of molecular clouds:

- composed mainly of H_2 and some He
- small fraction of other molecules:
CO, CN, CS, SiO, OH, H_2O , HCN, SO_2 ,
 H_2S , NH_3 , H_2CO , etc
- masses vary from ~ 1 to $\sim 10^5$ - $10^6 M_{Sun}$
- temperatures are ~ 10 - 30 K
- densities vary from $\sim 10^3$ to 10^5 cm^{-3}

Molecular clouds are usually stable against self-gravitational collapse because of internal gas pressure, magnetic fields, turbulent motions, rotation, ...

Collapsing interstellar molecular clouds

The *virial theorem* describes a system in equilibrium:

half the gravitational potential energy of the system equals its kinetic energy:

$$E_G = 2 E_K$$

Assume that the kinetic energy of our cloud equals its thermal energy:

$$E_K = N \frac{3}{2} kT \quad N: \text{total number of atoms, } k: \text{Boltzmann's constant, } T: \text{temperature}$$

The gravitational potential energy of our cloud equals:

$$E_G = \frac{3}{5} \frac{GM^2}{R} \quad G: \text{gravitational constant, } M: \text{total mass, } R: \text{radius of the cloud}$$

Thus, for a cloud in equilibrium (no collapse, no expansion) we have:

$$\frac{3}{5} \frac{GM^2}{R} = 3 N kT$$

A cloud will expand, if: $\frac{3}{5} \frac{GM^2}{R} < 3 N kT$

A cloud will collapse, if: $\frac{3}{5} \frac{GM^2}{R} > 3 N kT$

Collapsing interstellar molecular clouds

Assuming the cloud has a constant density ρ , and m is the mass per atom, we can solve for the critical mass of a cloud, the Jeans mass, M_J :

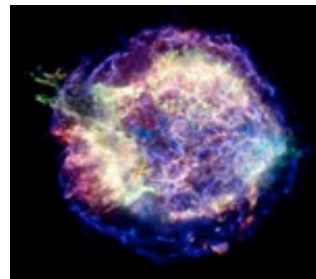
$$M_J = \left(\frac{5 kT}{Gm} \right)^{3/2} \left(\frac{3}{4 \pi \rho} \right)^{1/2}$$

A cloud with $M > M_J$ will collapse if its only means of support is internal thermal pressure.

Typical values for $M_J \sim$ several 1000 solar masses.

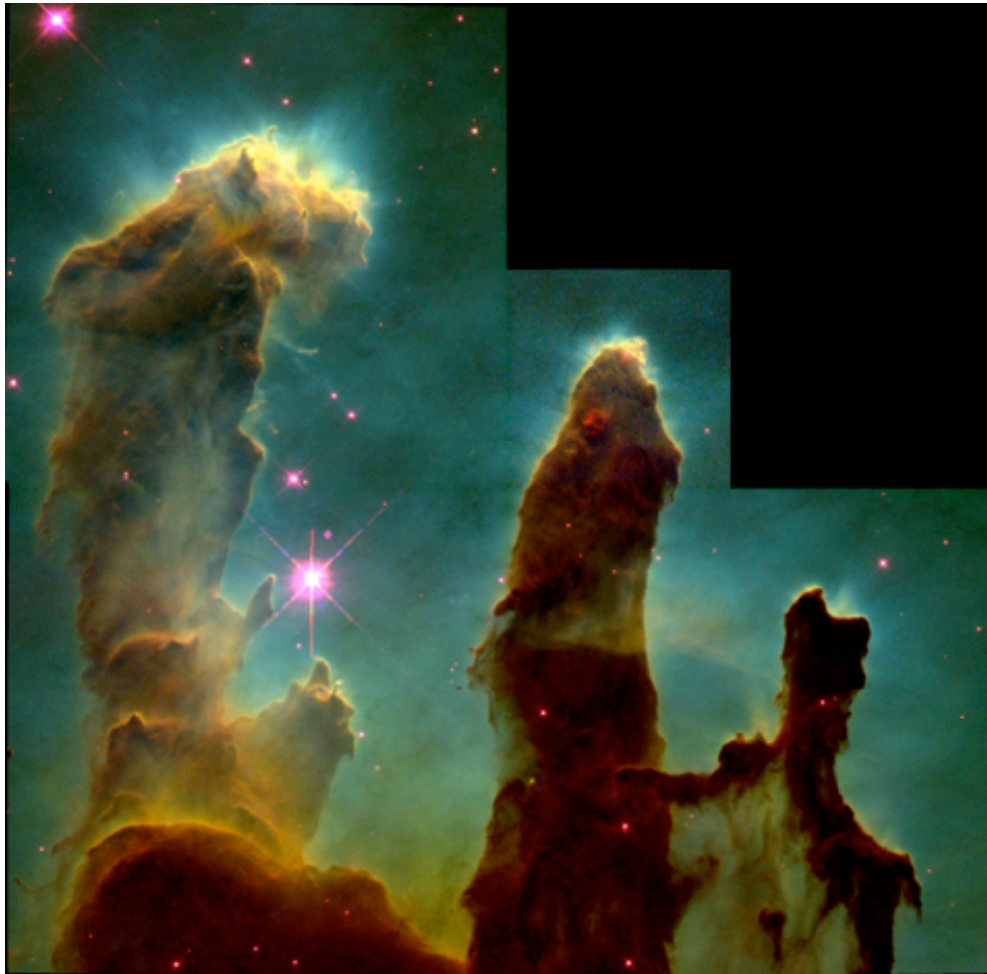
Possible triggers for collapse:

- galactic pressure changes
- nearby supernova explosions
- expanding hydrogen-regions



The birth of stars:

Stars are born in the densest regions of the collapsed molecular clouds:



Starforming region in the Eagle Nebula

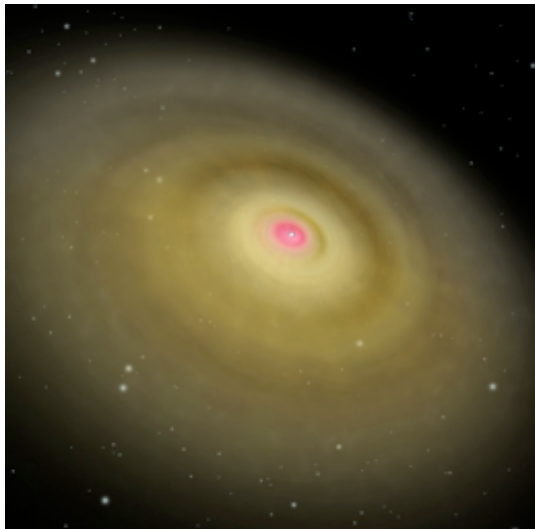
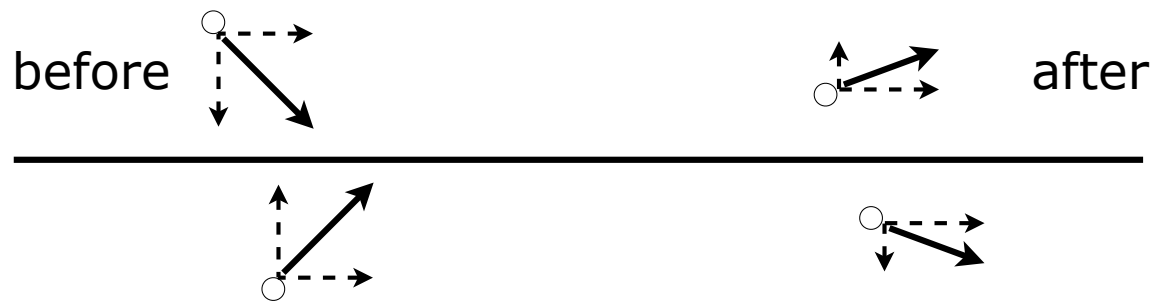
Credit: HST NASA/ESA

The evolution of young stars:

- the core temperature T increases
- as long as the cloud is sufficiently transparent, excess heat is radiated away, and T remains relatively low
- because of the increasing density, the transparency decreases, and T increases
- once $T > 10^6$ K, deuterium fusion starts, and contraction slows down
- once deuterium runs out, the contraction starts again, T increases
- once $T > 10^7$ K, hydrogen fusion starts, and contraction slows down

The formation of protoplanetary disks

Material from the molecular cloud collapses towards the core, but conservation of angular momentum prevents it from falling onto the protostar. It will orbit the core and settle into a disk around it:

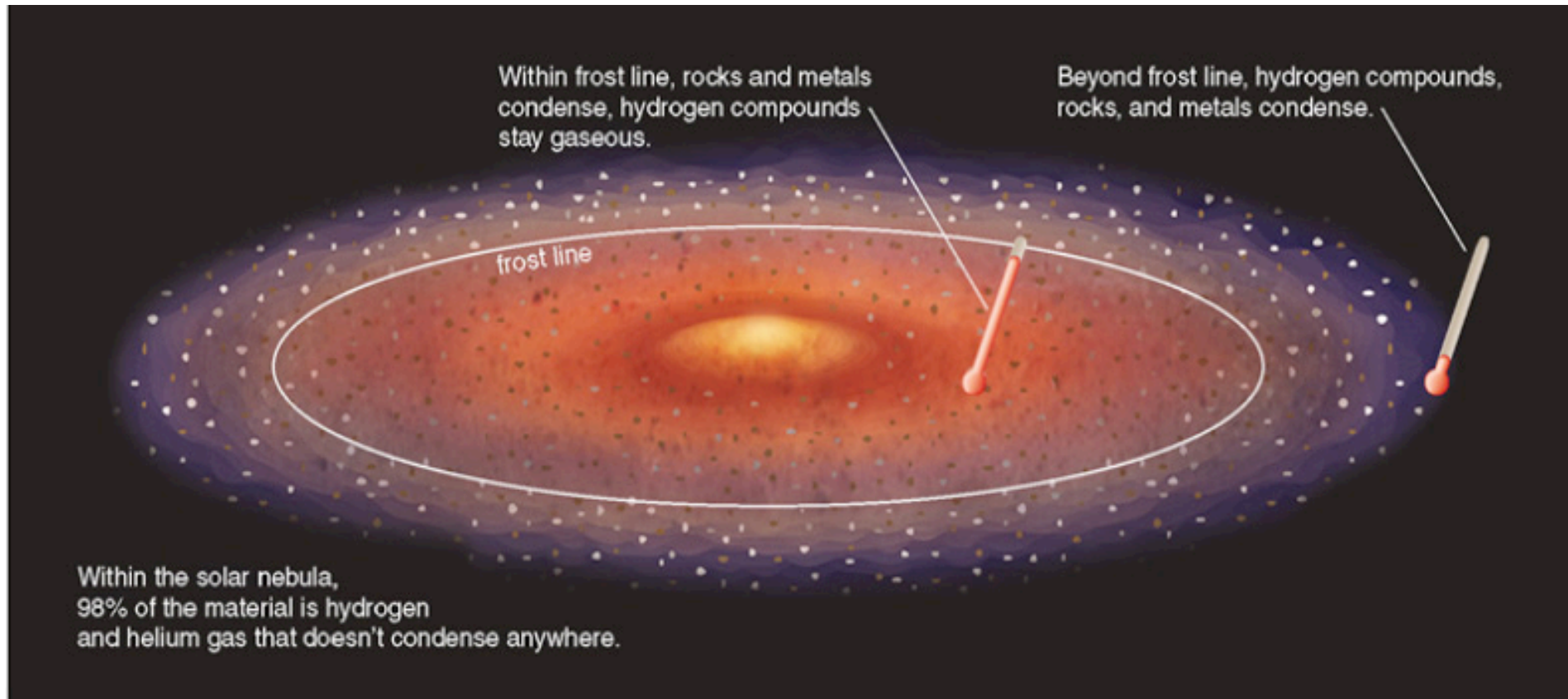


Credit: Subaru Telescope

Excess gravitational potential energy is dissipated as heat in the forming disk. The closer to the core, the higher T , because the more potential energy is released ...

Chemistry in a protoplanetary disk

The chemical reactions in the protoplanetary disk depend strongly on the local temperatures, pressures, and mixing processes:

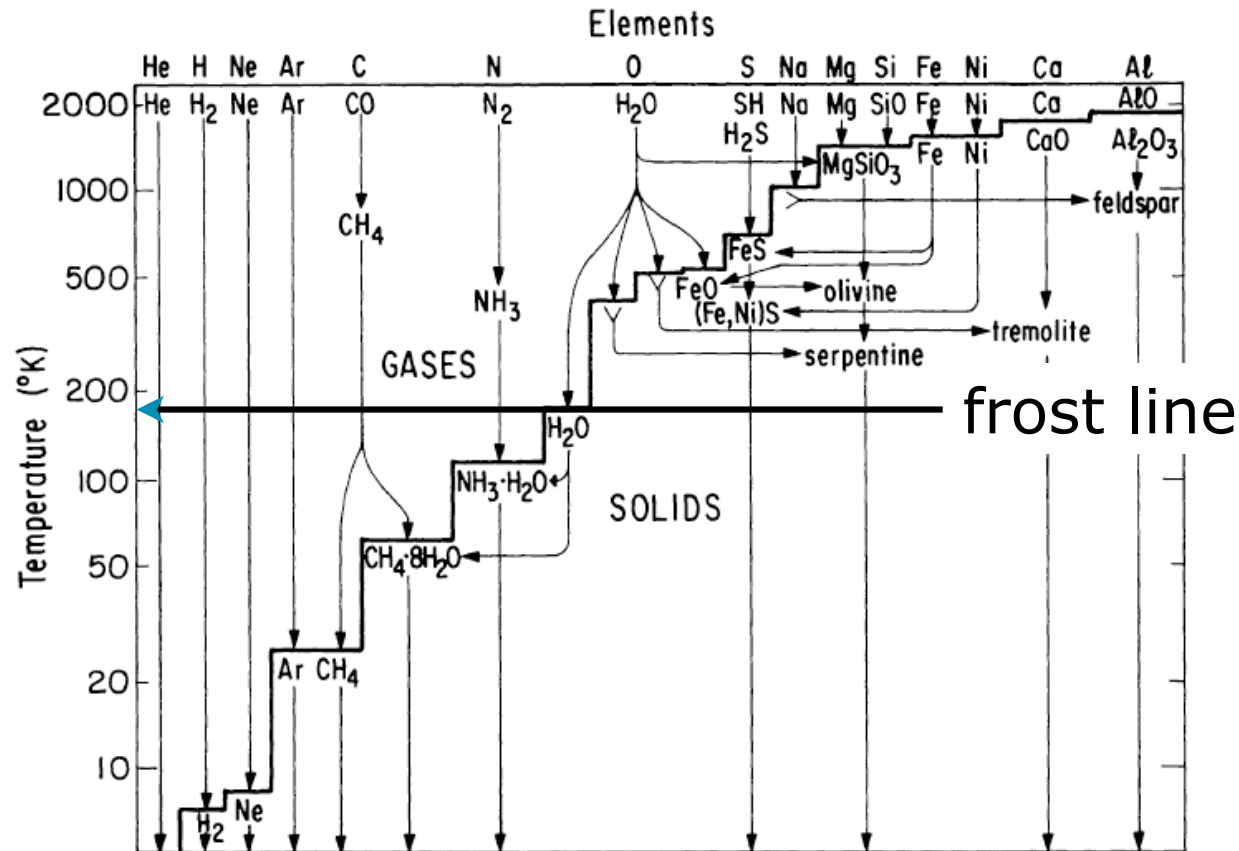


Credit: Addison & Wesley Publishing

The location of the frost/snow/ice line depends on the luminosity of the star, and the density and composition (mixing) of the disk

Condensation in a protoplanetary disk

The major reactions during fully equilibrated cooling of solar material from 2000 to 5 K, at a pressure of $\sim 10^{-3}$ bar (from Barshay & Lewis, 1976):

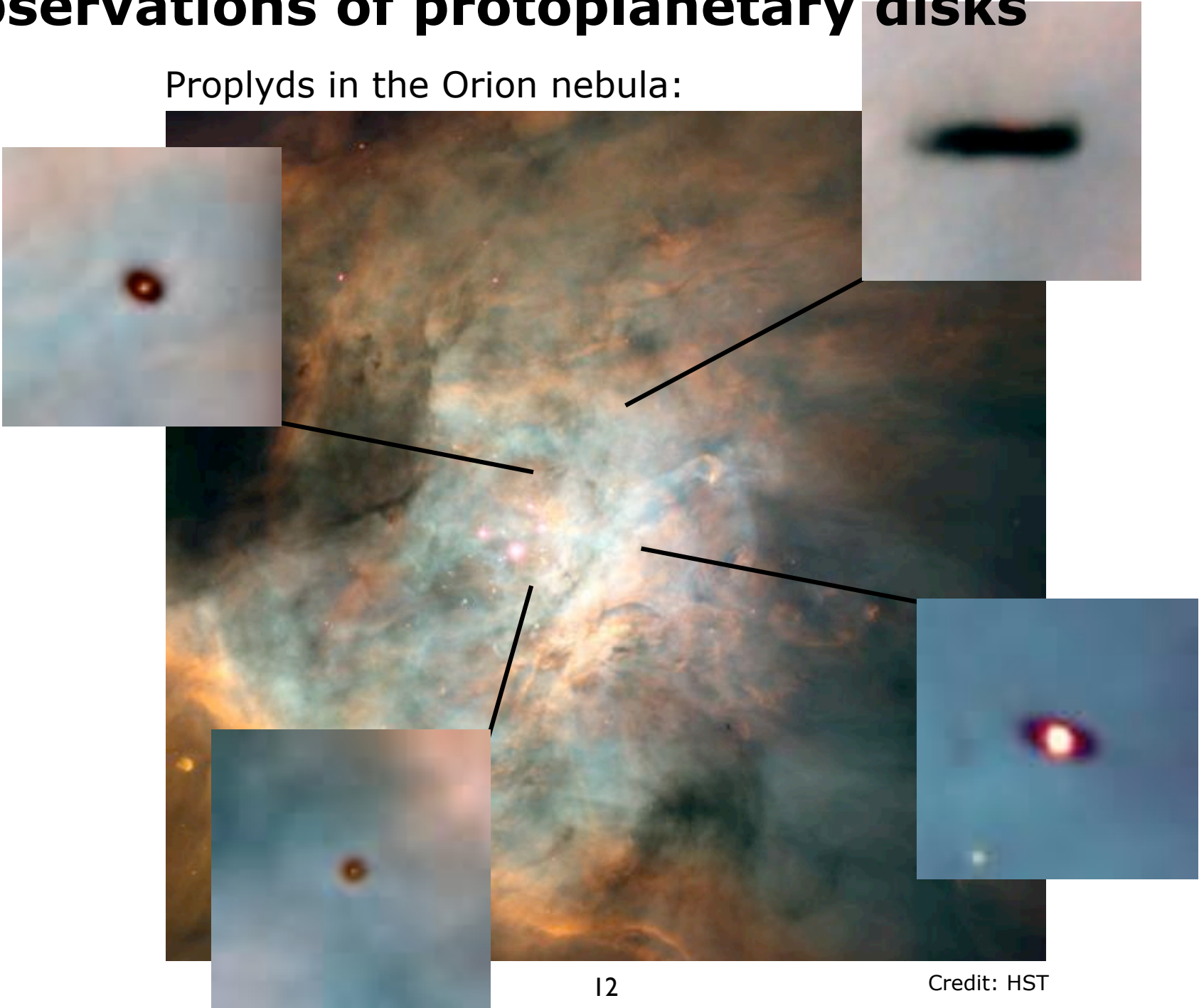


Note 1: at low temperatures (outer disk regions) there is no full chemical equilibrium

Note 2: dynamical mixing processes, transport, and e.g. cosmic ray ionisation are also important

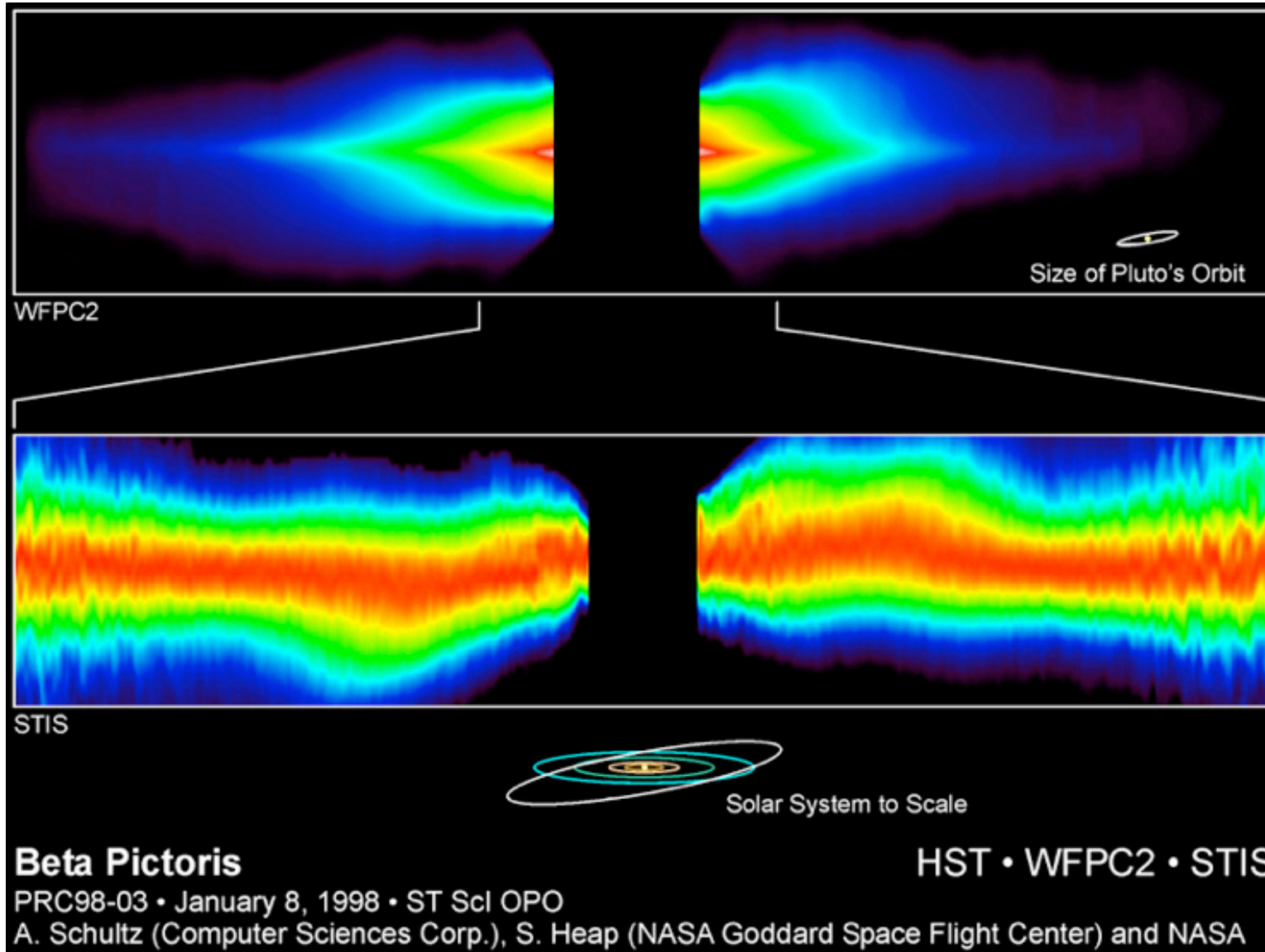
Observations of protoplanetary disks

Proplyds in the Orion nebula:



Observations of protoplanetary disks

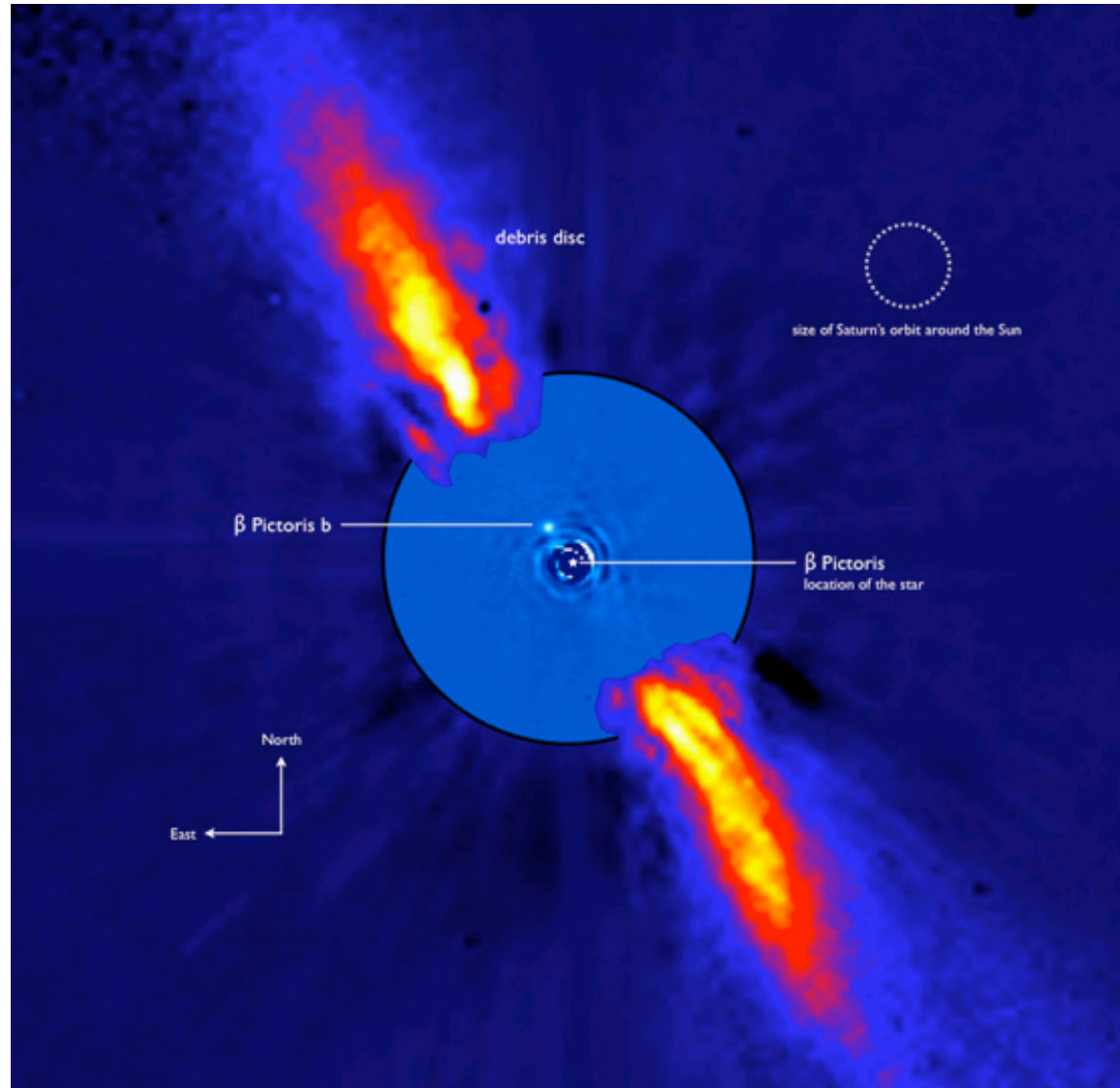
The dust disk around the star β -Pictoris:



This debris disk was first imaged in 1984

Observations of protoplanetary disks

A planet in the dust disk around the star β -Pictoris?



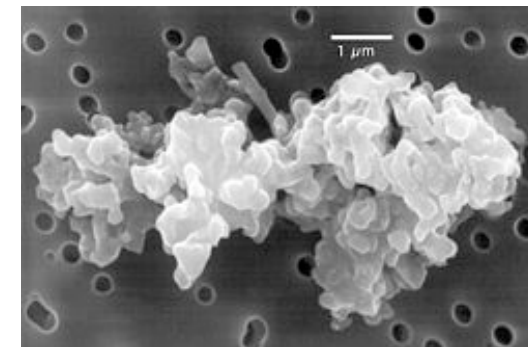
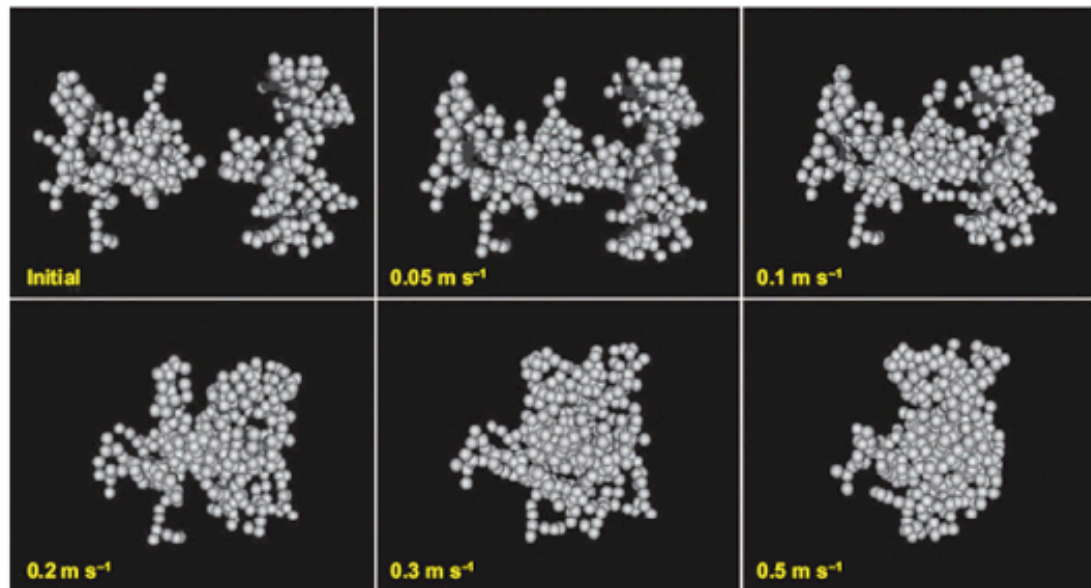
Credit: ESO/Lagrange et al. 2008

Early stages of planet formation

As a disk of gaseous matter cools, various compounds start to condense into microscopic grains that grow through collisions :

- At high temperatures: silicates and iron compounds
- At lower temperatures: water-ice and other ices

Simulations by Paszun & Dominik for collisional growth of dust aggregates assuming different impact velocities:



Interplanetary dust particle
(Jessberger & Brownlee)

Small (< cm) particles move with the flow of the gas ...

Formation of small planetesimals

Intermediate bodies (~ 1 m) are formed through collisions of the fluffy, small particles.

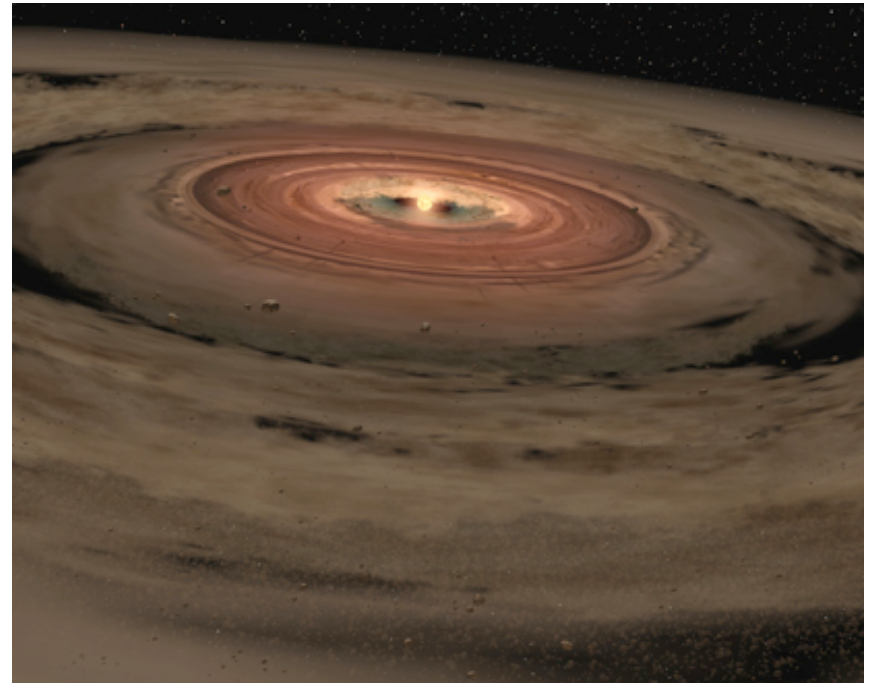
A problem:

The gas rotates slower than the keplerian speed of these bodies ...

Intermediate bodies will suffer from gas drag and spiral towards the star ...*

While large ($> \text{km}$) planetesimals will not suffer from the gas drag ...

The transition from m to km sized bodies should take place rather quickly and is not understood yet!



* A meter-sized body at 1 AU would spiral into the Sun in ~ 100 years!

Formation of protoplanets

Kilometer-sized bodies will attract each other gravitationally and grow through collisions.

The largest planetesimals will accrete almost everything that they collide with: *runaway* or *oligarchic accretion*

The isolation mass, M_i , the largest mass (in grams) to which a planetesimal orbiting a $1 M_{\text{Sun}}$ star can grow by runaway accretion is:

$$M_i \sim 1.6 \times 10^{25} (r^2 \sigma)^{3/2}$$

Here, r is the orbital distance and σ is the disk's surface density (in g cm^{-2})

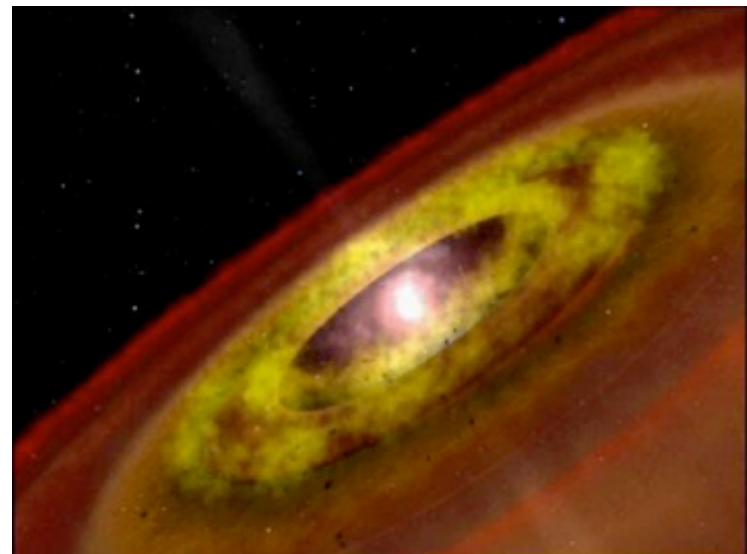
Without 'fresh' material and assuming a disk mass of $0.02 M_{\text{Sun}}$, $M_i = 0.08 M_{\text{Earth}}$ for the Earth, and $\sim 0.1 M_{\text{Jupiter's core}}$ for Jupiter's core ...

Gas-clearing stage of the disk

Since there is no gas left between the planets, the gas must have been cleared away at some stage during the evolution process.

Possible explanations:

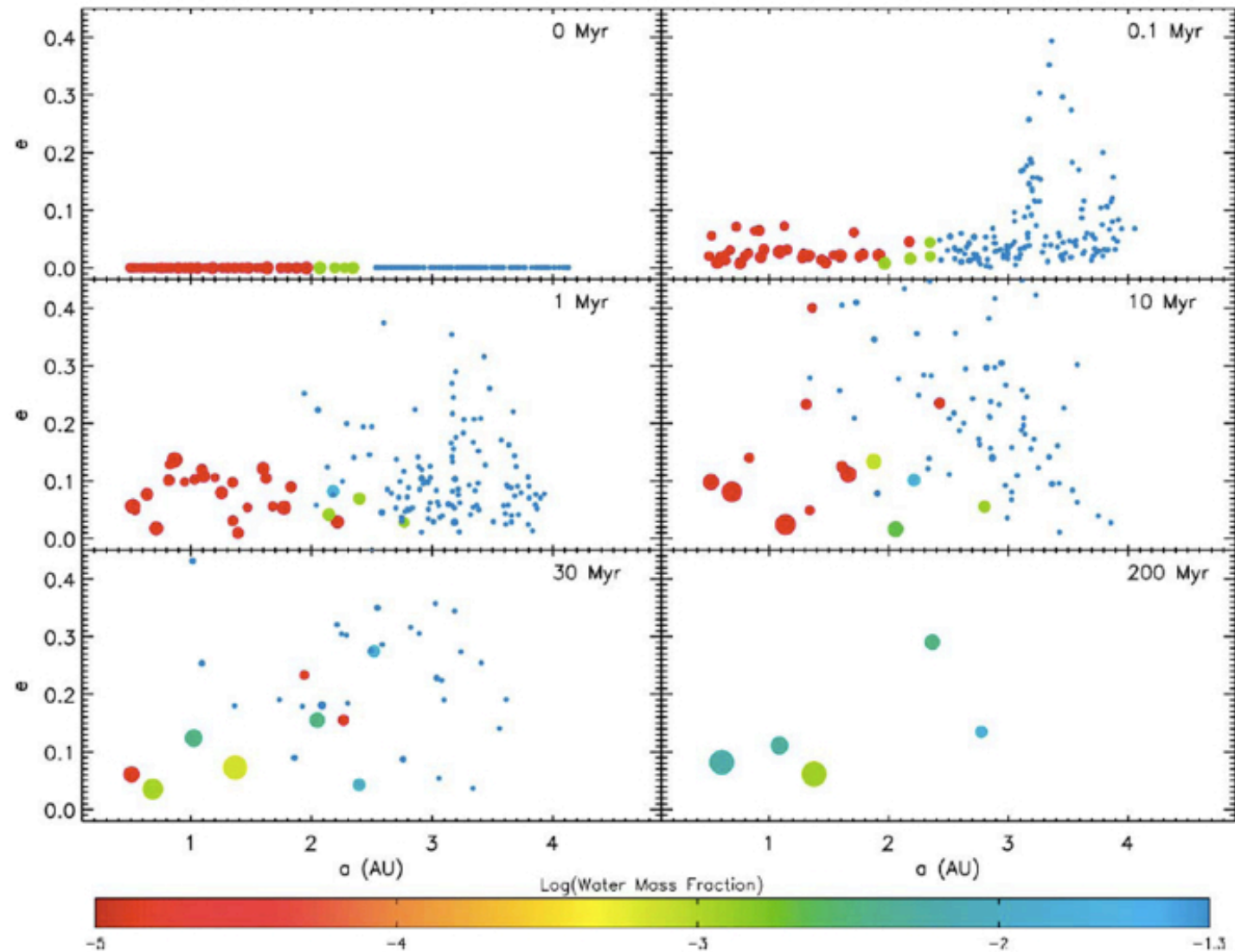
- $\sim 10^6$ - 10^7 yr after its formation, a pre-main sequence star goes through its T Tauri phase with very strong stellar winds that 'blow' the gas out of the system
- the gas 'evaporates' from the disk due to the UV flux of the young star itself or of nearby massive stars



The sun as a T Tauri star. Credit: NASA

Terrestrial planet formation

Numerical simulations of planetesimal accretion (with Jupiter at 5.2 AU):



Giant planet formation

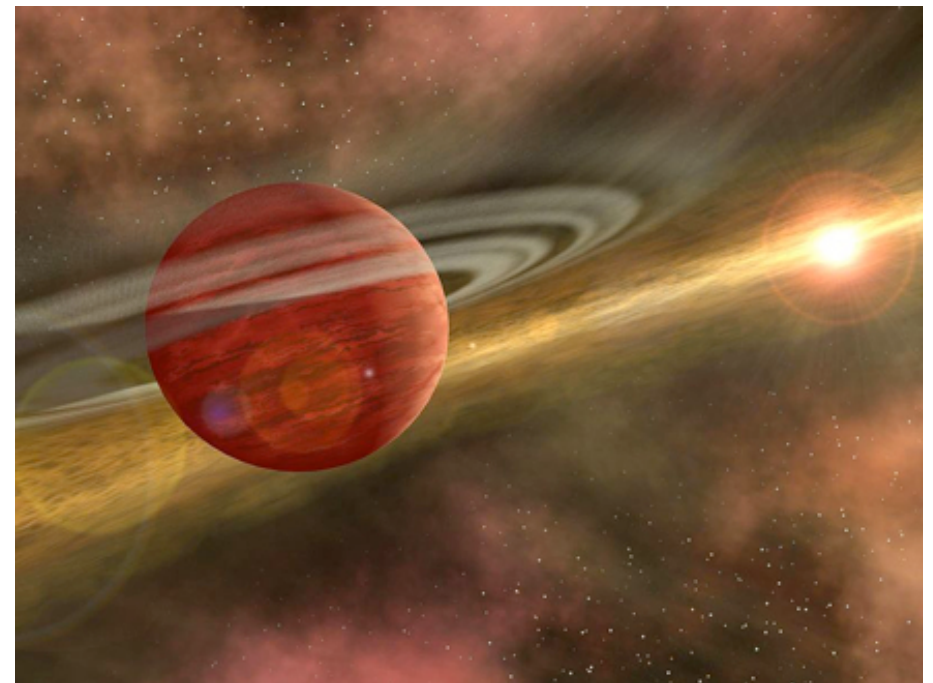
Heavy elements constitute less than 2% of the mass of a solar composition mixture.

The giant planets are enhanced in heavy elements by:

- 5x Jupiter
- 15x Saturn
- 300x Uranus and Neptune

Two main methods are proposed for giant planet formation:

- disk-instability
- core-accretion



Disk-instability hypothesis

Gravitational instabilities in the protoplanetary disk form clumps of material, the protoplanets, with the solid parts settling in their cores.

“Top-down planet formation”

Main advantages of the disk-instability hypothesis:

- It explains the similarities between stars, brown dwarfs, and gas giants
- It is fast enough to explain the gas envelopes of the giant planets (gas giants form within \sim several hundred years)

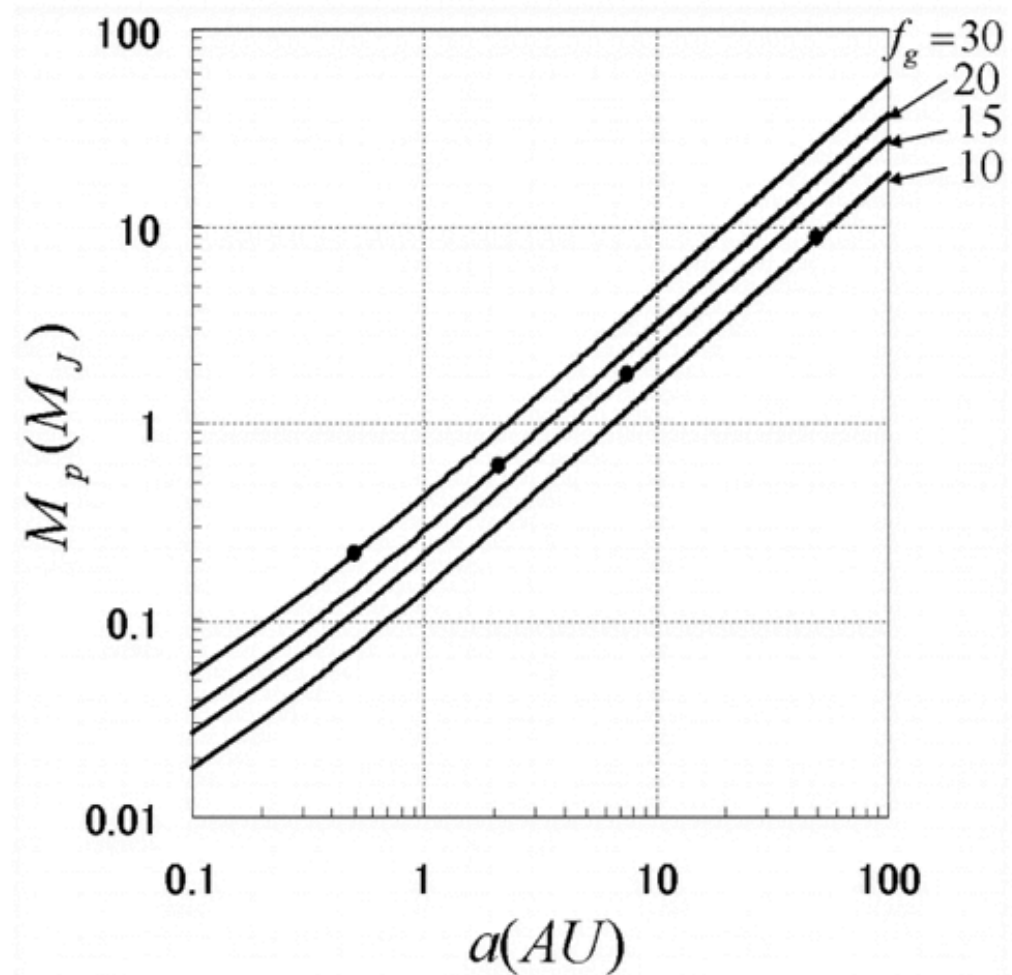
Main disadvantages of the disk-instability hypothesis:

- No explanation for the heavy element enrichment of the gas giants
- No explanation for the observed increased metallicity of host stars
- The instabilities require a very heavy protoplanetary disk ($\sim 1 M_{\text{Sun}}$)

Disk-instability hypothesis

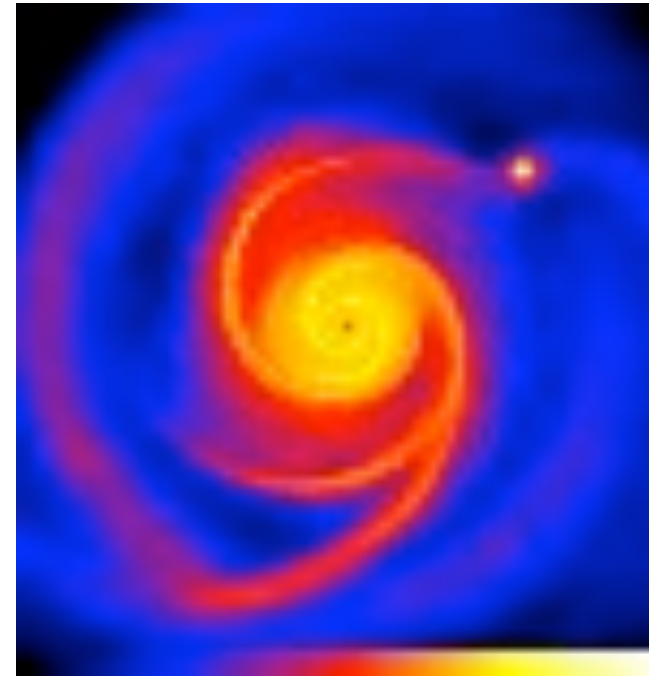
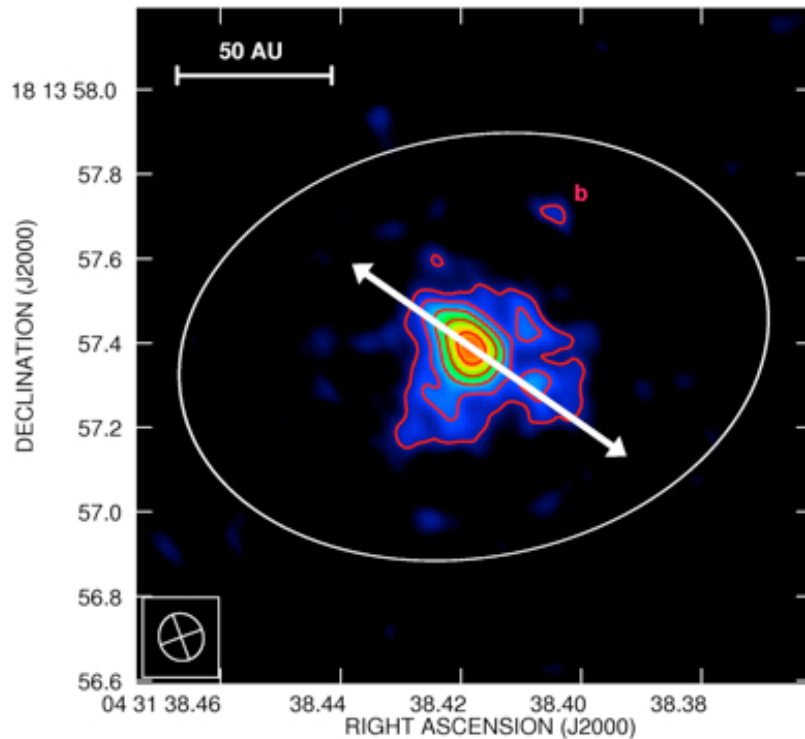
The disk-instability hypothesis predicts a minimum mass for the gas giants that form, depending on the disks surface density and the semi-major axis:

The minimum mass of planets that form due to disk-instability. Factor f_g is a scaling factor with respect to the minimum mass solar nebula (MMSN) model. The dots indicate where the planets should form [Matsuo et al. 2007].



Disk-instability hypothesis

Tentative (radio) detection of a protoplanet around the young ($<10^5$ yr) star HL Tau that could have formed out of disk-instabilities:

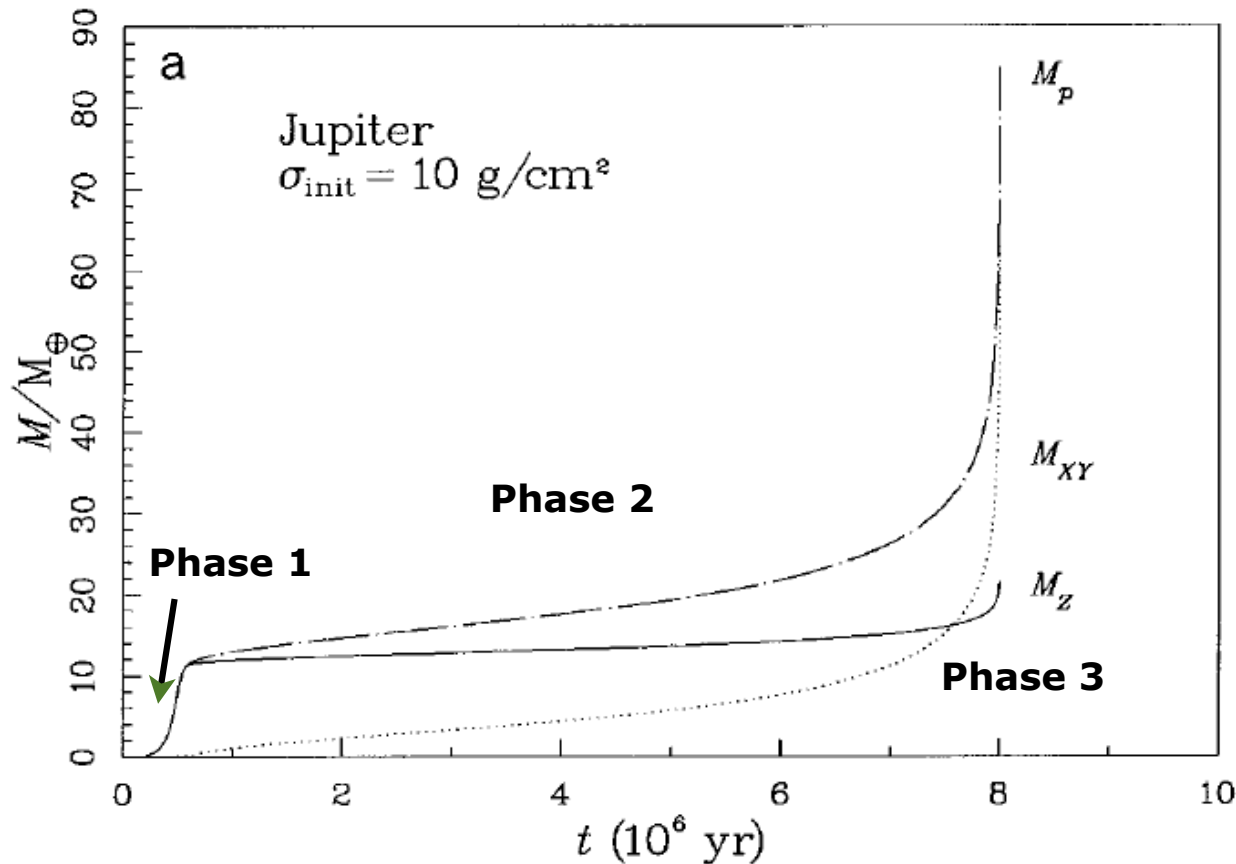


A numerical simulation of the clump's formation

Another star, XZ Tau, may have passed HL Tau about 16,000 yr ago, triggering the instabilities (Greaves et al. 2008)

Core-accretion hypothesis

The core of a giant planet forms by accretion of planetesimals, and its huge atmosphere by run-away accretion of stellar nebulae gas:



Phase 1 ($\sim 5 \cdot 10^5$ yr):

The planet accumulates solids by rapid run-away accretion, until it has severely depleted its feeding zone

Phase 2 ($\sim 7 \cdot 10^6$ yr):

The planet accumulates solids and volatiles at similar rates

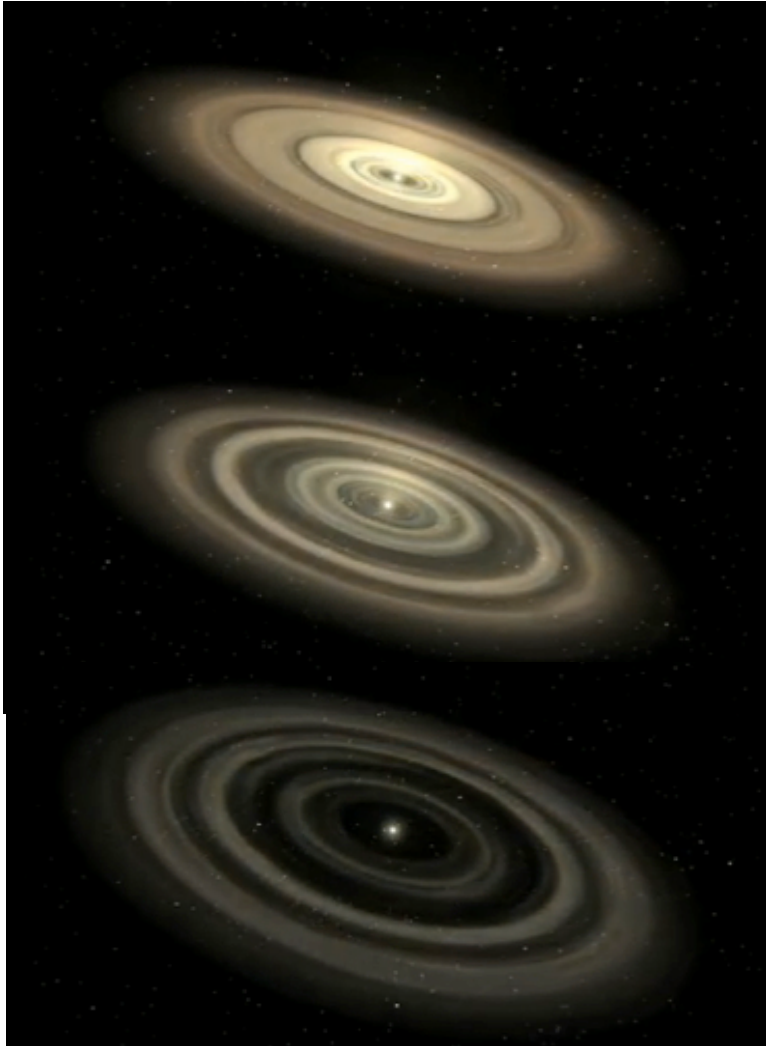
Phase 3:

The planet accumulates volatiles by rapid run-away accretion

From Pollack et al. (1996): M_p is the total planetary mass, M_Z the mass of the core, and M_{XY} the mass of the atmosphere. The density of the protoplanetary disk is 10 g/cm^3 .

Core-accretion hypothesis

While planets accrete gas, they open up gaps, not only because they attract the gas, but also because of exchange of angular momentum:



What is left is what we know as the zodiacal dust between the planets, and the Kuiper belt at the outer edge of the planetary system

Credit: NASA/JPL

Core-accretion hypothesis

At the end of the run-away gas accumulation phase, the planet cools down, and the planet starts to contract. This contraction releases gravitational energy, which is emitted as radiation:



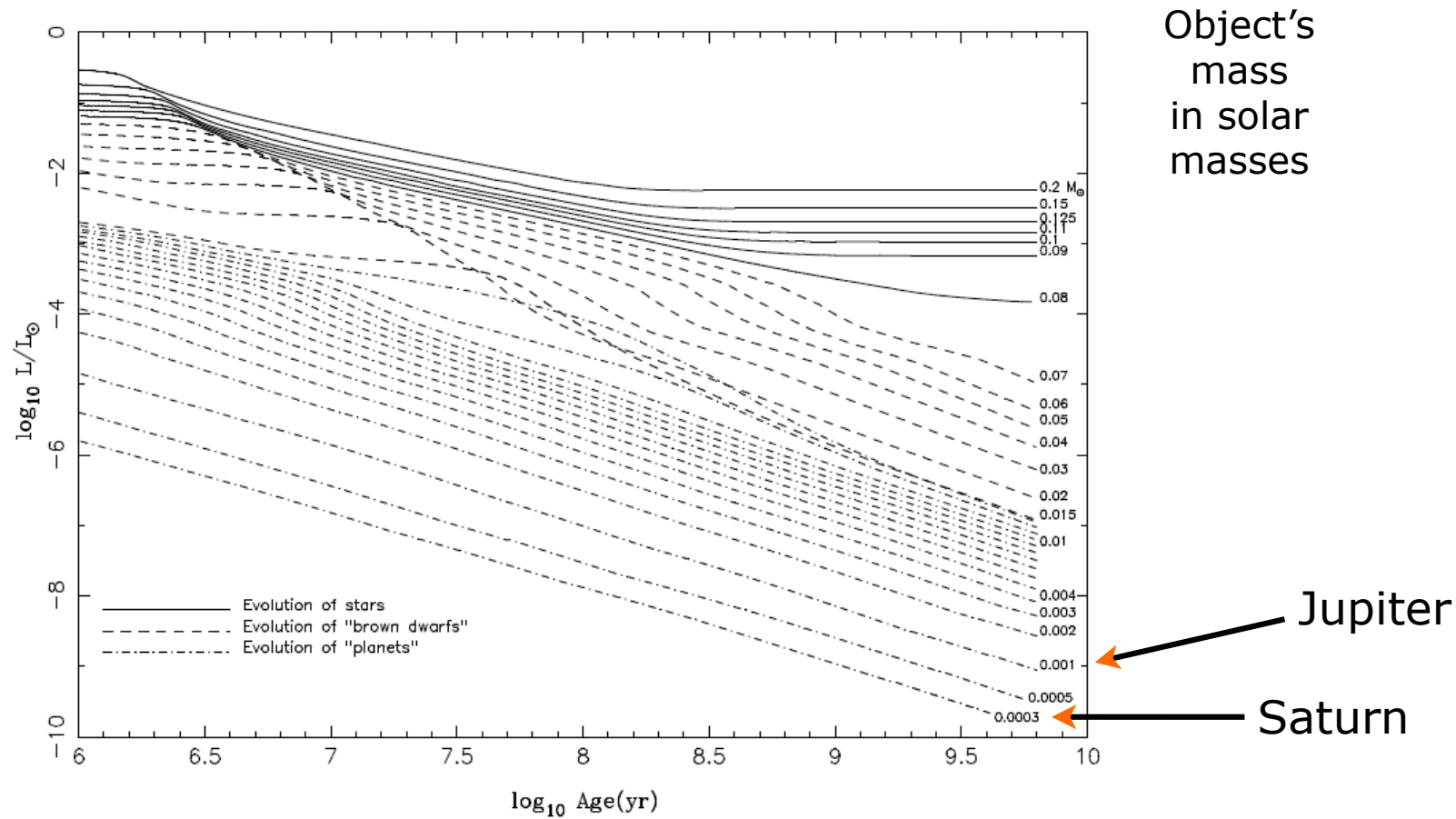
The ratio of the gravitational energy E_G to the planet's luminosity L gives the initial 'Kelvin-Helmholtz' time-scale t_{KH} of the contraction:

$$t_{KH} = \frac{E_G}{L} \sim \frac{GM^2}{RL}$$

G : the gravitational constant, M : the planet's mass, R : its initial radius.

Core-accretion hypothesis

After $\sim 10^4$ years for Jupiter and Saturn, and $\sim 10^5$ years for Uranus and Neptune, the contraction slowed down, and the temperature and luminosity decreased:



From Burrows et al. (1997): The evolution of the luminosity of stars, brown dwarfs and planets measured in Solar luminosities versus time after formation.

Core-accretion hypothesis

“Bottom-up planet formation”

Advantages of the core-accretion hypothesis:

- Seems the best mechanism for terrestrial planet formation
- Leads to large core masses
- Can explain non-solar compositions (gas giants and ice giants)
- Can explain the observed increased metallicity of host stars
- Would lead to circular orbits due to random accretion of planetesimals

Disadvantages of the core-accretion hypothesis:

- Might require unrealistically large core masses
- The total time required for forming the planets might be longer than the lifetime of the disk
- Gap formation in the disk might cut-off the atmospheric accumulation

Core-accretion versus disk-instability

Which is the 'true' planet formation method?

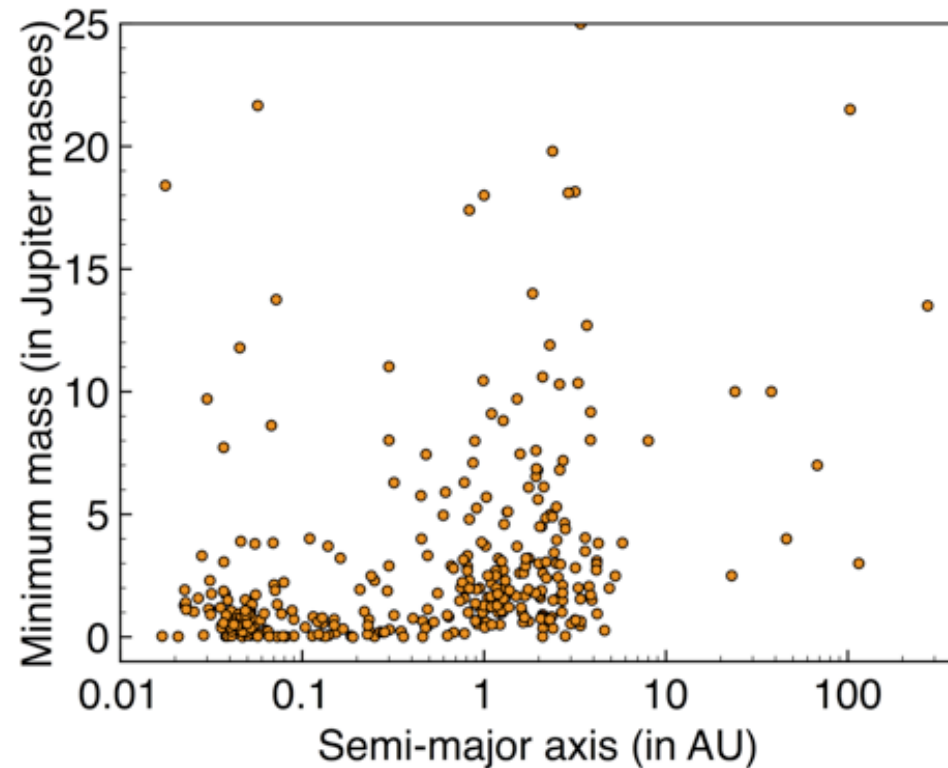
Planets might be formed through both methods.

It appears that about 90% of the planets is explained by core-accretion, and 10% by disk-instability ...



Planet migration

Many exoplanets are found to orbit their star in orbits much tighter than common in the Solar System*:



Planets and planetesimals are thought to form in the outer regions of a circumstellar disk, and then to migrate inward (sometimes outward) because of gravitational interaction with the gas and dust in the disk.

*This can very well be an observational bias! In time, more planetary systems like our own can be discovered

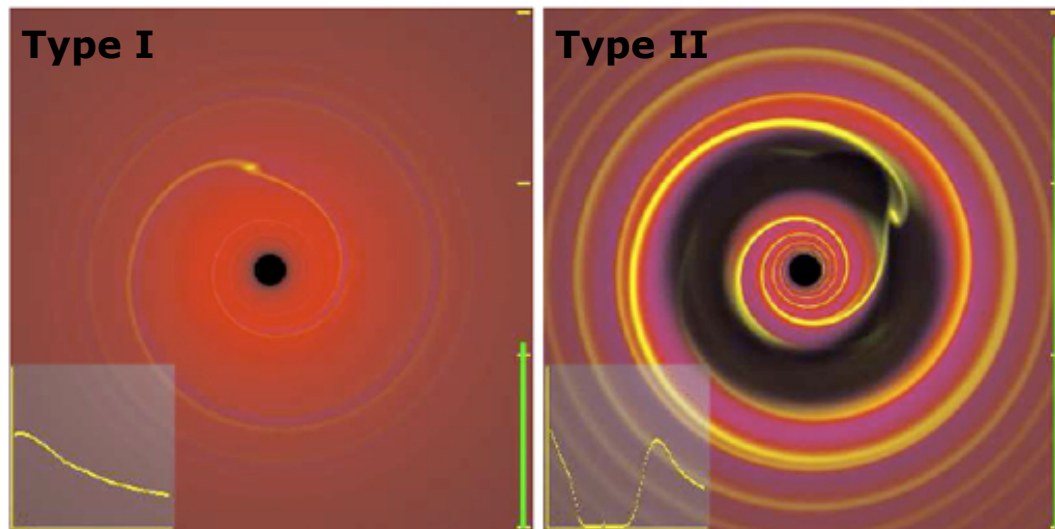
Planet migration

Orbital migration of a solid body involves a change of angular momentum to either gas or other solid bodies in the system, most likely due to gravitational interactions.*

Besides gravitational scattering, we distinguish two types of migration:

Type I: the planetary mass is smaller than the disk

Type II: the planetary mass is large enough to open up a gap

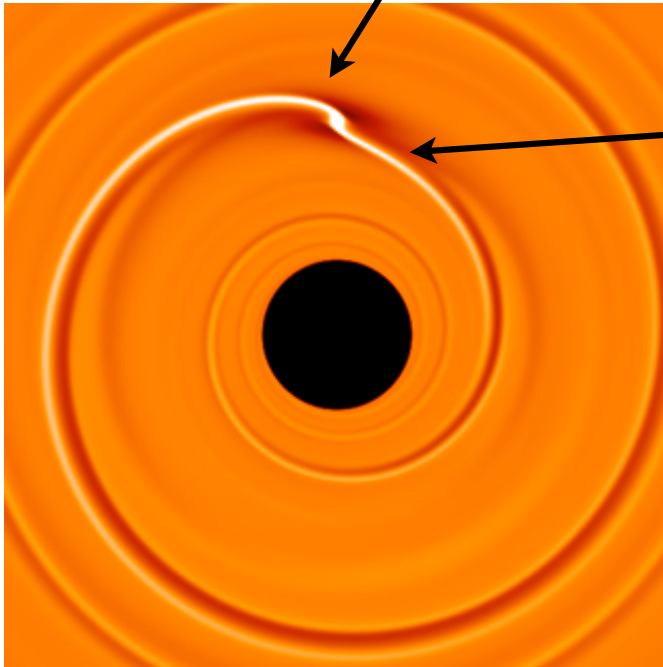


Armitage & Rice (2005)

* Aerodynamic drag, which is important for the orbital evolution of meter-sized planetesimals, is negligible for planetary masses

Type-I migration

The outer disk region orbits slower than the planet, and the gravitational attraction of the spiral waves tries to slow down the planet, and thus move it inwards.



The inner disk region orbits faster than the planet, and the gravitational attraction of the spiral waves tries to speed up the planet, and thus move it outwards.

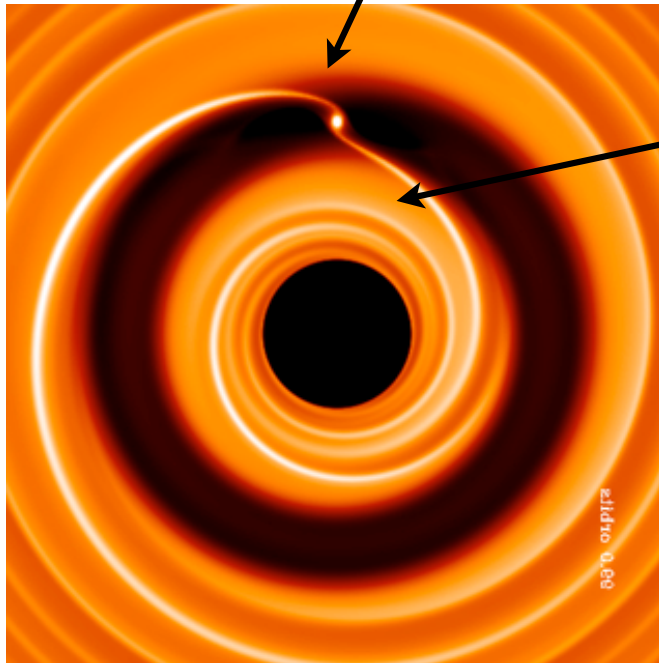
F. Masset (www.maths.qmul.ac.uk/~masset/moviesmpegs.html)

The inward force is generally larger than the outward force ...
the planet thus moves towards its star ... fast!

The time scale for a planet with mass M_p is $\sim 10^5 (M_p/M_\oplus)^{-1}$ yr

Type-II migration

The outer disk region tries to accrete onto the star: if the disk is heavy enough it will 'push' the planet towards the star



The inner disk region accretes onto the star and will deplete if the planet is not spiralling towards the star

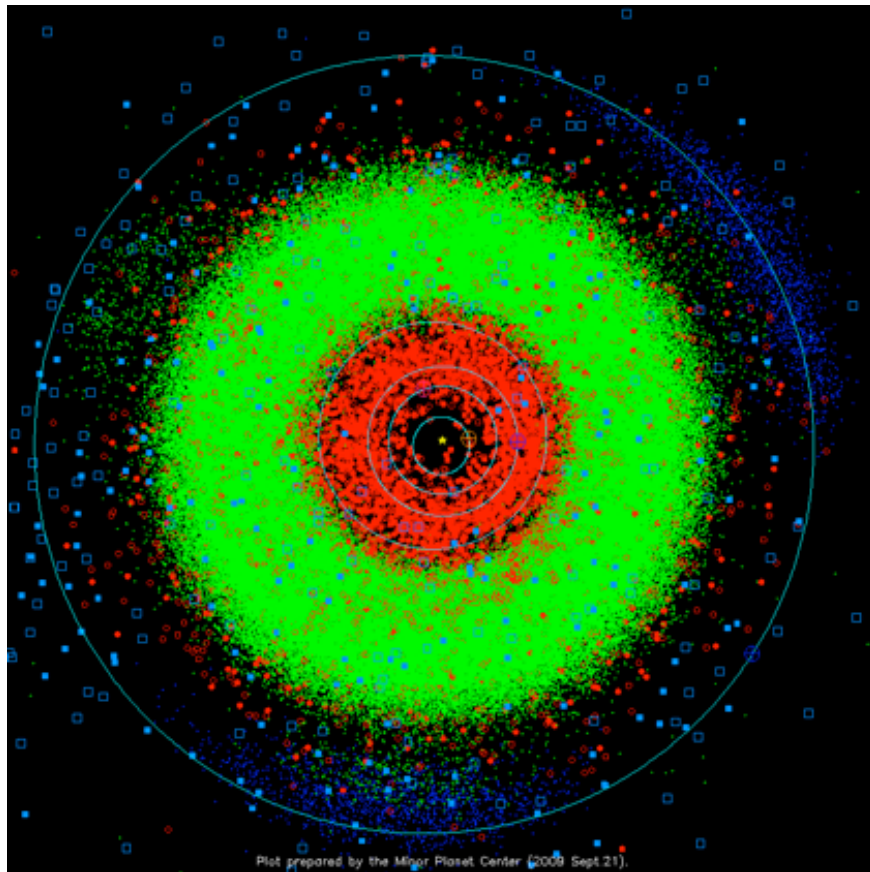
F. Masset (www.maths.qmul.ac.uk/~masset/moviesmpegs.html)

The time scale for type II migration is much longer than for type I ...

Question: when does this type of migration stop in our system?

Origin and evolution of asteroids

Asteroids or minor planets can be viewed as remnant planetesimals although some have been melted and there has been substantial collisional evolution among bodies in the main asteroid belt.



Current asteroid locations projected onto the ecliptic plane. The orbits of Jupiter and the inner planets are shown. The blue squares are comets. Credit: IAU Minor Planet Center

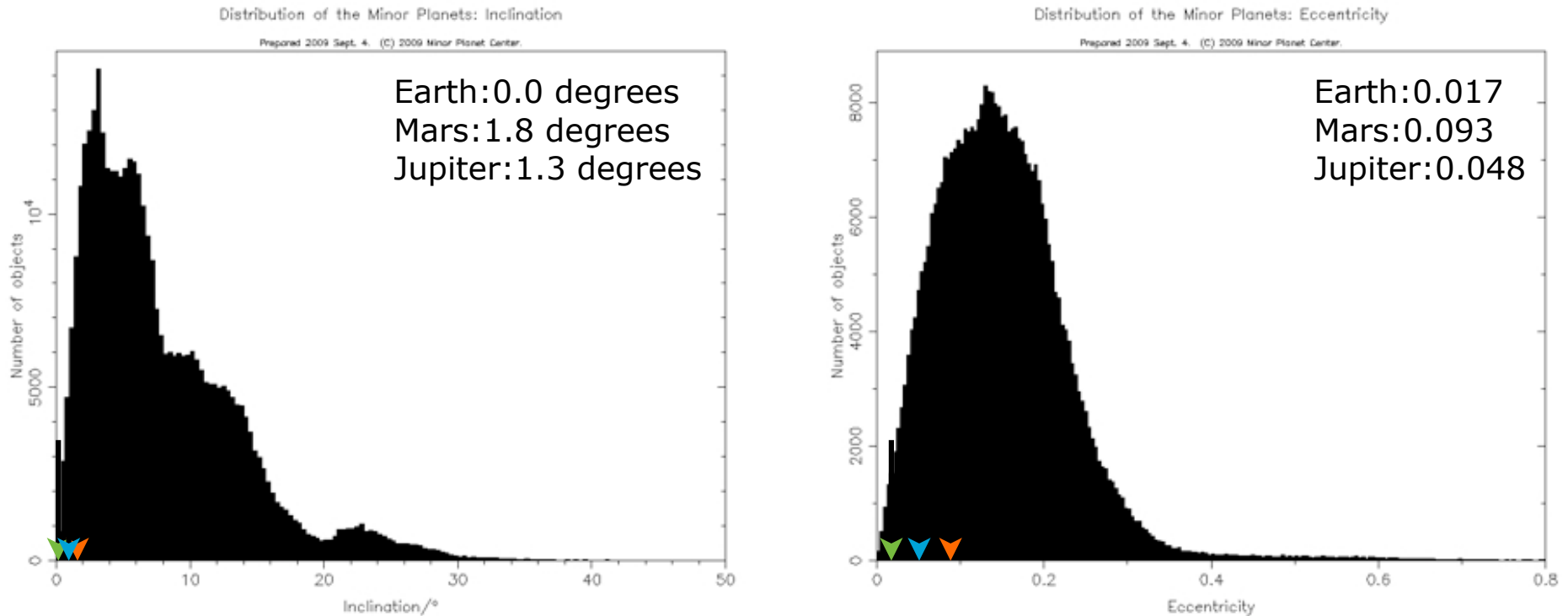
- There are more than 10,000 known asteroids
- They show a large diversity in composition
- The largest asteroid, Ceres, has radius $R \sim 470 \text{ km}^*$
- Number N of asteroids varies with R as $N_{R+dR} \sim R^{-3.5}$
- Total mass in asteroid belt $\sim 5 \cdot 10^{-4} M_{\oplus}$
- Too little mass to form a full-sized planet
- Less mass than expected at these orbital distances



* A rocky body with a typical density of 3.5 g/cm^3 is approximately spherical if $R > 350 \text{ km}$

Origin and evolution of asteroids

Gravitational disturbances of Jupiter are most likely responsible for the mass depletion in the asteroid belt. This would also explain the observed orbital properties of asteroids (large inclination angles and high eccentricities):



The distribution of orbital inclination angles (left) and eccentricities (right) of known asteroids. Also indicated: the inclination angles and eccentricities of the Earth, Mars, and Jupiter. Credit: IAU Minor Planet Center

Origin and evolution of comets

Comets consist primarily of water-ice and dust, but more volatile species (S_2 , N_2^+ , CO) have also been detected. Most comets must have formed at the outer edges of the protoplanetary disk, > 20 AU, where $T < 20 - 30$ K.

The main comet reservoirs in the Solar System:

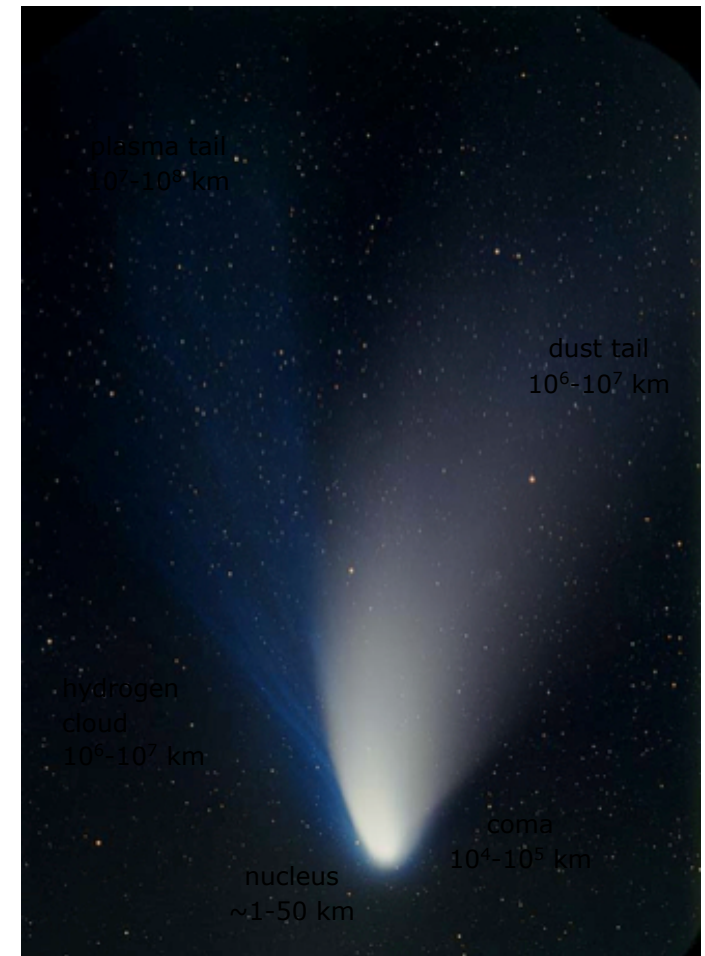
The Kuiper Belt*

- 40-50 AU
- Short Period (SP) (< 200 yr) comets
- $10^9 - 10^{10}$ objects (total mass $< 1 M_{\oplus}$)

The Oort Cloud

- $> 10^4$ AU
- Long Period (LP) (> 200 yr) comets
- $\sim 10^{12}$ objects

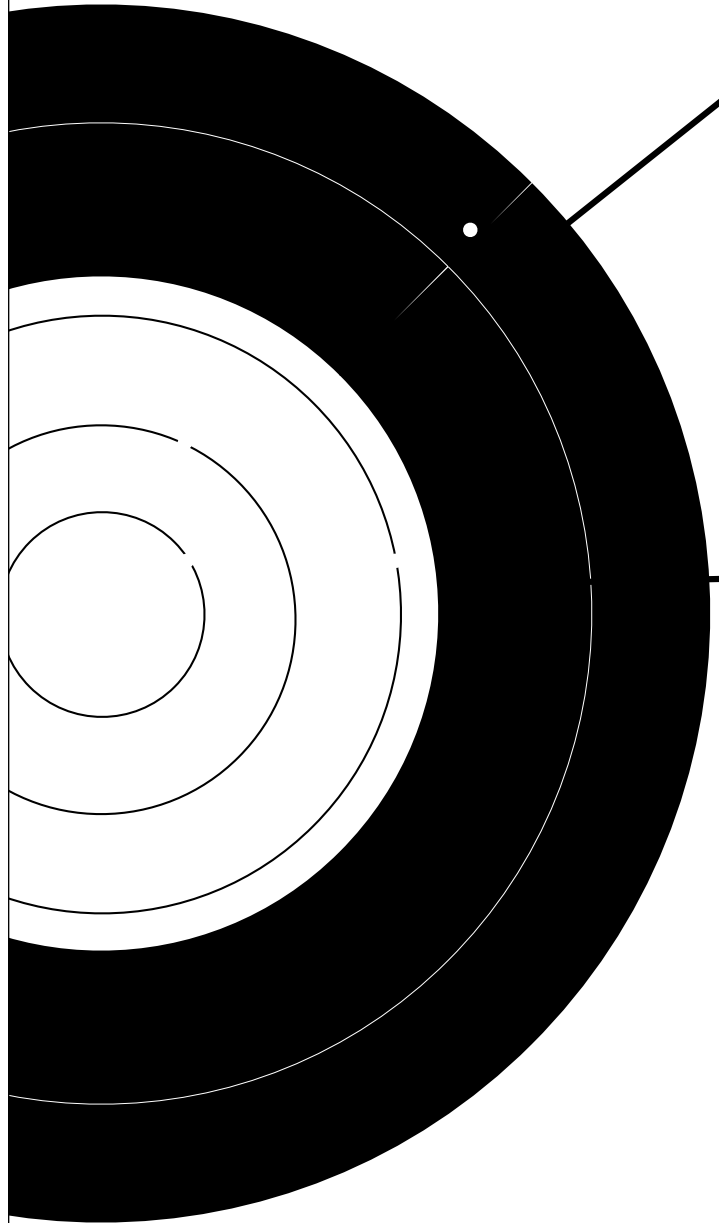
* Also called the Edgeworth-Kuiper Belt



Comet Hale-Bopp (1997)

Credit: Loke Kun Tan/starryscapes.com

Origin and evolution of comets



origin of Kuiper Belt comets

- current gravitational perturbations by the outer planets and local collisions

origin of Oort cloud comets

- gravitationally perturbed inwards by Uranus & Neptune, then outwards mainly by Jupiter
- current gravitational perturbations by passing stars and galactic tides

Planetary satellites

Regular satellites:

- prograde, low eccentricity orbits near the planet's equatorial plane
- well within the planet's Hill sphere*

Formed in-situ, e.g. from circumplanetary disk

Irregular satellites:

- prograde or retrograde, high eccentricity, high inclination orbits
- outside the regular satellites system
- usually quite small (non-spherical)

Formed elsewhere, were captured by the planet

*The Hill sphere is the area (with radius R_H) in which the gravity of a secondary body (m_2) exceeds that of a primary body (m_1) that is located at distance a :

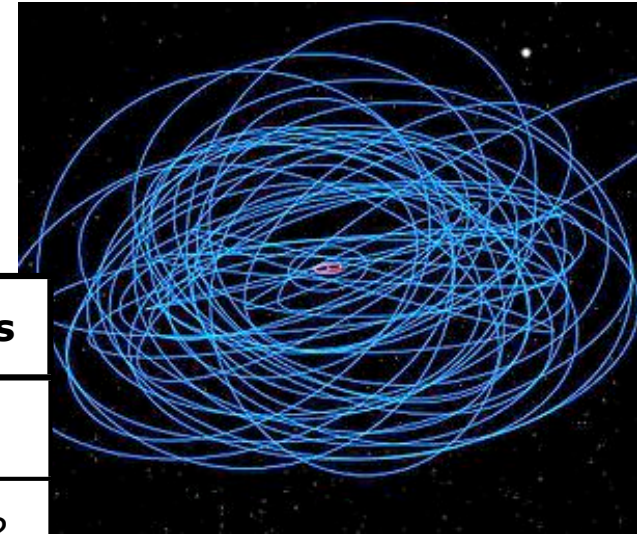
$$R_H = \left[\frac{m_2}{3(m_1 + m_2)} \right]^{1/3} a$$



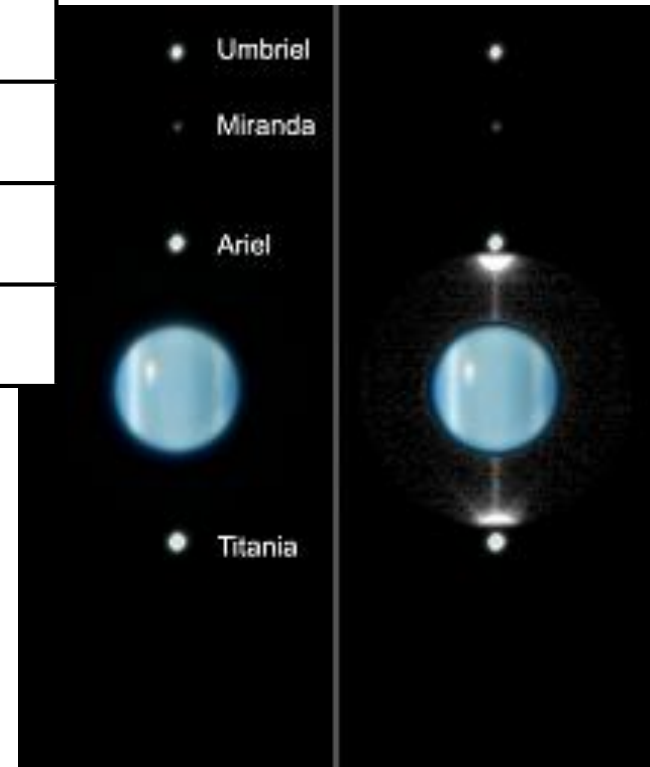
Planetary satellites

Satellites of the Solar System planets:

Planet	# regular satellites	# irregular satellites
Earth	1 (Moon)	-
Mars	2 (Phobos & Deimos)?	2 (Phobos & Deimos)?
Jupiter	8 (the Amalthean and Galilean moons)	55
Saturn	23 (largest: Titan)	38
Uranus	18 (largest: Titania)	9
Neptune	6	7 (largest: Triton)

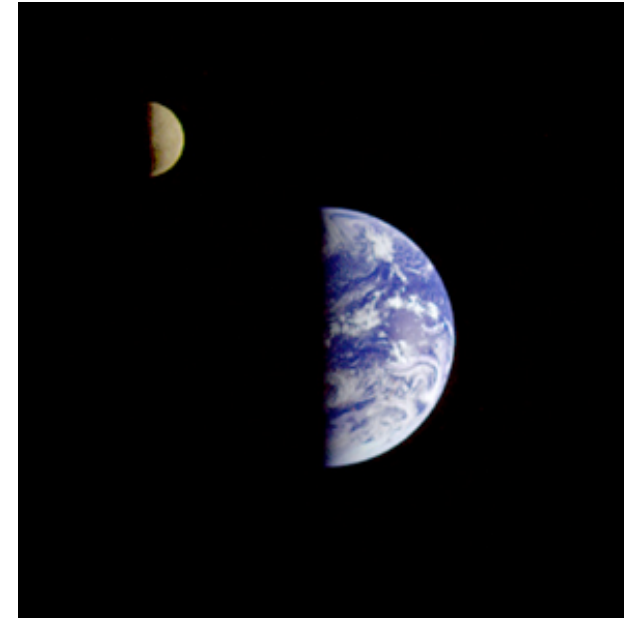


The moons of Saturn



Uranus' regular moons. Credit: ESO/VLT

The origin of the Earth's moon



Credit: NASA/Galileo mission

The origin of our moon has long been debated. Some early ideas:

- The Earth and moon formed at the same time and location (as a double-planet system)

The average density of Earth is 5.5 g/cm^3 , that of the moon 3.3 g/cm^3 (Earth has a large iron core, the moon doesn't): these should be similar

- The moon formed somewhere else (where there was less iron) and was captured by the Earth

The moon has the same oxygen isotope ratio as the Earth, while Mars and meteorites have different ratios. The moon formed thus near the Earth

- The early Earth rotated so rapidly that the moon spun off from the Earth

This would give the moon its composition, but leaves too much angular momentum in the Earth-moon system

The origin of the Earth's moon

The Earth's moon likely formed about 4 - 4.5 billion years ago, through a giant impact of the half-built Earth with a body about half its size:

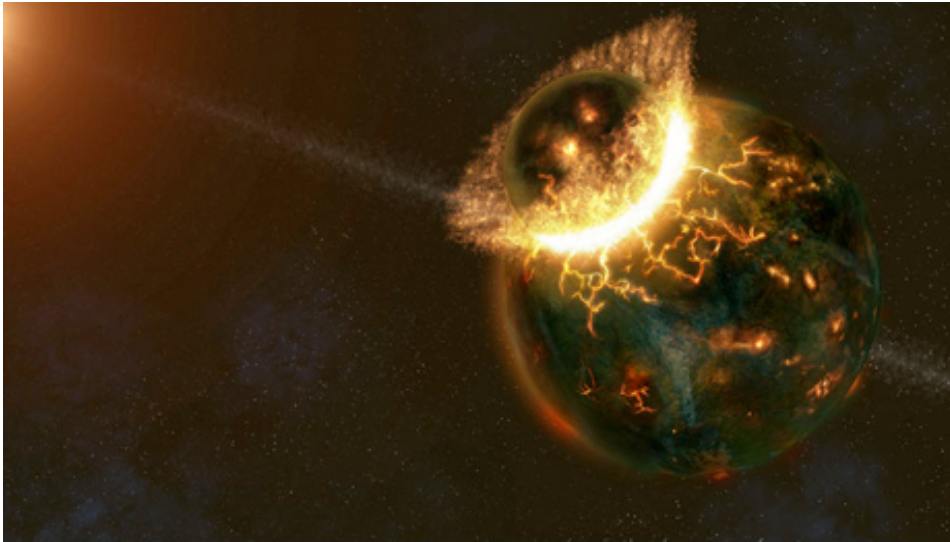


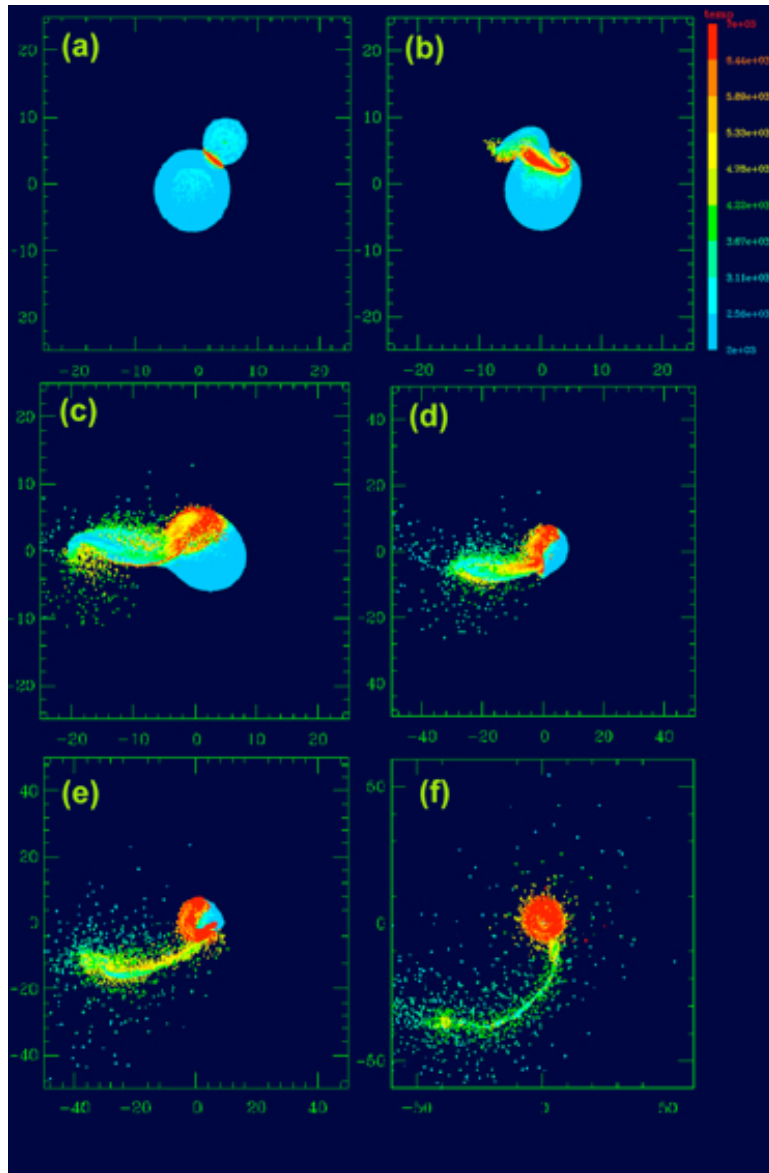
Image credit: National Geographic

After the impact, the Earth had about 2/3 of its current size and the moon formed from the ring of debris. Both the Earth and the moon continued to accumulate impact material after this event.

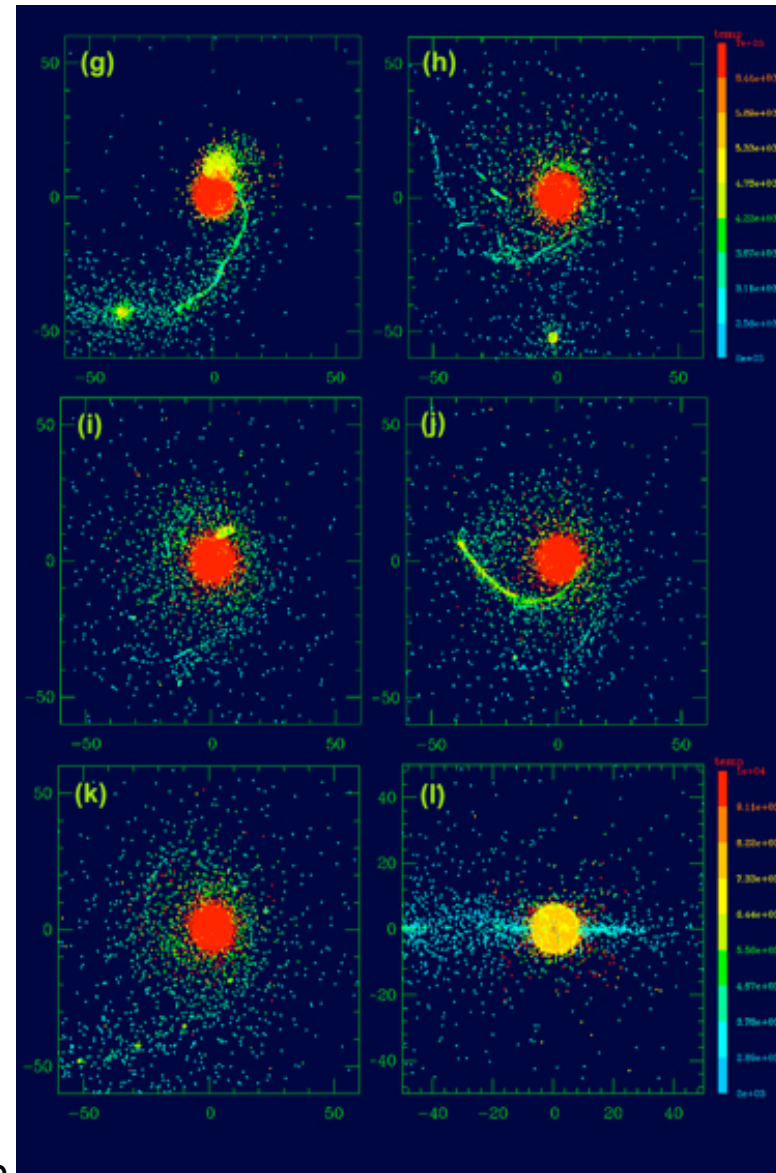
This chance event would explain why Venus and Mars lack large moons! ...

The origin of the Earth's moon

A numerical simulation of the possible events during the giant impact:



Credit: R. Canup, Southwestern Institute, Boulder



An example of an irregular satellite:

Phobos, one of the two Martian moons. The porosity and composition of Phobos reveal that it might have originated from an impact on Mars, its orbit is also quite circular and not very inclined:



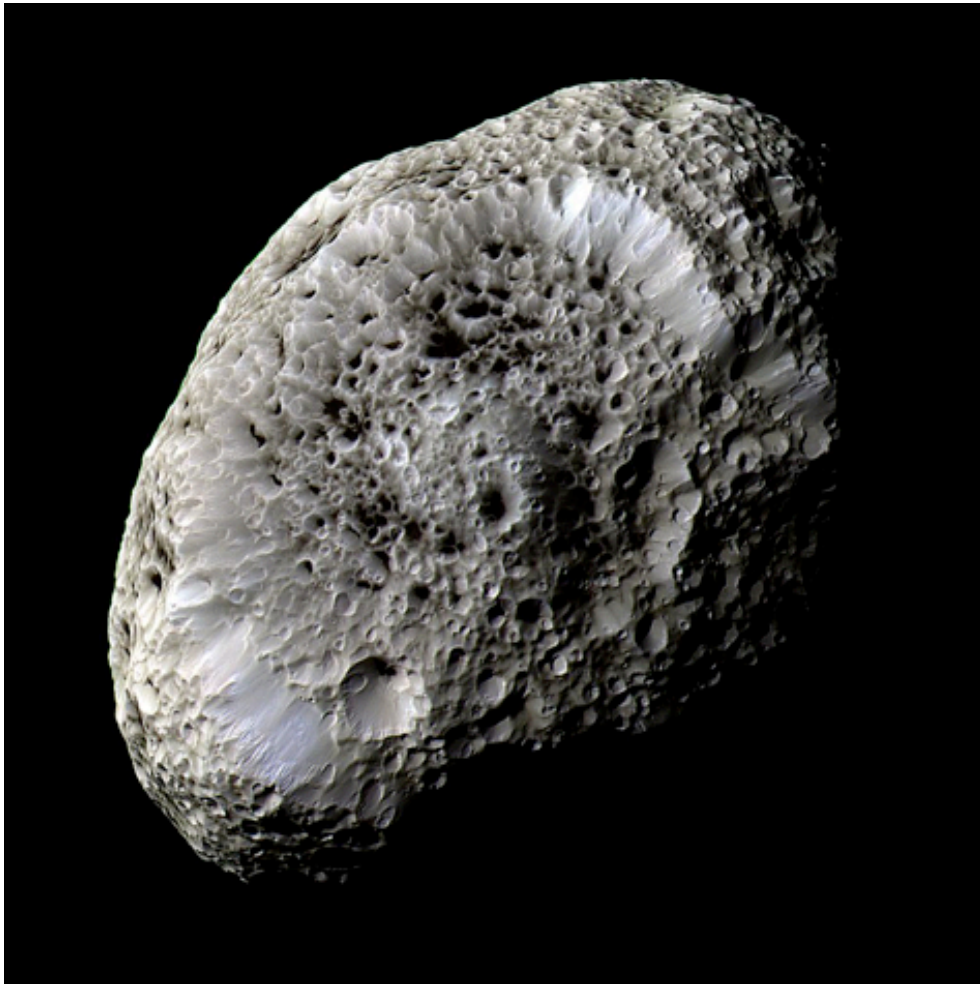
Phobos. Dimensions ~ 30x20x20 km. The large crater is called Stickney and its existence indicates Phobos' porosity. Credit: NASA/JPL



Deimos and Phobos were long thought to be captured asteroids.

An example of an irregular satellite:

A false-colour image of Saturn's moon Hyperion obtained during a fly-by of Cassini at a distance of $\sim 60,000$ km:



Credit: NASA/JPL/ Space Science Institute (Sept. 26, 2005)

Year of discovery	1848
Mass	$1.8 \cdot 10^{19}$ kg ($3 \cdot 10^{-6} M_{\oplus}$)
Dimensions	360x280x225 km
Mean density	1.4 g/cm ³
Mean orbital radius	1,481,000 km
Orbital period	21.3 days
Orbital direction	prograde
Orbital eccentricity	0.1042
Orbital inclination angle	0.43°
Rotational period	chaotic
Geometric albedo	0.3
Visual magnitude	14.19

Another irregular satellite:

A false-colour image of Neptune's largest moon Triton:



Credit: NASA/Voyager 2 (1989)

Year of discovery	1846
Mass	$2.1 \cdot 10^{22}$ kg (0.004 M_{\oplus})
Diameter	2700 km
Mean density	2.06 g/cm ³
Mean orbital radius	354,760 km
Orbital period	5.877 days
Orbital direction	retrograde
Orbital eccentricity	0.000016
Orbital inclination angle	157°
Rotational period	synchronous (5 d, 21 h)
Geometric albedo	0.76
Visual magnitude	13.47