

Lecture 2: Habitable Zones and Conditions in the Early Solar System

Aims of Lecture

- (1) To explain the concepts of galactic and circumstellar habitable zones
- (2) To outline the early evolution of the Sun and the ‘faint young sun paradox’
- (3) To discuss the period of heavy bombardment and the possible ‘impact frustration’ of life
- (4) To discuss the role of comets in delivering volatiles and organic compounds into the inner Solar System
- (5) To discuss the possible role of giant planets in shielding the inner regions of planetary systems from too many comets!

Contents of Lecture

1. The concept of habitable zones: Page 2
2. The ‘faint young sun’ paradox: Page 7
3. The period of heavy bombardment: Page 9
4. The role of comets and giant planets: Page 11

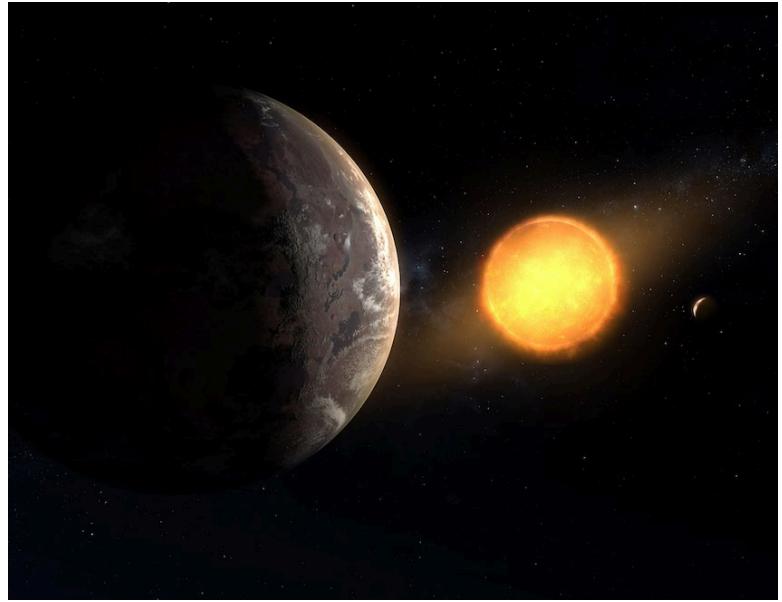


Fig. 1. An artist’s drawing of a planet orbiting a young star; note light scattering in the remnants of the original protoplanetary disk. The habitability of such a planet will depend on multiple factors: The star’s luminosity, the distance of planet from the star, the planet’s atmospheric composition and albedo, the abundance of asteroids and comets in the system, and possibly the presence of nearby giant planets and the location of the planetary system in the Galaxy (NASA/Daniel Rutter).

1. The concept of habitable zones

The term ‘habitable zone’ in astrobiology usually refers to spatial locations where astronomical conditions may permit life to exist. There are two rather different concepts:

1.1 A galactic habitable zone?

It has been suggested that a planet’s habitability may depend on its location in the Galaxy, giving rise to the concept of a Galactic Habitable Zone (GHZ; Fig. 2). The inner boundary of the GHZ could be governed by several factors related to the higher density of stars, including:

- Increased occurrence of supernovae (SN) and gamma-ray bursts (GRBs), the radiation from which may disrupt the evolution of life in their vicinity.
- Close encounters between stars may perturb planetary orbits (e.g. make orbits very elliptical, or even eject planets) and/or perturb comet clouds leading to increased risk of ‘impact frustration’ of evolution.
- High heavy element abundances might favour giant planets at the expense of terrestrial-type planets.

The outer boundary of the GHZ would be governed by the relative rarity of heavy elements required to build planets owing to reduced star formation rates. This will evolve with time, however, so the outer boundary of the GHZ would be expected to migrate outwards as the Galaxy ages.

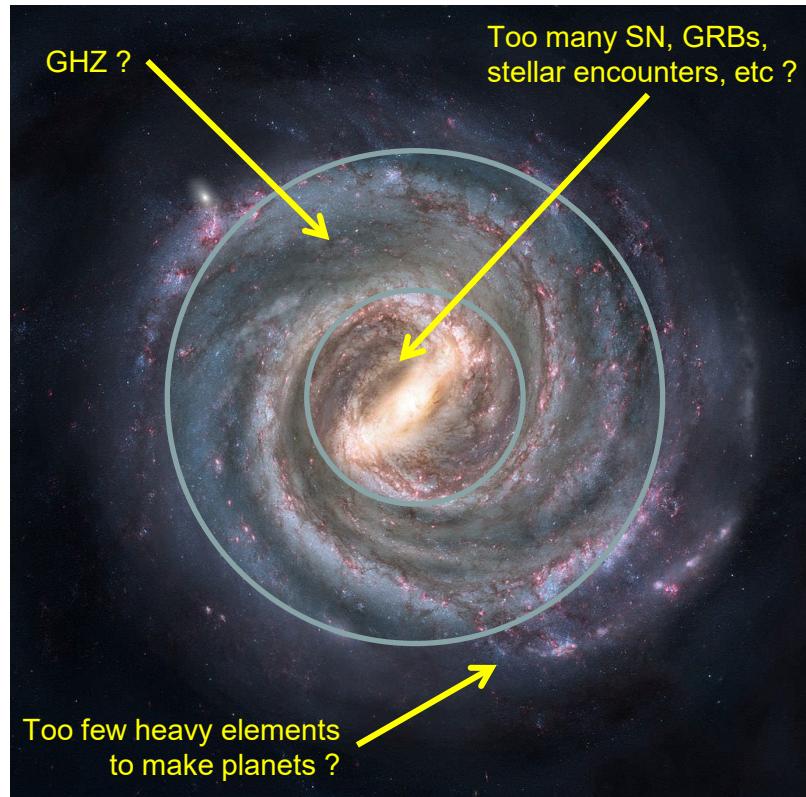


Fig. 2. Concept of a Galactic Habitable Zone superimposed on an artist’s impression of the Milky Way Galaxy. See text for discussion (background image: NASA).

The GHZ concept is open to criticism for several reasons:

- The definition of ‘habitability’ is vague. For the circumstellar habitable zone (discussed below) ‘habitability’ is *defined* (for better or worse!) by the stability of liquid water on a planet’s surface. For the GHZ there is no corresponding definition, just a vague sense that too many perturbing external influences (SN explosions, etc) would diminish habitability. But this must depend on the *complexity* of the life being considered – for example, microbes living in ice-covered oceans (like Europa’s probable ocean, see Lecture 9) will be less susceptible to nearby SN explosions than multi-celled life exposed on a planet’s surface.
- The dependence of planet formation on heavy element abundance is not known reliably – recent work indicates that rocky planets can probably form with heavy element abundances just a few percent of the Sun’s value, implying that the outer boundary of the GHZ in disk galaxies like the Milky Way may have encompassed the whole galactic disk even at relatively early times.
- The Galaxy is a chaotic place – star formation, supernova explosions, etc, all fluctuate in time and space and simplistic models that assume a purely radial dependence of these phenomena with distance from the centre do not allow for this complexity. For example, some early-formed planets close to the centre of the Galaxy may have been ‘lucky’ and missed stellar encounters or nearby supernovae, and some regions of the outer Galaxy may have been more heavy element rich than average permitting the early formation of rocky planets.

Even if it exists as envisaged, the GHZ will encompass a significant fraction of the total volume of the Galaxy (see Fig. 2), so is unlikely to impose a strong constraint on the total number of habitable planets.

1.2 The circumstellar habitable zone

The ‘circumstellar habitable zone’ (CHZ, often just HZ for short) is defined to be the region of space around a star where liquid water will be stable on the surface of an orbiting planet. Note that this is a very restricted definition of ‘habitability’. The inner boundary of the HZ is the distance from the star where the planet will be just too warm for liquid water to be stable, and the outer boundary where it is just too cold. However, the HZ depends on other factors which also determine the stability of liquid water on a planet’s surface:

- Brightness (spectral type) of the star
- Atmospheric greenhouse effects (depends on atmospheric composition);
- Planetary albedo (i.e. reflectivity; e.g. as controlled by cloud and ice cover); and
- Atmospheric pressure (governs the temperature range over which water remains liquid).

The inner and outer limits of the HZ move away from the star as the star evolves off the zero-age main sequence (Fig. 3). The region of a planetary system that remains ‘habitable’ over a specified period of a star’s life is known as the ‘continuously habitable zone’ for that period. These calculations assume an ‘Earth-like’ atmosphere, with reserves of CO₂ and H₂O to act as greenhouse gases. Despite the sophistication of the models, many uncertainties remain and Fig. 3 shows a range of possibilities.

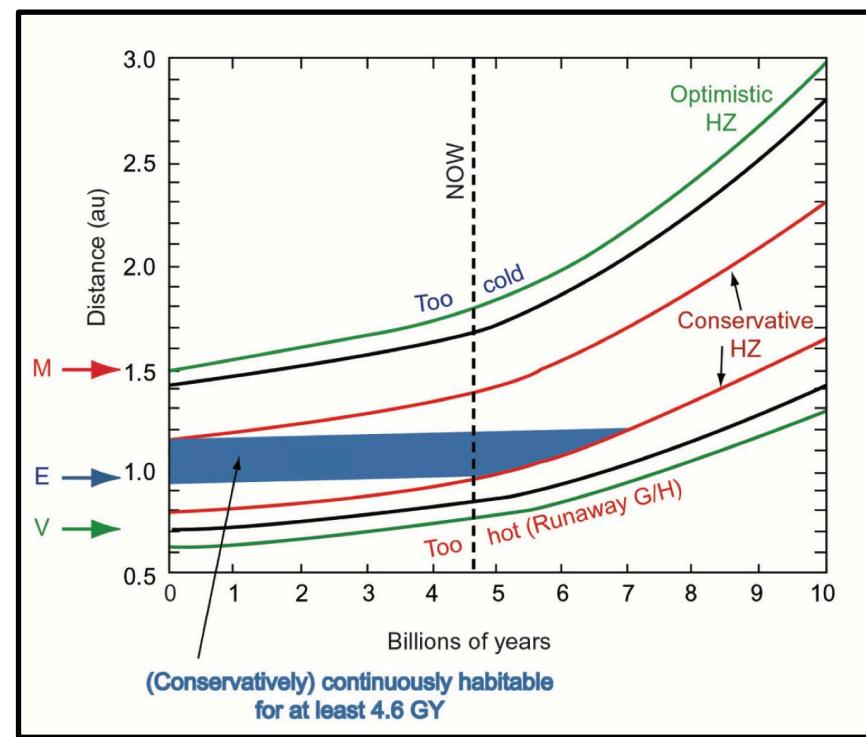


Fig. 3. Evolution of the habitable zone for an Earth-like planet over the main-sequence lifetime of the Sun. A range of models are shown, illustrating ‘conservative’ (red), ‘best guess’ (black) and ‘optimistic’ (green) scenarios. The width of the continuously habitable zone for the last 4.6 billion years is also shown (based on Kasting et al., *Icarus*, vol. 101, p. 108, 1993).

The ‘conservative’ case (red lines in Fig. 3) has the narrowest habitable zone predicted by the models, in which a runaway greenhouse effect occurs early, even for planets far from their stars. According to this model, Venus has always suffered from a severe greenhouse effect, and the Earth is only a few hundred million years away from one!

Conversely, in the ‘optimistic’ case (green lines), planets close to the Sun are able to postpone the onset of a runaway greenhouse effect (e.g. by burying CO₂ in carbonate rocks) whereas planets far from the Sun still have sufficient greenhouse gases in their atmospheres to maintain liquid water on their surfaces (even with a faint young sun). This model would predict that Venus may have been ‘habitable’ for the first two billion years of its history (for which there is no direct evidence), and is also consistent with the evidence for liquid water on the surface of Mars 3.5 to 4 billion years ago (Lecture 8).

The ‘best guess’ model lies between the two. This model implies that the Earth will remain ‘habitable’ for another 1.5 to 2 billion years, after which the increased concentration of water vapour in the atmosphere, resulting from higher temperatures due to the increasing solar luminosity, will trigger a runaway greenhouse effect. However, the reduction in atmospheric CO₂ will stress terrestrial photosynthetic organisms and will result in a reorganisation (likely mass extinction) of the biosphere *prior* to surface temperatures becoming uninhabitable, approximately 0.8 Gyr from present.

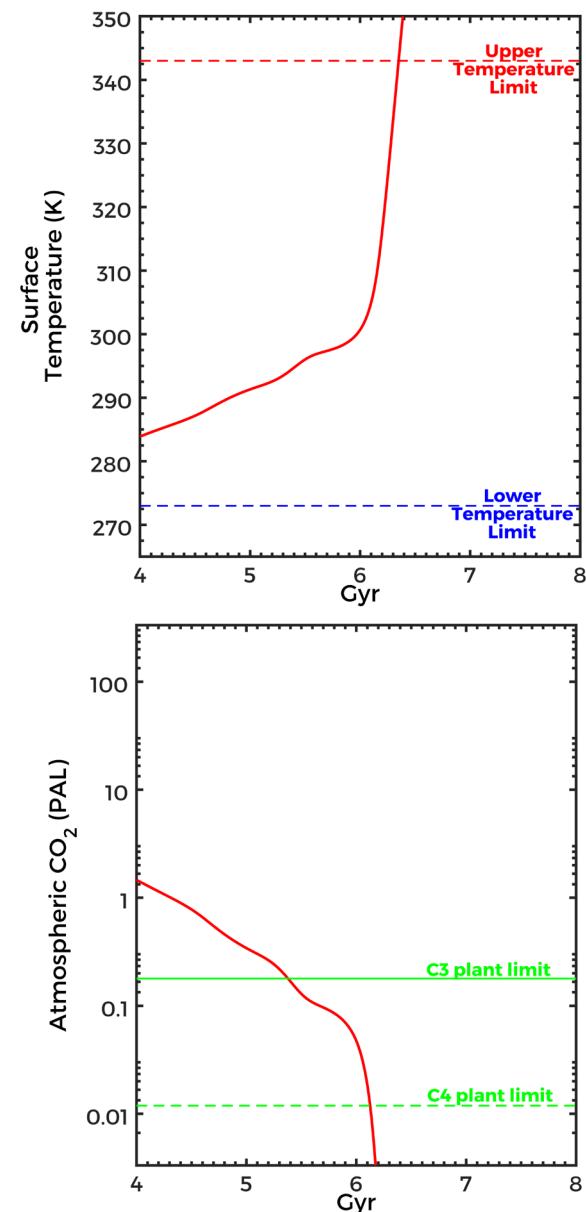


Fig. 4. Possible future climatic (top) and atmospheric CO₂ (bottom) evolution of the Earth. (based on Rushby et al., *Astrobiology*, vol. 18 (5), p.469, 2018). 5

Note that all these models rely on CO₂

and H₂O as greenhouse gases.

Introducing other atmospheric constituents (e.g. CH₄ or NH₃) could extend the outer boundaries of the HZ considerably.

The HZ will be both wider, and further out, for stars that are more massive, and thus more luminous, than the Sun. Conversely, the HZ will be narrower, and lie closer in, for less massive stars. Although the broader HZs of more massive stars may increase the chances of finding planets within them, note that more massive stars have shorter lifetimes which will reduce the opportunities for biological evolution.

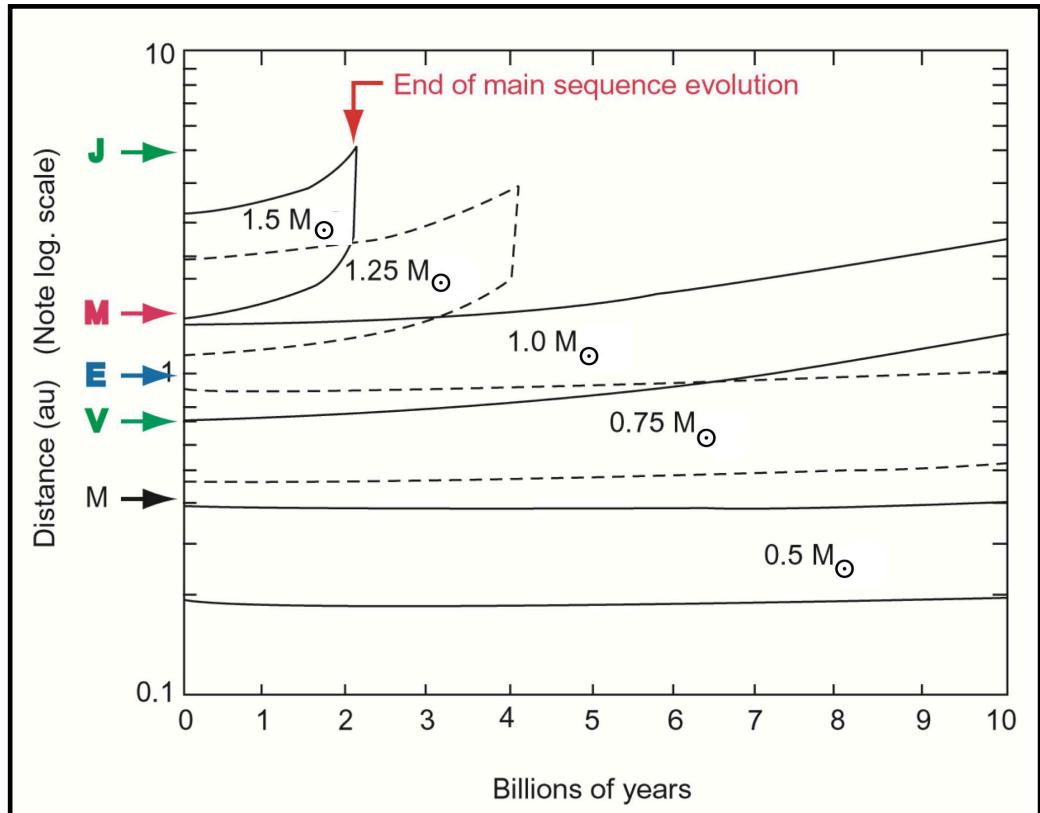


Fig. 5. The locations and widths of the habitable zones for stars of different masses. Those for the more massive stars come to an abrupt end when they leave the main sequence and become red giants.

Also, despite their narrower HZs, studies of exoplanets have actually found that planets within the HZs of very long-lived red dwarf stars are actually quite common (Lecture 10). The locations and widths of habitable zones (assuming ‘best guess’ parameters) for stars of different masses are shown in Fig. 4. The orbital distances of planets in our own solar system are shown for comparison. Note that the distance scale in Fig. 4 is logarithmic: for a recently formed 0.5 solar mass star the HZ extends from 0.2 to 0.4 AU, a width of only 0.2 AU (where 1 ‘astronomical unit’, AU, is the distance of the Earth from the Sun); on the other hand, for a 1.5 solar mass star the HZ extends from 1.2 to 3.2 AU, and is thus ten times wider.

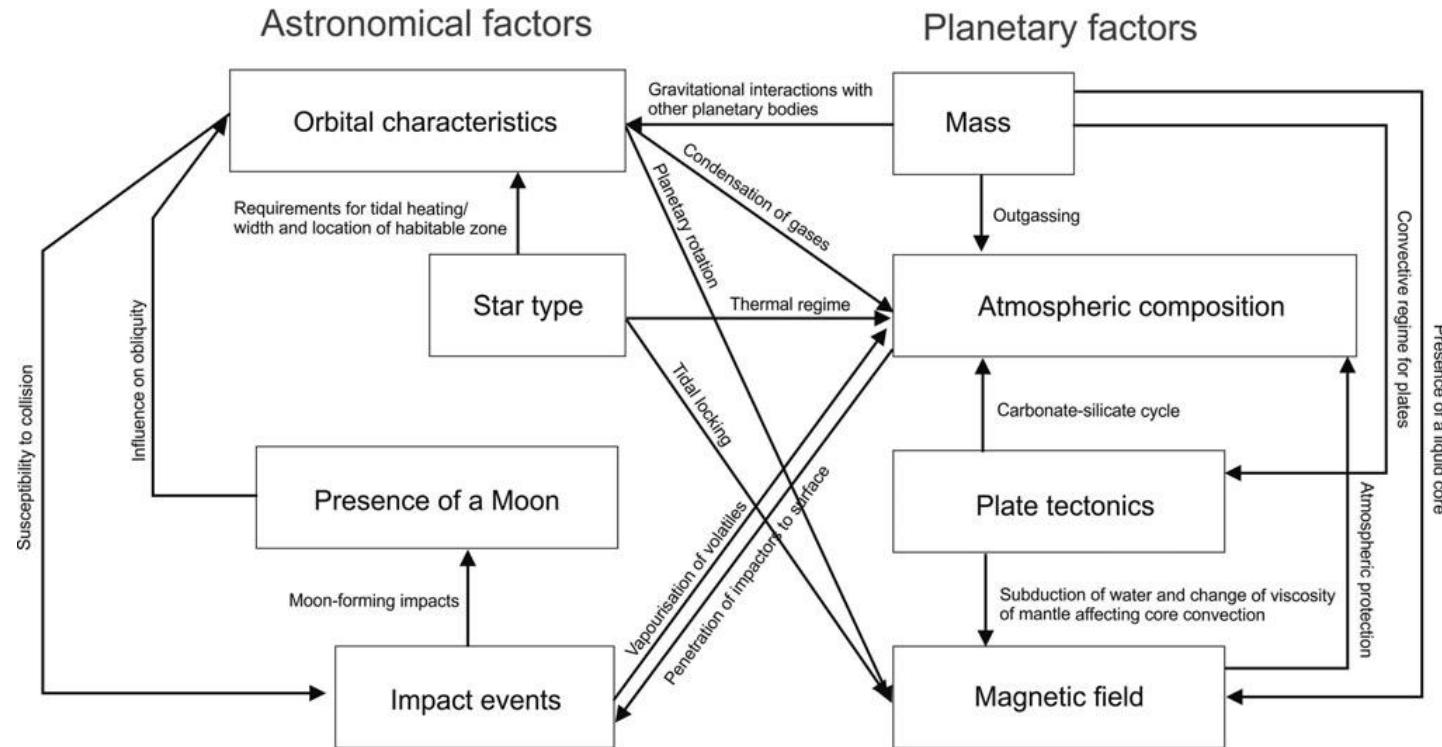


Fig. 6. 'Habitability' is a complex, emergent property that is affected by many interconnected astrophysical and planetary factors (Cockell et al., *Astrobiology*, vol. 16, p. 89, 2016).

Despite the usefulness of the habitable zone concept, it suffers from a major limitation: it only applies to planetary surfaces that are warmed by sunlight. Thus, it does not apply to underground or sea floor environments where life could be maintained by geothermal heat and chemical energy sources. For example, Jupiter's moon Europa lies well outside the habitable zone as defined here, but may nevertheless have an ocean of liquid water capable of supporting life under an icy crust (as discussed in Lecture 9).

'Habitability' is a complex, emergent property that is affected by many interconnected astrophysical and planetary factors; the HZ is valuable in that it is an astronomically observable property of a star that may allow us to prioritise planet characterisation, but should not be taken as ground-truth.

2. The faint young sun paradox

As discussed above (see also Lecture 1), stars brighten as they evolve off the zero-age main sequence in the HR diagram (Fig. 5). The initial brightening is very slow, but the Sun is now $\sim 30\%$ brighter than when it formed 4.6 billion years ago. This means that all the planets would have received less sunlight than they do today, and calculations indicate that the Earth should have been frozen solid. However, the existence of 3.8 Gyr old metamorphosed sedimentary rocks (and 4.3 Gyr old detrital zircon grains, discussed below), indicate that the Earth's surface did possess large quantities of liquid water at these early times. This discrepancy between theory and observation has become known as the 'faint young sun paradox'.

A major determinant of planetary surface temperatures, in addition to the brightness of the Sun and distance from it, is the greenhouse effect of planetary atmospheres. Briefly, atmospheric molecules such as CO_2 , H_2O , and CH_4 transmit visible light from the Sun, which results in heating of the ground, but trap the resulting infra-red radiation within the atmosphere, thus causing it to warm up (Fig. 6). The solution to the faint young sun paradox is therefore to postulate a more powerful greenhouse effect in the past.

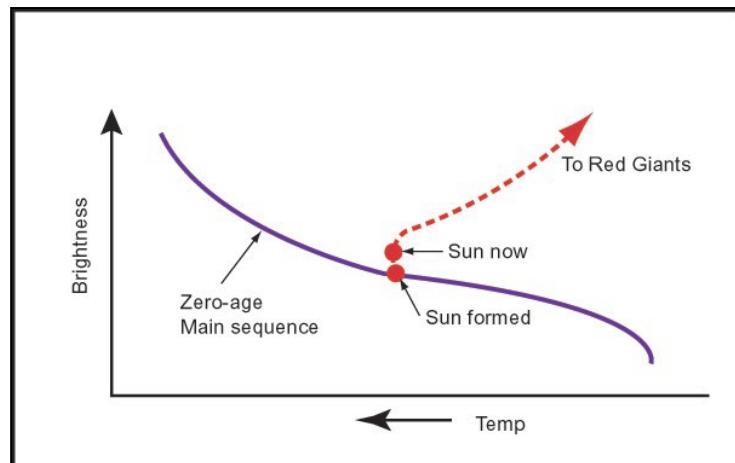


Fig. 7. Schematic Hertzsprung-Russell diagram showing the initial stages of the evolutionary track of a solar-mass star off the zero-age main sequence.

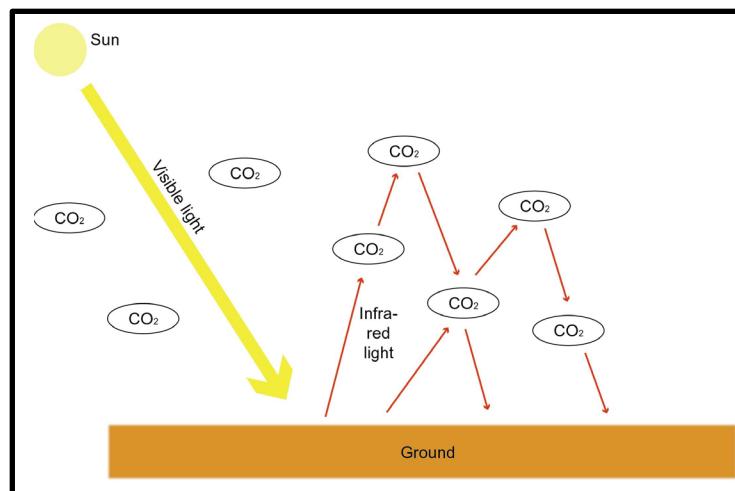


Fig. 8. Schematic illustration of the Greenhouse Effect: CO_2 molecules allow visible light from the Sun to heat the ground, but the resulting infra-red light is trapped within the atmosphere.

Even today, the Earth is some 35°C warmer than it would be in the absence of a greenhouse effect, and a stronger greenhouse effect, probably as a result of more CO₂ in the atmosphere, would have been required to compensate for the fainter Sun. If CO₂ levels were not significantly higher, then some additional greenhouse gas (e.g., CH₄) would also be required.

Estimates for the partial pressure of CO₂ in the Earth's atmosphere 4 Gyr ago vary from about 0.01 bar up to perhaps 1 bar (Fig 7) but are all much higher than the present value of 3.5×10^{-4} bar. As the Sun has brightened with time, CO₂ levels must have declined, or the Earth would have suffered a runaway greenhouse effect like that on Venus. This convenient 'planetary thermostat' may be unique to the Earth in our Solar System, and its implications will be discussed further in Lecture 7. A thicker CO₂ (with or without CH₄) atmosphere is also the most likely explanation for the presence of liquid water on the surface of Mars 3.5 to 4 Gyr ago.

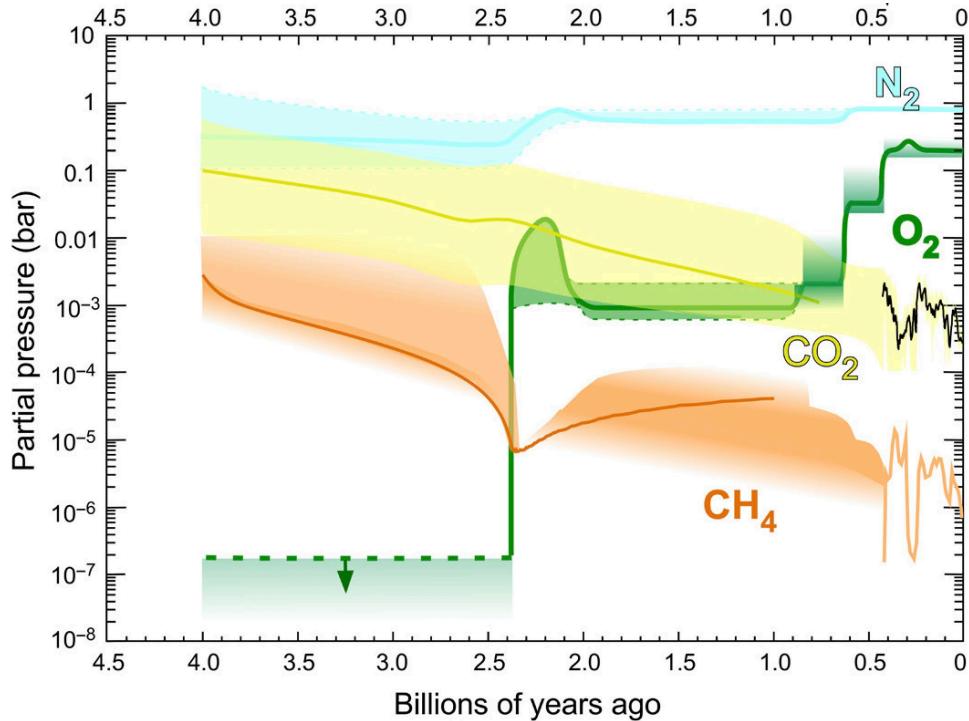


Fig. 9. Estimated partial pressures of gases in Earth's atmosphere over geological time (N₂: blue; O₂: green; CO₂: yellow; and CH₄: orange). Ranges of uncertainty are shown by shading (note the logarithmic scale). Both CO₂ and CH₄ levels are thought to have been higher in the past, giving rise to a stronger greenhouse effect and compensating for the lower solar luminosity. (Diagram from D.C. Catling and K.J. Zahnle, *Science Advances*, 6, eaax1420, 2020. Open Access; Creative Commons 4.0 (CC BY); <https://www.science.org/doi/10.1126/sciadv.aax1420>).

3. The period of ‘heavy bombardment’

One of the major scientific legacies of the Apollo missions to the Moon was the calibration of the impact cratering rate in the inner Solar System. This shows that the rate of meteorite impacts prior to about 3.8 Gyr ago was much higher than it has been since (Fig. 8), and this early period of Solar System history is known as the period of ‘heavy bombardment’.

The last large, basin forming, impact on the Moon was that which produced the Orientale Basin, approximately 3.8 Gyr ago, and most of the other large impact basins for which evidence survives were produced over the preceding few hundred million years. This intense impact flux may be understood within the context of our theories of Solar System formation as the tail end of the accretion of the terrestrial planets.

Given the proximity of the Moon to the Earth, it seems certain that the Earth experienced an equally intense bombardment – indeed, probably it would have been more severe, as the Earth is both a larger target and has a stronger gravitational field likely to attract incoming projectiles. It is estimated that over the period 4.1 to 3.8 Gyr ago the Earth would have acquired more than 200 impact basins over 1000 km in diameter.

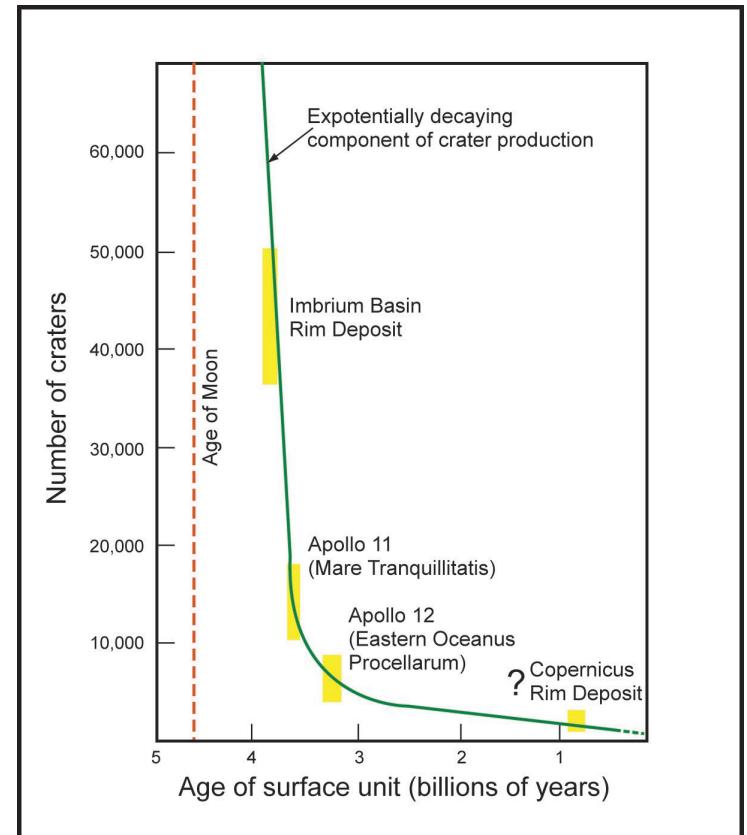


Fig. 10. The impact cratering rate for the inner solar system deduced from the lunar record, as calibrated from radiometrically dated samples collected by the Apollo missions. The large impact flux prior to about 3.8 billion years ago is known as the ‘period of heavy bombardment’ and would have caused severe difficulties for the origin and early evolution of life on Earth. As discussed in Lecture 3, the oldest actual evidence for life on Earth dates from the period 3.5 to 3.8 billion years ago, immediately after the heavy bombardment.

This period of heavy bombardment would have made the early environment of our planet extremely inhospitable for the origin and evolution of life (Fig. 9).

Had the Earth possessed oceans of liquid water at this time (and there is evidence from 4.3 Gyr old detrital zircon grains that it did), then these very large impacts were probably capable of boiling them. For example, an asteroid with a radius of 100 km striking the Earth at 10 km / s (the minimum likely speed) has sufficient kinetic energy to raise the temperature of a 3 km deep ocean by 100°C. Had life evolved in such an ocean, there is a significant chance that it would have been rendered extinct by such an impact. In principle, this could have happened several times, until the impact flux subsided approximately 3.8 Gyr ago – a possibility that has become known as the ‘impact frustration’ of the origin of life.

It is also possible that these violent events may have driven early life forms into deep refuges, either in the crust or to hydrothermal vents at the bottom of the oceans, where they may have been protected somewhat from the bombardment. Colonisation of such refuges may have favoured the evolution of life able to survive at high temperatures, which is consistent with the evidence (discussed in Lecture 7) that life on Earth appears to be descended from thermophilic (meaning ‘high-temperature-loving’) ancestors. It is also interesting to note that the earliest actual evidence we have for life on Earth (3.5 to 3.8 billion years ago; discussed in Lecture 3) dates from very soon after the end of the heavy bombardment. Thus it appears that life became established on this planet almost as soon as the planet became habitable.

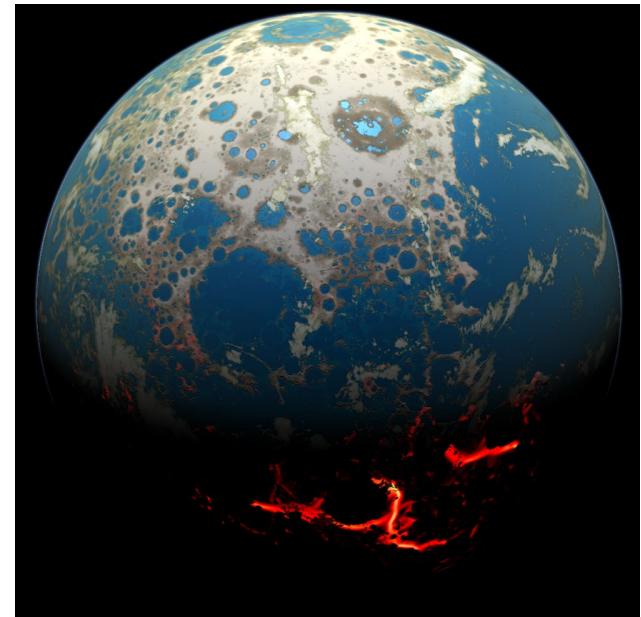


Fig. 11. Artist's conception of what the Earth may have looked like about 4 Gyr ago during the heavy bombardment (image: Simone Marchi/SwRI/NASA).

4. The role of comets and giant planets

4.1 Comets

Comets (Fig. 10) formed in the ice-rich outer regions of the Solar System, and are rich in volatiles and organic molecules. These materials were either inherited directly from the interstellar cloud from which the Solar System formed (Lecture 1) or synthesized in the outer regions of the protoplanetary disk itself, or both.

Given that the terrestrial planets formed in the volatile-poor inner Solar System, the question arises as to how they obtained their volatiles, and it is likely that these materials were added, at least in part, by collisions with comets (or other ice-rich planetesimals). From an astrobiological perspective, the cometary delivery of organic material may have been particularly important, especially if conditions on the early Earth were not conducive to the *in situ* production of organic compounds; we will return to this point in Lecture 6.

4.2 Giant planets

While cometary impacts may have played an important role in delivering volatiles and organics to the early Earth, too many collisions would have been disruptive once life became established. A giant planet at Jupiter's location appears to be effective at clearing planetesimals from the 'middle' regions of the Solar System where these objects are most likely to form. It does this by either sweeping them into itself through collisions, by gravitationally ejecting them from the Solar System.

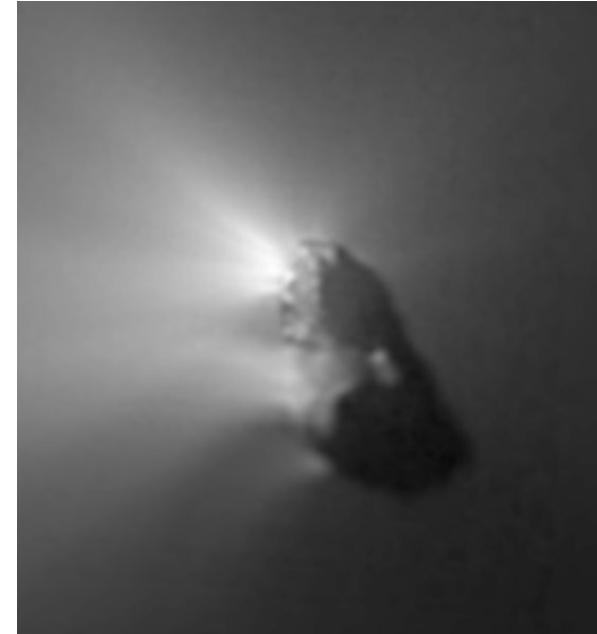


Fig. 12. The nucleus of Comet Halley (about 15 km long) imaged by the Giotto spacecraft in 1986. Cometary nuclei are mostly composed of water ice and other volatiles, and also contain much organic material. They may have been important in delivering volatiles and organics to the newly-formed terrestrial planets in the volatile-poor inner solar system (ESA).

Thus there are probably fewer such bodies in the Solar System than there would be if Jupiter did not exist. Moreover, many comets that today approach the inner Solar System from the Oort Cloud are also captured or deflected by Jupiter (Fig. 11).

Currently, the Earth is struck by one 10-km sized object (of which a significant number are comets, although some will be asteroids) every 100 million years or so. An example is the collision thought to have contributed to the Cretaceous-Paleogene (C-Pg) mass extinction 66 million years ago.

One estimate is that, without Jupiter, the impact rate on Earth may have been 10,000 times higher than it actually is. This would result in a K-Pg sized impact every 10,000 years rather than every 100 million! It seems unlikely that the 'higher' forms of animal life could survive such a frequency of mass extinctions, or that natural selection would have the time in the brief intervals between impacts to build up anything approaching the diversity of life that exists on the Earth today. Thus it has been argued that terrestrial-mass planets in planetary systems which lack a Jupiter-mass planet in a Jupiter-like orbit may never evolve complex forms of life, even if they are well-situated within a habitable zone (although such planets would doubtless have remained habitable for micro-organisms).

However, this protective role of Jupiter is far from certain. While protecting the inner Solar System from comets, Jupiter may increase the impact hazard by sending more asteroids into the terrestrial planet region. These two effects could cancel each other out, and much may depend on the overall 'architecture' of the planetary system concerned (see J. Horner and B.W. Jones, "[Jupiter: Friend or Foe?](#)", *Astronomy & Geophysics*, 51, 6.16, 2010).

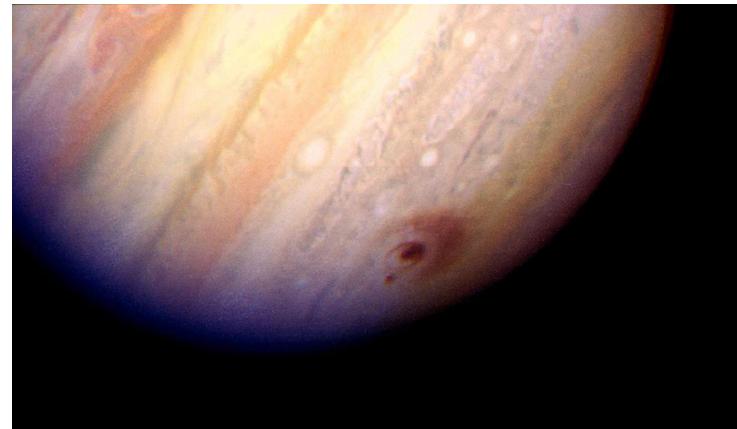


Fig. 13. HST image of one of the scars left on Jupiter by the impact of Comet Shoemaker-Levy 9 in 1994 (NASA).