Atmospheric Escape Processes and Planetary Atmospheric Evolution

G. Gronoff, P. Arras, S. Baraka, J. M. Bell, G. Cessateur, O. Cohen, S. M. Curry, J. J. Drake, M. Elrod, J. Erwin, K. García-Sage, C. Garraffo, A. Glocer, N. G. Heavens, K. Lovato, R. Maggiolo, C. D. Parkinson, C. Simon Wedlund, D. R. Weimer, and W. B. Moore

1Chemistry and Dynamics Branch, NASA Langley Research Center, Hampton, VA, USA, 2Science Systems and Application, Inc., Hampton, VA, USA, 3Department of Astronomy, University of Virginia, Charlottesville, VA, USA, 4National Institute of Aerospace, Hampton, VA, USA, 5Heliophysics Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA, 6The Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium, 7Lowell Center for Space Science and Technology, University of Massachusetts Lowell, Lowell, MA, USA, 8Space Sciences Laboratory, University of California, Berkeley, CA, USA, 9Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA, 10Department of Atmospheric and Planetary Sciences, Hampton University, Hampton, VA, USA, 11Space Science Institute, Boulder, CO, USA, 12Geological and Planetary Sciences (GPS), California Institute of Technology, Pasadena, CA, USA, 13Space Research Institute (IWF), Austrian Academy of Sciences, Graz, Austria, 14Center for Space Science and Engineering Research, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Abstract The habitability of the surface of any planet is determined by a complex evolution of its interior, surface, and atmosphere. The electromagnetic and particle radiation of stars drive thermal, chemical, and physical alteration of planetary atmospheres, including escape. Many known extrasolar planets experience vastly different stellar environments than those in our solar system: It is crucial to understand the broad range of processes that lead to atmospheric escape and evolution under a wide range of conditions if we are to assess the habitability of worlds around other stars. One problem encountered between the planetary and the astrophysics communities is a lack of common language for describing escape processes. Each community has customary approximations that may be questioned by the other, such as the hypothesis of H-dominated thermosphere for astrophysicists or the Sun-like nature of the stars for planetary scientists. Since exoplanets are becoming one of the main targets for the detection of life, a common set of definitions and hypotheses are required. We review the different escape mechanisms proposed for the evolution of planetary and exoplanetary atmospheres. We propose a common definition for the different escape mechanisms, and we show the important parameters to take into account when evaluating the escape at a planet in time. We show that the paradigm of the magnetic field as an atmospheric shield should be changed and that recent work on the history of Xenon in Earth’s atmosphere gives an elegant explanation to its enrichment in heavier isotopes: the so-called Xenon paradox.

Plain Language Summary In addition to having the right surface temperature, a planet needs an atmosphere to keep surface liquid water stable. Although many planets have been found that may lie in the right temperature range, the existence of an atmosphere is not guaranteed. In particular, for planets that are kept warm by being close to dim stars, there are a number of ways that the star may remove a planetary atmosphere. These atmospheric escape processes depend on the behavior of the star as well as the nature of the planet, including the presence of a planetary magnetic field. Under certain conditions, a magnetic field can protect a planet’s atmosphere from the loss due to the direct impact of the stellar wind, but it may actually enhance total atmospheric loss by connecting to the highly variable magnetic field of the stellar wind. These enhancements happen especially for planets close to dim stars. We review the complete range of atmospheric loss processes driven by interaction between a planet and a star to aid in the identification of planets that are both the correct temperature for liquid water and that have a chance of maintaining an atmosphere over long periods of time.

1. Introduction

The discovery of rocky exoplanets at distances from their host stars that might allow stable surface liquid water has led to a blossoming of studies of the habitability of such objects (Anglada-Escudé et al., 2016;
While the ultimate objective of this work is the discovery of life on an exoplanet, detailed investigations of such planets may also shed light on the evolution—both past and future—of the planets in our own solar system (Arney & Kane, 2018), in particular, how they came to be, remain, and/or ceased to be habitable for life as we know it (Editors of Nature Astronomy, 2017; Moore et al., 2017; Tasker et al., 2017).

The usual definition of the “habitable zone” (HZ) (Kasting et al., 1988; Lammer et al., 2009; Ramirez, 2018, and references therein) is where a planet like the Earth would be able to maintain liquid water at its surface; however, it says nothing about whether the planet actually has any liquid water or the necessary atmospheric pressure to stabilize the liquid state. This definition fails to take into account the necessary pathways to habitability: A planet forming in the HZ of a star will have to accrete volatiles from the protostellar nebula to be able to have an atmosphere and liquid water, and it will also have to keep them, which is not necessarily the case for the previously mentioned exoplanets—even if we suppose they have a strong intrinsic magnetic field (Airapetian et al., 2017; Garcia-Sage et al., 2017; Howard et al., 2018). The concept of the HZ is therefore distressingly incomplete, which led to the concept of space weather affected HZ (Airapetian et al., 2017).

One of the best examples of the problems with this definition comes from our understanding of the early Earth: the so-called “Faint Young Sun” (FYS) paradox. Four to 3 Gyr ago, the Sun was fainter by about 30% (Claire et al., 2012), and our models predict that surface water should have been frozen and therefore that Earth was not in the HZ. There is, however, considerable evidence for an active hydrological cycle and exceptionally warm and/or clement temperatures at that period (Kasting & Ono, 2006; Knauth & Lowe, 2003; Lammer et al., 2009; Mojzsis et al., 2001). The typical solution to the FYS paradox has been to propose that the Earth’s early atmosphere had a higher concentration of greenhouse gases such as CO₂, CH₄, NH₃, and N₂O, in a perhaps thicker atmosphere than now (Airapetian et al., 2016; Pavlov et al., 2000; Sagan & Mullen, 1972; Walker et al., 1981). Greenhouse gas levels have overall implications for geological activity, cloud/aerosol formation, and atmospheric chemistry and escape that can preclude their existence, stability, or positive contribution to habitability altogether (Kuhn & Atreya, 1979; Trainer et al., 2006a; Pavlov et al., 2000). Several hypotheses remain concerning the nature of the early Earth’s atmosphere; a major problem lies with the uncertainties on the nitrogen cycle in the past, and on the actual ground pressure that recent studies suggest being closer to 0.5 bar (Laneuville et al., 2018; Som et al., 2016; Zerkle & Mikhail, 2017). For a simple example of the complexity to extend research to exoplanets, consider recent work by Airapetian et al. (2016), which suggests that the higher solar activity has led to chemical reactions creating N₂O, a very efficient greenhouse gas in the early Earth’s troposphere.

Another uncertainty comes from the magnetic activity of the host star, responsible for the space weather conditions of close-in planets, and expected to be much stronger for lower mass stars such as the Trappist-1 system star and M dwarfs in general. Since those stars could remain as active as the young Sun throughout their lifetime (Airapetian et al., 2019), it is theoretically possible that some of the planets orbiting them are currently subject to a N₂O greenhouse effect while at the same time being out of the standard HZ.

In order to produce a more useful concept of habitability, we must contend with all the processes that lead to the habitability of a planet and how the different variables (such as the type of star and the rotation rate of the planet) affect it. The formation of a planetary atmosphere is a balance between the amount of volatiles brought during the accretion phase, and subsequently outgassed, and the subsequent escape or fixing of volatiles as the planet evolves (Lammer et al., 2009; Figure 1).

Atmospheric escape is often overlooked in this type of analysis or only approximated by an energy-limited hydrodynamic escape. Modeling based on this approximation led a fraction of the community to conclude that Pluto’s atmosphere was greatly outgassing until the observations of New Horizons measured an
escape rate 4 orders of magnitude lower than predicted (Gladstone et al., 2016; Zhu et al., 2014). This leads to major questions concerning atmospheric escape that need to be solved.

1.1. The Outstanding Questions of Atmosphere Escape

Several major questions about the evolution of planetary atmospheres have been asked (Ehlmann et al., 2016), such as “are [their] mass[es] and composition[s] sustainable,” and “how do [they] evolve with time?” Recent studies of atmospheric escape have led to the following major questions, specific to escape, that are being answered through experimental studies (e.g., satellites such as Venus Express (VEX), Mars Express (MEX), Rosetta, and Mars Atmosphere and Volatile and EvolutioN missioN (MAVEN)), and theoretical work.

1. What is the current escape rate of planetary atmospheres? How does it vary with forcing parameters? Measurements by different spacecraft enable estimates of the flux of ions and neutrals escaping a planet. However, limitations in temporal and spatial resolution render some observations very difficult; for example, the ion plume of Mars was inferred from MEX observations, but only MAVEN could fully observe and characterize it (Dong et al., 2015; Liemohn et al., 2014); ionospheric outflow at Earth is observed, but the fraction of ions coming back, and the variation of outflow with latitude, magnetic local time, and solar and geomagnetic activity, is difficult to address accurately (Strangeway et al., 2000).

2. What was the escape rate in the past? How did it vary with the varying forcing parameters and the varying atmospheres of planets? The isotopic composition of an atmosphere hints at changes in its composition and can be used to evaluate the total atmospheric loss. However, if some major parameters of the composition have changed, extrapolating the current atmosphere to the past can be problematic. The Earth’s atmosphere is an emblematic example of an atmosphere that has greatly changed, with the appearance of oxygen in large quantities after about 2.5 Gyr ago (Catling, 2014). Observations of Sun-like stars in different stages of their evolution suggest that the Sun had more sunspots and flares in the past, which, undoubtedly, changed the escape conditions of the planets in the solar system (Lammer et al., 2009).

3. How will escape and other atmospheric evolutionary processes shape the future of the planetary atmospheres we observe today? For example, what will the habitability of the Earth and Mars be in a billion years? Variation of the Earth’s magnetic field may affect escape rates and dramatically change the atmosphere of the Earth. At Mars, the atmospheric photochemistry may lead to H2O escape with the oxidation of the crust if O is not escaping enough (Lammer et al., 2003). Recent modeling shows that CO2 + dissociative recombination is also an efficient loss channel (Lee et al., 2015). The adsorption of CO2 into the crust (Hu et al., 2015; Mansfield et al., 2017; Takasumi & Eiichi, 2002; Zent & Quinn, 1995) implies that future change in Mars’ obliquity will increase the outgassing and therefore the surface atmospheric pressure of the planet. But what will happen if no more H2O compensates for the escape? Is it possible for all remaining Martian H2O to escape? How much CO2 could escape?

4. Does a magnetic field protect an atmosphere from escaping? Polar ionospheric outflow is an efficient process to accelerate ions to escape speed. Since it is driven by the energy of the solar wind, funneled by the magnetic field, the stronger the magnetic field, the more energy is available for ionospheric outflow. In that sense, a planet with a magnetic field could be more sensitive to escape (Gunell et al., 2018). However, the returning component of the polar outflow is increasing, and therefore, the net escape should be addressed in different conditions; there are many questions regarding how this component may evolve, and it may be so that it prevents an effective escape altogether. Do the similar escape rates measured at Earth, Venus, and Mars (Gunell et al., 2018) mean that there is no effective shielding, or is the comparison between these planets flawed because the upper atmosphere composition, and therefore, the exospheric temperature are extremely different? Is it just a coincidence that both the Earth and Titan are able to sustain a nitrogen atmosphere despite relatively large exospheric temperatures (more precisely low λex parameter, see section 2.1.2) while being immersed in a magnetosphere? Is the question of the magnetic field protection actually the relevant one?

5. What is the escape rate from exoplanets? Can we test our models against exoplanetary observations? Some observed exoplanets are in hydrodynamic escape (Ehrenreich et al., 2015). It is possible to observe more extreme regimes for exoplanets than for planets in the solar system; therefore, the models developed for the current solar system conditions are likely to be inadequate for exoplanets. One of the main
advantages of these tests is to be able to validate the conditions likely encountered in the early solar system. One example of such a process that is believed to be more important in the past is sputtering, but how could we detect its efficiency at exoplanets?

1.2. Analytical Approach

The solar system has a large variety of planetary bodies, with very different atmospheres, including Mars with a thin CO₂-rich atmosphere, Venus with a thick CO₂-rich atmosphere (both of those presenting evidence of substantial escape), or Earth with a N₂/O₂ atmosphere. The difference between these planets is, in a large part, determined by how they are losing their atmospheres. Several missions, such as MAVEN, MEX, and VEX, have been giving insights on the evolution of planetary atmospheres through their escape to space and have led to a better understanding of which important processes are active to date, and maybe in the past. In addition, work on comets, such as 67P with Rosetta, highlight some of the fundamental processes that lead to escape in slightly different regimes (Brain et al., 2016). Unfortunately, these results cannot be simply extrapolated to exoplanets, since they may be subject to very different conditions.

To that extent, it is necessary to know (1) what the possible mechanisms by which planets lose their atmosphere into space are, (2) how these mechanisms behave with different conditions, (3) how they produce different observables, and (4) what our current understanding of these mechanisms is. Ultimately, one would like to

- Determine what the escape processes are: Review all the processes that have been suggested in the literature, review what their suggested rates were, and, since definitions may vary between authors, decide for a standard definition.
- Determine what the key parameters are for each escape process, that is, what variations will be of importance, and how these parameters couple with each other.
- Determine the unknown parameters that need to be addressed to answer the questions of section 1.1.
- Determine the observable for each escape process and determine how to disentangle the observations of escape in different solar/stellar conditions to determine the relative importance of each processes.

This is why, in the present paper, we start by reviewing the different escape processes and their limitations (section 2), what the major parameters that we need to know to calculate these escape processes and know their importance are (section 3), before looking at how they influence the solar system planets (section 4) and some exoplanets (section 5) in time. We will finally look at which measurements and models are needed to better understand the escape processes at planets and exoplanets (section 6) before concluding.

2. The Escape Processes

The escape processes are usually separated into two parts: the thermal and nonthermal processes. The thermal processes are dependent on the temperature of the upper atmosphere, usually controlled by the host star’s extreme and X ultraviolet (EUV-XUV) flux. The nonthermal processes are the result of more complex interactions, such as plasma interactions. Some nonthermal processes (such as sputtering) have a consistent nomenclature in the literature whereas others (such as ion outflow) have variable definitions depending on the authors. In Table 1, we summarize these escape processes and in Table 2 their main parameters. Those escape processes are sketched on Figure 2, and an evaluation of the current escape rates can be found in Table 3.

Nonthermal escape processes can be separated into photochemical loss (section 2.2), ion loss (section 2.3), ionospheric outflow (section 2.4), and other losses (section 2.5). Moreover, in order to compute the total loss of an atmosphere into space, it is necessary to take into account the problem of the ion return (section 2.6). It is important to note that, while we are separating these processes, they do influence each other, and sometimes, one leads to the other. For example, an ionospheric outflow process at Venus can produce fast particles involved in ion pickup and sputtering (Luhmann et al., 2008).

2.1. Thermal Escape

Thermal escape is one of the most important escape processes (Chassefière & Leblanc, 2004; Selsis, 2006). It takes place in two regimes, Jeans escape, and hydrodynamic escape, with a transition regime that is the subject of recent studies (e.g., Erkaev et al., 2015; Strobel, 2008b; Volkov, Johnson, et al., 2011; Volkov,
Most of the observed isotopic fractionation in planetary atmospheres is interpreted as originating from thermal escape because of its energy efficiency at escaping large amount of gases.

### 2.1.1. Fundamental Theory

#### 2.1.1.1. Jeans Regime

The neutral atmospheric constituents in the upper atmosphere are in local thermodynamic equilibrium (LTE; or close to it). Therefore, their distribution function can be approximated by a Maxwellian function (Mihalas & Mihalas, 1984):

$$f(x, v) = N \left(\frac{m}{2\pi kT}\right)^{3/2} e^{-\frac{mv^2}{2kT}}$$

$$= N \left(\frac{1}{u_i\sqrt{\pi}}\right)^3 e^{-\frac{v^2}{u_i^2}},$$

(1)

where $u_i = \sqrt{\frac{2kT}{m_i}}$ is referred to as the thermal speed for the species $i$. 

### Table 1

**The Escape Processes**

| Process                        | Origin                                      | Key parameters                                      |
|--------------------------------|---------------------------------------------|-----------------------------------------------------|
| Jeans escape                   | Temperature accelerate particle above the escape velocity | Temperature, gravity, $T_c\lambda_{ex}$ parameter > 2.5 |
| Hydrodynamic escape            | Thermal acceleration in a fluid way         | Temperature, gravity, $T_c\lambda_{ex}$ parameter < 2.5 |
| Photochemical/ion recombination | Ion recombination releasing kinetic energy  | Low gravity, molecular ion, requires ionosphere densities |
| Photochemical/dissociations    | Molecular photodissociation release kinetic energy | Requires thermosphere densities and low gravity |
| Ion pickup                     | Solar wind picks up ions from ionosphere    | Requires compressed/no magnetosphere |
| Ion sputtering                 | Accelerated ions from the ionosphere translate their kinetic energy | Requires compressed/no magnetosphere, B, $U_{sw}$ |
| Charge exchange/trapped        | Fast ion trapped in magnetosphere becomes ENA through charge exchange | Requires magnetosphere, ion density and temperature, and neutral densities |
| Charge exchange/solar wind     | Solar wind ion becomes ENA that can access thermosphere and increases heating | Requires large coronae, $U_{sw}$ and $N_{sw}$ |
| Charge exchange/particle precipitation | Particle precipitating in thermosphere becomes ENA and translate kinetic energy | Requires precipitaton fluxes and cross sections |
| Ionospheric outflow           | creation of ion upward wind through ambipolar diffusion | Requires fields and ionosphere |
| (often called polar wind in magnetized planets) | | |
| Other ion escape               | Plasma instabilities leading to ions going upward and being picked by the solar wind | Requires fields, ion density, and temperature |

### Table 2

**Escape at Planets, Parameters Compiled From Hinson et al. (2017), Young et al. (2018), and Johnson et al. (2013b)**

| Planet   | $\lambda_{ex}$ parameter | $T_c$ (K) | $T_e$ (K) | $g$ (m/s$^2$) | $R$ (km) | $H_{esc}$ (km) | $B$ (Gauss$^2$) | Average solar EUV (W/m$^2$) | Solar wind pressure (nPa) | $Q_c$ (W) |
|----------|--------------------------|-----------|-----------|--------------|---------|---------------|----------------|--------------------------|----------------------------|----------|
| Mercury  | 2.2                      | 500       | 725       | 3.70         | 2,439.7 | 0.002         |                | 9,082.7                  | 13.8–21.0               | 7.31 $\times 10^{10}$ |
| Venus    | 22.3                     | 290       | 4307      | 8.87         | 6,051.8 | 15.9          |                | 2,601.3                  | 1.0–12.0                 | 2.64 $\times 10^{12}$ |
| Earth    | 9.4–5.0                  | 800–1600 | 5020      | 9.80         | 6,378.1 | 8.5           | 0.306          | 1,361.0                  | 1.0–6.0                  | 3.51 $\times 10^{12}$ |
| Moon     | 0.8                      | 226       | 400       | 1.62         | 1,738.1 |               |                | 1,361.0                  | 1.0–6.0                  | 1.03 $\times 10^{10}$ |
| Mars     | 6.3–5.0                  | 240–300   | 1014      | 3.71         | 3,396.2 | 11.1          |                | 586.2                    | 0.1–1.1                  | 1.68 $\times 10^{11}$ |
| Jupiter  | 311–218                  | 700–1000  | 145000    | 24.79        | 71,492  | 27.0          | 4.30           | 50.26                    | 0.05–0.10                | 5.92 $\times 10^{15}$ |
| Saturn   | 157–98                   | 500–800   | 52200     | 10.44        | 60,268  | 59.5          | 0.215          | 14.82                    | 0.01–0.09                | 1.06 $\times 10^{15}$ |
| Titan    | 2.3                      | 180       | 280       | 1.35         | 2,575   |               |                | 14.82                    | 0.01–0.09                | 1.84 $\times 10^{10}$ |
| for CH$_4$ | 37.3                   | 4475      |           |              |         |               |                |                          |                            | 3.7 $\times 10^{11}$  |
| Uranus   | 34                       | 800       | 18300     | 8.87         | 25,559  | 27.7          | 0.228          | 3.69                     | 0.001–0.02               | 1.14 $\times 10^{14}$ |
| Neptune  | 48                       | 700       | 22250     | 11.15        | 24,764  | 19.1–20.3     | 0.142          | 1.508                    | 1.26 $\times 10^{14}$  |
| Pluto    | 15.1                     | 68        | 408       | 0.62         | 1,184   | 78            |                | 0.873                    | 0.006                    | 8.4 $\times 10^{8}$   |
| for CH$_4$ | 8.5                    | 384       |           |              |         |               |                |                          |                            | 1.6 $\times 10^{10}$  |
The exobase is quantitatively defined as the level where $l_i$, the mean free path of the $i$th constituent is equal to the scale height ($H$) (Hunten, 1973; Shizgal & Arkos, 1996). At the exobase, we consider that a molecule of the $i$th constituent going upward at the escape velocity, $v_{esc} = \sqrt{2GM/r}$ will not impact another molecule and therefore will escape. This approximation is the equivalent of considering an atmosphere collisionless above the exobase and fully collisional below (Fahr & Shizgal, 1983).

By integrating the vertical flux, $v_i \cos(\theta) \times f_i$, at the exobase, for the velocities greater than the escape velocity ($v_{esc}$), and neglecting the collisions above it, we retrieve the flux of escaping molecules.

### Table 3
The Present Escape Values

| Process                                | Venus                        | Earth                        | Mars                          |
|----------------------------------------|------------------------------|------------------------------|-------------------------------|
| Jeans escape                           | $2.5 \times 10^{19} - 5.1$ $^{(1)}$ | H: $6 \times 10^{26} - 10^8$ (solar max) $^{(2,6)}$ | H: $6 \times 10^{26} - 10^8$ (solar min) $^{(2,6)}$ |
| Charge exchange/trapped                | H: $5 \times 10^{24} - 5 \times 10^{25} - 10^6$ $^{(2)}$ | H: $6 \times 10^{26} - 10^8$ (solar min) $^{(2,6)}$ | 10$^{-22}$ - 10$^{-23}$ - 10$^{-5}$ $^{(2)}$ |
| Ion pickup                             | $H^+ : 10^{25} - 2 \times 10^{6}$ $^{(1)}$ | Small $^{(2)}$              | O$^+$ : $10^{24} - 10^{6}$ $^{(2,3)}$ |
| Sputtering                             | He$^+$ : $5 \times 10^{24} - 10^{6}$ $^{(2)}$ | Small $^{(2)}$              | C$^+$ : $1.6 \times 10^{24} - 10^{-5}$ $^{(2)}$ |
| Photochemical escape                   | $3.8 \times 10^{25} - 7.7 \times 10^6$ $^{(1)}$ | Small $^{(2)}$              | O : $3 \times 10^{24} - 1.8 \times 10^{6}$ $^{(3)}$ |
| Magnetized ion outflow (polar wind)    | N/A                          |                               | C : $10^{24} - 10^7$ (solar min) $10^{-5} - 10^{-7}$ (solar max) $^{(2)}$ |
| Unmagnetized ion outflow/K-H/clouds   | $5 \times 10^{24} - 1 \times 10^{25} - 1 \times 10^6 - 2 \times 10^6$ $^{(1)}$ | 10$^{-23} - 10^6$ $^{(3)}$ | O : $5 \times 10^{24} - 3 \times 10^3$ $^{(2)}$ |

Note: Total escape, in s$^{-1}$, followed by the fluxes, in cm$^{-2}$ s$^{-1}$. Both are reported in the literature. While fluxes show the magnitude at a given planet, highlighting the intensity, total escape highlights the overall aspect of escape; when comparing planet to planet, none are satisfying since comparing the total escape of a Mars with, for example, Venus hides the size effects. On the other hand, comparing the fluxes from Earth with, for example, Mars, hides local effects like exospheric temperature. We decided to show both values. References: $^{(1)}$ Lammer et al. (2008) and references therein. $^{(2)}$ Tian et al. (2013) and references therein. $^{(3)}$ Jakosky et al. (2018) and references therein. $^{(4)}$ Gunell et al. (2018) and references therein. $^{(5)}$ Inui et al. (2019) and references therein. $^{(6)}$ At Earth, the Jeans flux for the solar max is basically the same flux as for charge exchange at the solar min because H escape is diffusion limited; see section 2.3.2. NB: we considered the fluxes at the exobases, at Venus, a 200-km exobase has a surface of $4.9 \times 10^{18}$ cm$^2$; at Earth, for a 500-km exobase, $5.9 \times 10^{18}$ cm$^2$; at Mars, for a 200-km exobase, $1.6 \times 10^{18}$ cm$^2$. 

---

**Figure 2.** The main processes of atmospheric escape, along with their typical efficient altitudes domains (near the thermosphere/exobase or away from it) and their conditions of efficiency/occurrence (magnetic field).
\[ \Phi_i(\text{escape}) = \int_0^{v_i} \int_0^{v_{esc}} v \cos(\theta) f(v_i) v_i^2 \sin(\theta) dv_i dv \theta. \] (2)

Carrying out this integration gives

\[ \Phi_i(\text{escape}) = N_i \left( \frac{kT_e}{2\pi m_i} \right)^{1/2} \left( 1 + \frac{m_i v^2_{esc}}{2kT_e} \right) e^{-\frac{m_i v^2_{esc}}{2kT_e}} \] (3)

\[ = N_i \left( \frac{u_i}{2\sqrt{\pi}} \right) \left( 1 + \frac{v^2_{esc}}{u_i^2} \right) e^{-\frac{v^2_{esc}}{u_i^2}} \] (4)

In Equations 3 and 4, it is important to use the values of \( v_{esc} \) and \( u_i \) at the exobase (using the temperature, \( T_{exo} \), and radius \( r_{exo} \) at the exobase) to get a correct estimation of the escape flux.

It is common to introduce the nondimensional Jeans parameter to express the escape flux, and we will see later that this parameter is very useful in understanding the thermal escape process. The Jeans parameter is the ratio of gravitational energy to thermal energy, expressed at \( \lambda_{exo} = \left( \frac{GMm_i/r}{kT} \right) = \frac{v^2_{esc}}{u_i^2} \). Using this, the escape flux becomes

\[ \Phi_i(\text{escape}) = N_i \left( \frac{u_i}{2\sqrt{\pi}} \right) \left( 1 + \lambda_{exo} \right) e^{-\lambda_{exo}}. \] (5)

Equation 2 assumes that we can approximate the distribution at the exobase by a Maxwellian despite the fact that molecules faster than \( v_{esc} \) are removed. When the escape rate is high enough, a non-Maxwellian correction must be applied to consider that the high-energy tail of the Maxwellian is depleted, following Chamberlain and Smith (1971). This correction lowers the escape rate by about 25%. However, this correction is based on the assumption of an isothermal atmosphere below the exobase and has been evaluated for H and He escape within a O or CO\(_2\) rich background atmosphere, that is, the thermosphere of the Earth and Mars/Venus.

A more realistic simulation, performed by Merryfield and Shizgal (1994), that considered the effect of increasing temperature with altitude shows that the escape from the deeper layer should also be considered (i.e., it cannot be assumed that a Maxwellian is a good approximation of the atomic/molecular distribution at the exobase). In that case, the correction is an increase of the order of 30%. Therefore, for extremely precise determination of escape, it is important to solve the Boltzmann equation; one of the most used techniques is the direct simulation Monte Carlo (DSMC) method (Tucker & Johnson, 2009; Volkov, Johnson, et al., 2011), whose results show that the source of escaping particles is distributed over a wide altitude range above and below the exobase.

Overall, Equation 4 is a good approximation for the thermal escape when the atmosphere is strongly gravitationally bound to the planet, and this formula is valid for all the constituents independently. Ideally, this equation would be evaluated at or near the nominal exobase but can be applied far below the exobase assuming no addition heating if a correction factor is applied (Johnson et al., 2016; Volkov, 2015).

It is important to note that while the Jeans parameter is the main parameter of thermal escape, the location of the exobase is extremely important. In the case of Titan or a possible early Moon atmosphere (Aleinov et al., 2019) the altitude of the exobase is nonnegligible compared to the radius of the planet, and while the flux per unit surface is small, it can become the most important source of loss when taking the whole exobase surface into account.

### 2.1.1.2. Hydrodynamic Regime

In cases where the internal energy of individual gas molecules approaches the kinetic energy required for escape (\( \lambda_{exo} \approx 1 \)), the gas will begin to escape as a flow of continuous fluid (Gross, 1972; Hunten, 1973; Watson et al., 1981).

Qualitatively, the fundamental distinctions between the Jeans and hydrodynamic regimes can be viewed in two helpful and complementary ways. First, the Jeans regime is “collisionless” (Shizgal & Arkos, 1996): It is
one where collisions between molecules define an exobase as a surface (or at least a narrow region). The atmosphere is retained not only by the gravitational pull on individual molecules but also by the effective force of collision with other atmospheric molecules. In the “collisional” hydrodynamic regime, the molecules are so energetic that collisions are insufficient to restrict escape. Indeed, the escaping flow of lighter gases (the ones that are most likely to be escaping) is capable of exerting an effective force and dragging heavier gas molecules such as water and the noble gases (Pepin, 1991; Zahnle & Kasting, 1986). And furthermore, hydrodynamic escape can take place far below the exobase.

Second, the distinction between Jeans and hydrodynamic escape is analogous to that between evaporation at temperatures below the boiling point and boiling. In this analogy, the exobase is like the surface of the evaporating fluid, the force of gravity is like the atmospheric pressure, and the effective pressure of other molecules is like the effective pressure of other molecules in the liquid. Quantitatively, hydrodynamic escape is approached by numerical solution of an appropriate system of inviscid fluid dynamical equations (Tian & Toon, 2005; Watson et al., 1981). For instance, the one-dimensional time-dependent inviscid equations for a single constituent, thermally conductive atmosphere in spherical geometry, are given by Tian and Toon (2005) as

\begin{align}
\frac{\partial (\rho v^2)}{\partial t} + \frac{\partial (\rho v^2 r^2)}{\partial r} &= 0, \\
\frac{\partial (\rho v^2 r^2)}{\partial t} + \frac{\partial (\rho v^2 r^2 + pr^2)}{\partial r} &= -\rho GM + 2pr, \\
\frac{\partial (E r^2)}{\partial t} + \frac{\partial [(E + p)vr^2]}{\partial r} &= -\rho v GM + q r^2 + \frac{\partial kr^2 \partial T}{\partial r},
\end{align}

where \( E = \rho (v^2/2 + e) \) is the equation for the total energy density, \( e = p/[(\gamma-1)] \) is the definition of the internal energy, and \( p = \rho RT \) is the ideal gas law. Here, \( \rho \) is the density of the gas, \( p \) is the pressure, \( \gamma \) is the polytropic index of the gas, \( R \) is the universal gas constant, \( k \) is the thermal conductivity, and \( q \) is the volume heating rate. Thus, Equation 6 is the continuity equation, Equation 7 is the momentum conservation equation, and Equation 8 is the energy conservation equation.

If energy conservation is neglected and the temperature is held constant, steady state solutions to the above system are possible. It is convenient in that case to rewrite \( p \) and \( GM \) such that

\begin{align}
p &= \rho v_0^2, \\
GM &= 2r_0 v_0^2,
\end{align}

where \( v_0 \) is the sound speed and \( r_0 \) is a critical radius based on the relative gravitational potential energy and kinetic energy of a particle at the sound speed.

A single differential equation is then obtained:

\begin{equation}
\frac{1}{v_0} \frac{dv}{dr} \left(1 - \frac{v^2}{v_0^2}\right) = \frac{2r_0}{r^2} - \frac{2}{r}
\end{equation}

Equation 11 has an obvious critical point at \((\pm v_0, r_0)\), and thus, various solutions to the differential equation can be derived by integrating from these two critical points to some other velocity and radius, assuming \( v_0 \) is constant. The example of interest here is the transonic outflow solution obtained by integrating Equation 11 from the critical point to higher velocity, \( v \), and radius, \( r \).

\begin{equation}
\log \frac{v}{v_0} = \frac{1}{2} \left(\frac{v^2}{v_0^2}\right) + \frac{2r_0}{r} + 2\log \frac{r}{r_0} = \frac{3}{2}
\end{equation}

A good discussion of the various solutions and their significance can be found in Pierrehumbert (2010).
It is possible to derive a theoretical upper bound for hydrodynamic escape of a single constituent atmosphere at a given temperature and atmospheric level. This bound is given by Equation 4 in the limit where $\lambda_\infty \approx 0$ (Hunten, 1973):

$$
\Phi_i(\text{escape}) = N \left( \frac{kT_e}{2\pi m} \right)^{1/2}.
$$

(13)

At this bound, escape takes the form of a one-dimensional outflow at the thermal velocity. In realistic models of atmospheres; however, hydrodynamic loss rates tend to be much lower than the theoretical limit for reasons to be discussed below. In addition, the use of an hydrodynamic escape approach is easily abused, especially when many assumptions have to be made on the nature of the atmosphere (such as the composition and the exospheric temperature). A solution to that problem is to estimate the critical heating rate (Johnson et al., 2013a, 2013b, and section 3.1.1).

2.1.1.3. Fluid-Kinetic Models

Applying the Jeans equation requires that the temperature and density to be known near the exobase. As an alternative to the hydrodynamic model, one can still use the fluid Equations 6–8 by utilizing the Jeans escape rate and energy escape rate as upper boundary conditions. This fluid-Jeans model has been adapted to hot gas giants (Yelle, 2004), as well as to terrestrial planets like Earth (Tian et al., 2008). One advantage of these methods is the solution is valid up to the exobase, so that heating, photochemistry, and diffusion can be included and their effects on the escape rate investigated. Using this fluid-Jeans model, Tucker et al. (2012) and Erwin et al. (2013) refined the escape rate using DSMC to get a fluid-DSMC result. This extends the fluid-Jeans result to model the transition from the collisional to collisionless regimes and demonstrates the breakdown in the fluid equations below the exobase. These models predict that escape rates at Titan and Pluto are roughly consistent with Jeans escape even with low gravity or with high heating rates.

2.1.1.4. Limiting Factors to Thermal Escape

One limit arises from the impact of thermal escape on the energetics of the upper atmosphere. Removal of the escaping material either will cool the atmosphere around the exobase or lower the altitude of the exobase. Either way, some energy source will be necessary to maintain escape in a steady state. That energy source is whatever stellar EUV <90 nm that can be absorbed near the exobase (this absorption threshold for upper atmospheric species), which results in one form of the energy-limited escape rate (Erkaev et al., 2007; Watson et al., 1981):

$$
\Phi_i(\text{limited escape}) = \varepsilon F_{\odot, \text{EUV}} \left( \frac{GMm}{r} \right)^{-1},
$$

(14)

where $\varepsilon$ is the EUV heating efficiency and $F_{\odot, \text{EUV}}$ is the solar (or stellar) EUV flux.

When fluxes of EUV are high, energy-limited escape is defined by the balance between conductive heating of absorbed solar EUV from the exobase with adiabatic cooling of the thermosphere, as initially argued by Watson et al. (1981). Thus, at increasingly higher levels of EUV flux, the thermospheric temperature profile should evolve from one in which temperature increases monotonically to the exobase to one where peak temperature is significantly below the exobase. And after a certain point, the higher the incoming solar flux, the lower the exobase temperature (Tian et al., 2008): a regime thought to have limited the thermal escape rate on the early Earth (the authors refer to hydrodynamic escape as the regime where the adiabatic flow is important in the upper thermosphere, even though they are using Jeans escape to define the escape rate at the upper boundary). Where increased EUV flux simultaneously reduces other cooling mechanisms (such as infrared [IR] emission from CO$_2$ on Mars Tian, 2009), the adiabatic cooling-driven energy limit to thermal escape is less relevant. Erwin et al. (2013) showed that Pluto’s atmospheric escape is energy-limited even with the small EUV flux experienced at its orbit.

The final limit arises from the impact of thermal escape on the composition of the upper atmosphere. Escaping species typically cannot be supplied to the escaping region of the atmosphere at rates comparable to the various theoretical upper limits for either Jeans or hydrodynamic escape. Escape rates are then controlled by the flux of escaping species to the region of escape, a regime known as diffusion-limited escape because diffusion is the principal transport mechanism in the escape regions of the most familiar
planetary atmospheres (Hunten, 1973; Kasting & Catling, 2003). Consider a trace gas of density $n_i$ significantly lighter than the mean molecular mass of a planetary atmosphere and present at the homopause, where eddy diffusion is too weak to mix the atmosphere thoroughly. The separate gases will unmix by molecular diffusion and segregate. Unmixing at the homopause sets the limiting diffusion rate, which is dependent on the mixing ratio at the homopause itself as well as the diffusion coefficient of the light trace gas in the heavier principal constituents ($b_{i,\text{dom}}$) (Hunten, 1973). (In the following equation the mass of the trace gas is very small compared to the mass of the principal constituent, its number density is also very small, and we neglect thermal diffusion; we will explore diffusion-limited escape more in section 3.1.3; $n$ is the number density of the main constituent).

$$\Phi_i(\text{escape}) = \left( \frac{b_{i,\text{dom}}}{H} \right) \frac{n_i}{n} \text{n}(15)$$

Note that the dependence of the diffusion rate at the homopause on the concentration of the light species at the homopause makes diffusion-limited transport also dependent on all barriers to transport of the light gas lower in the atmosphere such as an atmospheric cold trap. Escape of H at Earth is a perfect example of diffusion-limited escape (Shizgal & Arkos, 1996).

Thus, thermal escape has three classes of rate limit: (1) an absolute one based on fluid dynamics at the exobase, (2) an energetic one based on the absorption of solar EUV near the exobase, and (3) a compositional one based on atmospheric vertical transport below the exobase.

### 2.1.2. Key Parameters

The most important parameters controlling thermal escape are the atmospheric scale height in the thermosphere, $H$, which depends upon the exospheric temperature $T_e$ (section 3.3.2) and the mass of the atmospheric constituents, $m_i$.

The regime of thermal escape is governed by the dimensionless Jeans parameter $\lambda_{\text{ex}} = (GMm_i/r)(kT)$ with $r$ being taken either as the distance from the center of the planet to the surface, exobase, or the location of the molecule(s) in question. There is a critical value for $\lambda_{\text{ex}}$, below which there is a transition between hydrodynamic and Jeans escape. Simply equating the internal energy and the escape velocity would suggest that the critical value of $\lambda_{\text{ex}}$ is

\[ \frac{1}{\gamma - 1} \]

with $\gamma$ begin the heat capacity ratio, directly linked to the degree of freedom of the molecule/atom. This would correspond to 1.5 for ideal monoatomic gases and 2.5 for ideal diatomic gases. Thus, Selsis (2006) refers to a critical value of 1.5 for simplicity. Simulations by Erkaev et al. (2015) of an atmosphere dominated by $H_2$ show a transition in escape rates near $\lambda_{\text{ex}} = 2.5$, which implies that

\[ \frac{1}{\gamma - 1} \]

is indeed a good estimate of $\lambda_{\text{ex}}$.

Following Selsis (2006), we can define a critical temperature $T_c$ for which $\lambda_{\text{ex}} = 1.5$ for the different planets, which is valid for a $H$ atom.

At Titan, the eventual escape of material to space is determined by the combined effects of the deep atmosphere limiting flux and the effects of photochemical loss (for $CH_4$) or production (for $H_2$) above the homopause region (Bell et al., 2014). Each of the major species, $N_2$, $CH_4$, and $H_2$, possesses separate critical points, but the nominal exobase is located near 1,500 km, which is a significant fraction of the radius of Titan (2,575 km).

Selsis (2006) gives a table of critical temperature, $\lambda_{\text{ex}}$ (noted $\chi$ in that paper) and exospheric temperature for different objects in the solar system. Table 2 is an update taking into account the recent data, for example, from New Horizons.

In Johnson et al. (2013a, 2013b) a criterion for where the transition between Jeans escape and hydrodynamic escape should be considered, based on the heating rates, has been described.

### 2.1.3. Questions and Important Points

#### 2.1.3.1. How the Transition Between the Thermal Escape and the Hydrodynamic Escape is Done

Motivated by Cassini spacecraft data for Titan and New Horizons data for Pluto, there has been renewed interest in the physical assumptions underlying planetary escape. Following Hunten (1982, 2013b), it was
assumed that if the binding parameter $\lambda_{ex}<1$ near the exobase, an organized hydrodynamic flow would result, whereas if $\lambda_{ex}>10$, that collisionless Jeans escape would result. Intermediate models called \textit{slow hydrodynamic escape} including transport effects such as thermal conduction were also developed (Hunten & Watson, 1982; Strobel, 2008b; Watson et al., 1981) to bridge the intermediate values of $\lambda_{ex}$ between the two limits. Recently, Volkov, Johnson, et al. (2011) used DSMC to model atom/molecule motions under gravity and collisions. It was assumed that heating occurred below the base of the simulation domain, so that particles enter the domain with a Maxwell-Boltzmann distribution at a prescribed temperature. Subsequent collisions between particles then transport heat upward effectively by a heat conduction flux (although the Fourier law may be inaccurate to describe this flux). The particle density at the base of the simulation domain was parameterized through the ratio of mean free path to the scale height (the Knudsen number), which is a measure of the frequency of collisions. The surprising result of the simulations presented in Volkov, Johnson, et al. (2011) was that a sharp transition occurs from the hydrodynamic to the Jeans escape limits, near $\lambda_{ex} \sim 2-3$ depending on the particle interaction law. Analytic support of these results was given in Volkov, Johnson, et al. (2011). For $\lambda_{ex}>3$, the bulk fluid velocity never becomes supersonic, and the escape rate is near the Jeans escape rate. Hence, given the assumptions of that study, hydrodynamic outflow is limited to small values of the binding parameter. Early in the Cassini mission to the Saturn system, Strobel (2008b) posited that slow hydrodynamic escape could be occurring in the upper atmosphere of Titan, due the moon's low gravity and the extended nature of its atmosphere. Further still, the combined works of Strobel (2008b, 2012) and Yelle et al. (2008) went a step further and suggested that hydrodynamic escape was in fact the only mechanism that could adequately reproduce the observations of methane. However, later investigations by Bell et al. (2011) and later in Bell et al. (2014) demonstrated that, by self-consistently coupling dynamics, composition, and thermal structure calculations, the in situ measurements of methane by the Ion-Neutral Mass Spectrometer (INMS) (Magee et al., 2009; Waite et al., 2004) could be explained with the atmosphere in a nearly diffusive state without the need for invoking slow hydrodynamic escape of methane.

Similar to the situation at Titan, the data obtained by the New Horizons flyby of Pluto and Charon were not consistent with a previously posited hydrodynamic escape mechanism occurring at the dwarf planet (Gladstone et al., 2015). Prior to this observation, Pluto was suggested to be the archetype for a planetary atmosphere in a state of hydrodynamic escape. Instead, the DSMC simulation by Tucker and Johnson (2009) and Tucker et al. (2012), which suggested that Pluto's atmosphere could be simulated without invoking hydrodynamic escape, seems to better match observations made by New Horizons. Thus, despite being posited as occurring at several bodies in the solar system, there is no clear evidence for slow hydrodynamic escape occurring in our solar system during the current epoch.

\subsection*{2.1.4. Observables}

When observing escape in real time, thermal escape can be viewed as principally a function of the density of the escaping species and exospheric temperature (Equation 4). A typical technique is to infer density and temperature from airglow emission, which is also a function of density and exospheric temperature (e.g., Chaffin et al., 2014). In some cases, in situ mass spectrometry of neutrals can enable better constraints on density (e.g., Cui et al., 2008), while satellite drag can add yet another constraint jointly dependent on bulk atmospheric density and temperature (e.g., Krauss et al., 2012).

The central value of observing airglow emission for planets in the solar system and the difficulty of obtaining additional constraints on escape from exoplanets strongly suggests that airglow emission will be the key observable for quantifying thermal escape at exoplanets, whether by Jeans or hydrodynamic escape. The expected observable for intense hydrodynamic escape is of a highly extended hydrogen corona containing relatively large amounts of heavier atoms rather than a rapid fall-off in the concentration of such atoms beyond the exobase (Vidal-Madjar et al., 2003). Airglow, however, is extremely difficult to observe at exoplanets and can be affected by particle precipitation (Bernard et al., 2014). For small/rocky planets such as an Earth-like or a Mars-like exoplanet, a technique based on CO$_2$ or O$_2$ absorption due to stellar occultation in the near ultraviolet (UV) can be used but is extremely challenging (Gronoff et al., 2014).

The main observable for thermal escape in a planet's past is mass fractionation of the isotopic composition of the atmosphere from the stellar value. However, caution must be exercised. Isotopic composition can be affected by the outgassing of primordial materials and low-temperature chemical reactions unrelated to
escape (Pepin, 2006; Pope et al., 2012). Moreover, isotopic composition is strongly sensitive to Jeans escape but variably sensitive to hydrodynamic escape.

For Jeans escape, it can be inferred from Equation 4 that the escape rate is proportional to $m_i^{-1/2}$ for small values of $\lambda_{esc}$ and $m_i^{1/2}e^{-\lambda_{esc}}$ for large values of $\lambda_{esc}$. The former case would be hydrodynamic escape. So for Jeans escape, deuterium escapes at a rate less than atomic hydrogen.

In the case of hydrodynamic escape, the principal escaping species drags gases lighter than the “crossover mass” ($m_c$) (Hunten et al., 1987).

$$m_c = m_{esc} + \frac{kT_{esc}}{bgX_{esc}}$$  

(16)

where $m_{esc}$ refers to the principal escaping species, $b$ is the binary diffusion coefficient (the diffusion coefficient in a two-component gas), and $X$ is the mole fraction. If the escape flux of the principal escaping species can be defined at a reference altitude $\Phi_{esc}$ and is sufficiently small, then the escape flux of the trace species at the reference altitude $\Phi_{trace}$ is

$$\Phi_{trace} = \frac{X_{trace}}{X_{esc}} \Phi_{esc} \left[ \frac{m_c - m_{trace}}{m_c - m_{esc}} \right]$$  

(17)

(Hunten et al., 1987). It is in these slower hydrodynamic escape cases that significant fractionation is possible on geological timescales. Otherwise, the larger species are carried along with the flow. And the escape rate of each isotope is proportional to its mole fraction, lowering the fractionation effect.

$$\Phi_{trace} = \frac{X_{trace}}{X_{esc}} \Phi_{esc} \left[ 1 - \frac{bgX_{esc}}{kT_{esc}} \left( m_{trace} - m_{esc} \right) \right]$$  

(18)

(Hunten et al., 1987). In this case, fluxes are weakly dependent on mass at masses close to the mass of the principal escaping species but more strongly dependent on mass at masses much greater than that of the principal escaping species, resulting in minimal fractionation of low mass species but significant fractionation of high mass species (Hunten et al., 1987; Tian et al., 2013).

As noted in Pepin (1991), Shizgal and Arkos (1996), and Pepin (2006), the uncertainty in the hydrodynamic escape parameters, notably with the EUV output of the Young Sun, the noble gas reservoirs, the volatile outgassing, etc., is a problem to retrieve the whole history of a planetary atmosphere. In addition, other escape processes lead to isotopic fractionation.

### 2.2. Photochemical Escape

The dominant nonthermal loss processes vary for each planetary body. The relative significance of each process depends on planetary mass, atmospheric composition, and distance from the Sun. For instance, at Mars, the current dominant nonthermal loss processes are photochemical, while at Venus it is thought to be through ionospheric escape (Lammer et al., 2008).

The photochemical escape of a planetary atmosphere is a nonthermal loss process due to exothermic chemical reactions in the ionosphere that provide enough kinetic energy for the escape of the neutral constituents. Photochemical escape often includes direct interactions of photons and photoelectrons with thermospheric and exospheric molecules, as well as chemical reactions of ions with neutrals and electrons. In the following, we will add the symbol "*" to neutral and ionized species to show that they have a nonnegligible amount of kinetic energy. Such species are usually called “hot”; and for the neutral atoms, the term ENA, for energetic neutral atom, is often used.

The general method of computation for the escape of a fast atom or ion can be found in Shematovich et al. (1994). The general transport equation for any species in the atmosphere is
\[
\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \frac{\mathbf{F}_m}{m} \frac{\partial f}{\partial \mathbf{v}} = Q + H_{hv} + J_{el} + J_q + J_{ex},
\]  

(19)

where \( Q \) represents the productions, \( H_{hv} \) the spontaneous transition to another state—typically by light emission—\( J_{el} \) the loss due to elastic scattering (and therefore momentum transfer) (Lilensten et al., 2013), and \( J_q \) the loss due to quenching. Note the addition of an extra loss term, \( J_{ex} \), for charge exchange. The transport equation should be taken into account for all the species, and they can be coupled when the loss of one species creates another one. An example of that situation is the coupled transport between \( H \) and \( H^+ \), where a proton undergoing a charge exchange will become a fast \( H \), which can be re-ionized later. This equation is also valid for the excited state species, such as \( O(1S) \) and \( O(1D) \), that are notably responsible for the green line and the red line in aurorae (Gronoff, Simon Wedlund, Mertens & Lillis 2012; Gronoff, Simon Wedlund, Mertens, Barthélemy, et al., 2012).

In the following subsections, we review the main processes creating ENA/fast ions. Charge exchange is described in a later section. While the same equation should be solved to address atmospheric escape, approximations are often used for the coupled ion/ENA equations, angular diffusion, and upper atmospheric densities (Rahmati et al., 2018). These approximations are used for several reasons. One particularly problematic point in the simulations is the distance at which a particle is considered lost in space; some studies take a few planetary radii, other a few exospheric altitudes. Such approximations can create difficulties when comparing with observations (Baliukin et al., 2019).

### 2.2.1. Ion Recombination

An exothermic ion recombination (or chemical reaction) can give enough kinetic energy to one of its products so that it can escape. Ion recombination is the most effective channel to escape \( O \) in the present Martian atmosphere. It is, in general, an efficient way to heat up an atmosphere through nonthermal process. It is also a process leading to the escape of heavier atoms from light planets or bodies. The process has been largely studied in the past (Shizgal & Arkos, 1996) and is being refined in support of the MEX and MAVEN missions (Cipriani et al., 2007; Lillis et al., 2017; Valeille et al., 2010; Yagi et al., 2012; Zhao & Tian, 2015).

At Mars, the main photochemical escape process is the loss of oxygen through the reaction:

\[
O_2^+ + e^- \rightarrow O(^3P) + O(^3P) + 6.99eV \]  

(20)

\[
ightarrow O(^1D) + O(^1P) + 5.02eV \]  

(21)

\[
ightarrow O(^1D) + O(^1D) + 3.05eV \]  

(22)

\[
ightarrow O(^1S) + O(^3P) + 2.80eV \]  

(23)

\[
ightarrow O(^1D) + O(^1S) + 0.83eV \]  

(24)

Recent studies by MAVEN were able to show the hot oxygen corona produced by these recombination reactions (Deighan et al., 2015). A study by Cravens et al. (2017) shows that, in the limits of the current solar conditions at Mars, a linear dependence of the escape rate to the EUV flux can be made.

Another interesting reaction is \( N_2^+ + e^- \rightarrow 2N^+ \), which is efficient enough for the removal of \( ^{14}N \) but not \( ^{15}N \) at Mars and could explain the isotopic fractionation (Shizgal & Arkos, 1996).

At Earth and Mars, we also have (Gröller et al., 2014; Shizgal & Arkos, 1996) (the channel with \( O(^1D) \) has branching ratio close to 0)

\[
NO^+ + e^- \rightarrow N(^4S) + O(^1P) + 2.78eV \]  

(25)

\[
ightarrow N(^4D) + O(^1P) + 0.39eV \]  

(26)

\[
ightarrow N(^4S) + O(^1D) + 0.81eV \]  

(27)
To compute the photochemical escape through these processes, it is first necessary to compute the ion density. This involves, first, computing the ion productions (via photoionization, secondary electron ionization, etc.); second, computing the resulting chemistry and transport to get the ion densities; third, computing the hot atom production, using the densities and the reaction rate; and fourth, computing the actual escape by computing the transport of the hot atom. Such an escape should include collisions with other species; if the hot atom creation rate is important enough, it should be taken into account that these collisions heat up the upper atmosphere and therefore change its profile toward more escape.

Recent work at Mars shows that the CO$_2^+$ dissociative recombination is a nonnegligible source of hot oxygen (Lee et al., 2015; Zhao & Tian, 2015). In the following, the first reaction is believed to have a branching ratio between 96% and 100%:

\[
\text{CO}_2^+ + e^- \rightarrow \text{CO}(^1\Sigma) + \text{O}(^3P) + 8.27\text{eV} \\
\rightarrow \text{CO}(^1\Sigma) + 13.78\text{eV} \\
\rightarrow \text{C}(^3P) + \text{O}_2(^3\Sigma_g) + 2.29\text{eV} \\
\rightarrow \text{C}(^4P) + 2\text{O}(^3P) - 2.87\text{eV}
\]

\[
\text{CO}^+ + e^- \rightarrow \text{C}(^3P) + \text{O}(^3P) + 2.90\text{eV} \\
\rightarrow \text{C}(^1D) + \text{O}(^3P) + 1.64\text{eV} \\
\rightarrow \text{C}(^3P) + \text{O}(^1D) + 0.93\text{eV} \\
\rightarrow \text{C}(^1S) + \text{O}(^1P) + 0.22\text{eV} \\
\rightarrow \text{C}(^1D) + \text{O}(^1D) - 0.32\text{eV} \\
\rightarrow \text{C}(^3P) + \text{O}(^1S) - 1.28\text{eV}
\]

Hot oxygen in a planetary thermosphere can also induce escape of lower mass species by sputtering (Shizgal, 1999).

2.2.2. Photodissociation

Another process leading to the creation of fast ions or atoms is the direct dissociation by photon, electron, or proton impact.

In Shematovich et al. (1994) an example is given by the reaction O$_2$ $+$ hν $\rightarrow$ O$(^3P)$ + O$(^1P,^1D,^1S)$; the kinetic energy given to the products is the difference between the energy of the photon and the binding energy (i.e., the threshold energy for the reaction). Similar processes can be evaluated for N$_2$, CO$_2$, etc. Photodissociation reactions are seldom considered in evaluating escape rates since the production of fast enough particle to escape is small with respect to ion recombination processes. To properly evaluate these productions, it is necessary to have an accurate set of cross sections (see section 3.2.4). In general, thermospheric codes consider that the kinetic energy given by these photodissociations ends up in heating; therefore, one has to be careful to not count that loss of energy twice in their simulations.

2.2.3. Dication Dissociation

The dication dissociation effect on planetary escape has been proposed by Lilensten et al. (2013). It is a nonthermal processes that is based on the fact that the Coulombian dissociation of a molecular doubly charged ion may give enough energy to one or both of the ions to allow their escape.

The typical example for this process is CO$_2^+$ $\rightarrow$ CO$^+$+O$^+$, as described in Lilensten et al. (2013). Other processes such as N$_2^+$ $\rightarrow$ 2N$^+$ (Gronoff et al., 2007) or O$_2^+$ $\rightarrow$ 2O$^+$ (Gronoff et al., 2007; Simon et al., 2005) can
gives sufficient energies for the ion to escape. To account for the flux of escaping particles through that process, it is necessary to compute the transport of the fast ions from where they are created to the exobase. Since it is ions that are escaping, they are not necessarily escaping even if they reach the exobase with sufficient energy: The presence of magnetic fields could prevent their escape and return them into the atmosphere where they could create some additional heating (the Coulomb energy being in the range of several 10 eV, such ions could not efficiently sputter, except if they are further accelerated by the solar wind). A process not accounted for in the Lilensten et al. (2013) paper is the heating of the ionosphere and the creation of fast ENA through charge exchange of the fast ions with the atmosphere (a process similar to the one described in Chassefière, 1996a). On the contrary, ions with energy lower than the escape energy could escape due to electromagnetic forces, as will be explained in sections 2.4 and 2.5.

The calculation of the dication escape in a nonmagnetized atmosphere proceeds as follows: From escape due to electromagnetic forces, as will be explained in sections 2.4 and 2.5.

The ion loss mechanisms begin with the interaction of the upper atmosphere and ionosphere with the solar wind. Neutral atoms can be ionized by solar UV, charge exchange, and electron impact and can be scavenged by the solar wind. There are different processes and loss channels through which the planetary ions can escape to space, including pickup and sputtering, charge exchange, and outflow, which will have its dedicated subsection. Ion escape is believed to be one of the major sources of atmospheric escape in the current solar system and also at exoplanets around M dwarfs (García-Sage et al., 2017).

2.2.4. Key Parameters

Modeling photochemical loss requires the cross section for ionization by different processes (including elastic, inelastic, and charge exchange) and the chemical reaction rates for the density/recombination (including the branching ratio and the products speed probabilities). The ionospheric electron temperature is overall extremely important since the recombination cross section is likely to be extremely sensitive to it (Sakai et al., 2016). For the simulation of the ion/electron composition and temperature, it is necessary to perform a 3-D modeling of the ionosphere.

2.2.5. Questions

The evaluation of escape rates from photochemical reactions has mainly been done for solar system planets, especially Mars and Venus. Once we consider exoplanets or the Young solar system, questions remain about the efficiency of each processes. The ion recombination or the usually neglected processes such as particle impact dissociation could become more important when increase in XUV or precipitating particle flux occur. This question is difficult to answer since each process affects the state of the upper atmosphere and the efficiency of each other.

2.2.6. Observables

The recombination processes create ENA at very specific energies, typically in the 5-eV range. Since collisions occurs, changing the spectral shape of the energy distribution, the direct observation of these energies peaks is extremely challenging. Indirect techniques, based on modeling the hot oxygen corona, are used. At Mars, a technique to observe the product of photochemical reactions involved observing the hot oxygen geo-corona (Deighan et al., 2015). As explained in Shizgal and Arkos (1996), photochemical escape can explain the fractionation of $^{14}\text{N}/^{15}\text{N}$ at Mars. A more recent work from Mandt et al. (2015) shows that nonthermal processes except photodissociation can explain the isotopic enrichment. The work of Liang et al. (2007) shows that self-shielding effects can lead to an increase in heavier isotopes (here $^{15}\text{N}$ at Titan) escape from photodissociation.

2.3. Ion Loss

The ion loss mechanisms begin with the interaction of the upper atmosphere and ionosphere with the solar wind. Neutral atoms can be ionized by solar UV, charge exchange, and electron impact and can be scavenged by the solar wind. There are different processes and loss channels through which the planetary ions can escape to space, including pickup and sputtering, charge exchange, and outflow, which will have its dedicated subsection. Ion escape is believed to be one of the major sources of atmospheric escape in the current solar system and also at exoplanets around M dwarfs (García-Sage et al., 2017).
2.3.1. Pickup and Sputtering Escape

2.3.1.1. Pick-Up Escape

Pick up ion loss is due to the ionization of neutral constituents in the exosphere and upper atmosphere that sense an electric field and can be “picked up” and swept away. In the presence of the magnetic field, at Earth, for example, the polar wind drives pickup ion escape (Moore et al., 1997). At lower altitudes, this interaction can compress the magnetic field on the sunward side, forming a tail on the antisunward side. At high altitudes, the loss of H\(^+\), He\(^+\), and O\(^+\) can occur when thermal plasma originating from the polar regions in the ionosphere is accelerated into the magnetosphere and escapes downtail (Johnson et al., 2008). These processes will be detailed in sections 2.4 and 2.6.

At weakly magnetized planets, such as Mars and Venus, the lack of an intrinsic dipole magnetic field creates a scenario where the solar wind directly interacts with the upper atmosphere. In this situation, neutral constituents are ionized and picked up by the background convection electric field that is driven by the solar wind, where \( \vec{E}_{SW} = -\vec{U}_{SW} \times \vec{B}_{SW} \), where \( \vec{E}_{SW} \) is the electric field induced on an ion by the solar wind (and therefore that ion will be subject to a force \( \vec{F} = q(\vec{E}_{SW}) \)). \( \vec{U}_{SW} \) is the solar wind speed and \( \vec{B}_{SW} \) is the interplanetary magnetic field (IMF). The main channels for ionizing planetary neutrals are photoionization, charge exchange, and electron impact ionization. Curry et al. (2013) investigate these mechanisms as a function of solar zenith angle, bulk velocity, and plasma temperature, respectively, finding that the majority of pickup ions are formed in the corona and subsolar region of Mars. The origin of pickup ions plays a major role in their fate as escaping particles or precipitating particles (Fang et al., 2010). In the former case, the pickup ions can accelerate to twice the solar wind speed and their gyroradii are on the order of a planetary radius and are likely to escape. The maximum energy of a picked-up ion is \( E_{max} = 2mU_{SW}^2\sin^2(\theta_B) \), where \( \theta_B \) is the angle between the solar wind direction and the IMF (Rahmati et al., 2015). In the case of precipitating ions, the pickup ions will collide with neutrals in the exobase or thermosphere and transfer enough energy and momentum to the neutral that they could be able to exceed the escape velocity, a process known as sputtering. Ion precipitation also impacts the atmosphere through heating. The sputtering process can also happen at Earth, inside the polar regions (Shematovich et al., 2006), but it is a small process there.

2.3.1.2. Pickup Equations

If we consider \( n_{SW} \) as the solar wind density, \( n_0 \) the density of oxygen where that solar wind is located, \( \sigma_{CX} \) the average charge exchange cross section between the solar wind and oxygen, and \( \sigma_{PI}(\lambda) \) the photoionization cross section, we have an ion production of \( P_I = n_0 \int \sigma_{PI}(\lambda) \Phi_{EUV}(\lambda) d\lambda + \sigma_{CX} n_{SW} U_{SW} \) (and other ionization processes can be added such as electron impact) (Rahmati et al., 2015, 2017), that production is balanced by the pickup transport. If we consider \( P_I(\vec{v}) \) the production of ion at a speed defined by \( \vec{v} \) (so that \( \int P_I(\vec{v}) d\vec{v} = P_I \)), \( e \) the charge of the ion and \( m \) its mass, then the velocity distribution function \( f(\vec{x}, \vec{v}) \) for the picked-up ions is governed by (Hartle et al., 2011)

\[
\vec{E}_{SW} = -\vec{U}_{SW} \times \vec{B}_{SW} \tag{38}
\]

\[
\vec{v} \frac{\partial f}{\partial \vec{x}} + \frac{e}{m} (\vec{E}_{SW} + \vec{v} \times \vec{B}_{SW}) \frac{\partial f}{\partial \vec{v}} = P_I(\vec{v}) \tag{39}
\]

\[
\Phi(\vec{x}) = \int v f d\vec{v} \tag{40}
\]

with \( e \) the ion charge. Several techniques can be used for solving Equation 39; The complexity arises from the solar wind piling up around the planet (or the comet; Coates, 2004), creating complex magnetic field geometries. Typically, it has been solved using test particles (Monte Carlo simulations) in fields from magnetohydrodynamic (MHD) or self-consistent hybrid codes, as by Jarvinen and Kallio (2014).

2.3.1.3. Sputtering

The yield \( Y \) of sputtered neutrals is defined by the sputtering efficiency. This yield is the ratio of the number of escaping particles and the number of incident particles, which varies inversely with the planet's gravitational energy (Johnson, 1994; Johnson et al., 2008; Leblanc & Johnson, 2002). Sputtering is dependent on the incident particles' energy and angle of incidence, as well as the mass of the incident particle. For lighter incident pickup ions, the direct scattering of planetary neutrals is known as “knock-on,” which dominates at low,
grazing incidence angles. For heavier incident pickup ions, the additional momentum can create a cascade of collisions at high enough energies to cause a neutral to escape, where $Y \geq 1$ (Johnson et al., 2008; Leblanc & Johnson, 2001). This occurs for $O^+$ pickup ions at energies of approximately keV to approximately hundreds of keV. This is especially important when the pickup ion gyroradius is of the order of the planet radius, as at weakly magnetized bodies such as Mars, Venus, and Titan.

Sputtering is widely believed to be the dominant escape process at Mars and Venus during earlier epochs of our Sun, which has major implications for exoplanetary atmospheres. Luhmann et al. (1992) calculated the flux of precipitating pickup ions and ENAs using a 1-D exospheric model of the O density and a gas–dynamic model of the solar wind and found compared to pickup ion and photochemical escape, sputtering drove the highest rates of atmospheric erosion (see Figure 3). Other studies using MHD and hybrid models have found similar results (Chaufray et al., 2007; Wang et al., 2014). Sputtering as a dominant driver of atmospheric escape is further supported by current isotope ratios. Specifically, Ar is an important atmospheric tracer because once Ar is in the atmosphere, the only loss process is escape to space (as opposed to volcanic outgassing from the interior, impact delivery, and mixing with the crust), which limits the exchange pathways that become complicated for most planetary volatiles (Jakosky & Phillips, 2001). Thus, measurements of the present-day atmosphere reflect the importance of these exchanges over billions of years and emphasizes the need for understanding our own terrestrial planets’ atmospheric evolution as a ground truth for understanding exoplanetary atmospheres.

Unfortunately, sputtering is incredibly difficult to observe as the sputtered component at Mars is indiscernible from photochemically produced oxygen. Thus, models have typically predicted what the sputtered component would be in a variety of scenarios. The passage of the Siding Spring comet close to Mars (Bodewits et al., 2015) created a high flux of $O^+$ ions that impacted the atmosphere as predicted in (Gronoff et al., 2014). Observations by the MAVEN Solar Energetic Particle (SEP) instrument and the Mars Odyssey-High Energy Neutron Detector (HEND) indicate an increase in the $O^+$ pickup during the passage; however, an increase in solar activity at the same time prevents a clear conclusion on whether or not it was due to the comet (Sánchez-Cano et al., 2018). Wang et al. (2016) computed the impact of these ions on the escape rate and found that up to 10 t of atmosphere may have escaped, while 1 t of material was added. Another formulation of the escape by sputtering can be found in Shizgal (1999). In this paper, it is the sputtering by the hot oxygen of Venus that leads to the escape of H and D. The main difference with the usual pickup-sputtering process is the origin of the hot O, from the thermosphere itself, and therefore that the forcing by an external flux and the use of yield function cannot be applied. Shizgal (1999) developed a specific kinetic model for the escape.

### 2.3.1.4. Sputtering Equations

The rate of species $n_j$ escaping an atmosphere from sputtering is given by

$$\frac{\partial n_j}{\partial t} \approx 2\pi R_{\text{exo}}^2 < \Phi_a Y_j >$$

(Johnson, 1994), where $\Phi_a$ is the flux of the particle leading to the sputtering and $Y_j$ the sputtering efficiency for that peculiar species. For sputtering by an Incident Particle A and a Target Species B, and of respective masses $M_A$ and $M_B$, that is, a thermosphere whose main constituent is B, the yield can be computed as follows. First, for an incident particle of energy $E_A$, it is necessary to evaluate the elastic cross section $\sigma_d(E_A)$, which is related to the momentum transfer (or knock-on, elastic nuclear) stopping cross section $S_d(E_A)$ through:

$$\gamma = \frac{4M_AM_B}{(M_A + M_B)^2}$$

(41)

$$\sigma_d(E_A) = \frac{\gamma}{2} E_A S_d(E_A)$$

(42)
The overall yield is the result of single impact plus multiple impact momentum transfer at energy greater than the escape energy. It can be approximated by (Johnson, 1994; Johnson et al., 2000)

$$Y(\theta, E) \approx \frac{a_\beta \sigma(E)}{2U_\sigma \sigma_d(E) \cos \theta} \approx U_{es} \sigma(E) \cos \theta (43)$$

where $\theta$ is the incident angle, $U_{es}$ is the gravitational binding energy at the exobase, and $\sigma(E)$ is the average energy of the escaping particle. $a_\beta$, and $p$ are constants depending upon the impact particles; see Johnson (1994) for some numerical values in the literature. The sputtering yield may be enhanced by the sputtered particles that are picked-up and accelerated toward the atmosphere (equation 2 in Johnson, 1994). If the efficiency of escape for a sputtered particle is $Y_a$ and its ionization and return is $p$, then the effective yield is $Y_{eff} = Y / (1 - p(Y_a - 1))$.

### 2.3.1.5. Other Impact Processes

The classical sputtering process involves the impact of an ion that has been accelerated by pickup, that is, a nonthermal processes outside of the thermosphere. Gacesa et al. (2012) proposed a very similar mechanism where the impact of hot O from the Martian corona sputters light gases. Their computations suggest that it is the main channel for HD and D₂ direct escape. To validate that approach, it is suggested to observe the emission of H₂ ro-vibrationally excited by the impact. ENA impacts on the Martian atmosphere are also a source of escape, especially when they have been created by charge exchange from the solar wind (Lewkow & Kharchenko, 2014), which leads us the other class of escape processes.

### 2.3.2. Charge Exchange of a Magnetically Trapped Particle

The basic idea of charge exchange escape is that a magnetically trapped energetic ion, such as H⁺, exchanges its charge and becomes an ENA that can escape or sputter (an ion trapped in a magnetic mirror may be prevented to reach the thermosphere and therefore to efficiently sputter) (Shizgal & Arkos, 1996). The temperature dependence is complex: At Earth it decreases with exospheric temperature for H (Shizgal & Arkos, 1996) so that the escaping flux from charge exchange plus Jeans escape is constant, reaching the diffusion-limited value.

A simple approach adopted by Yung et al. (1989) was to consider that the ion H⁺ had a Maxwellian distribution at the temperature $T_{ion}$. Having exchanged its charge, the ENA could escape, and it would have the same energy as the initial ion. The efficiency for an escaping charge-exchanged atom, with respect to the temperature of the initial ion is

$$\alpha_i(R) = B_{CX} \left[ 1 - \frac{\nu_{ex}^2}{u_i(R)^2} \right] e^{-\frac{u_i}{2kT_{ion}}} (45)$$

with $u_i(R) = \sqrt{2kT_{ion}(R)/m}$. Considering $k_i \rightarrow j$ the charge exchange rate between the ionized species $i$ and a neutral species $j$ (rate that can vary with temperature), this gives the escape flux

$$\varphi = \int_{R_{ex}}^{R_{p}} \frac{R}{2\pi k} \alpha_i(R) \sum_{k_i \rightarrow j} [k_i] [j] dR. (46)$$

The $B_{CX}$ factor in the definition of $\alpha_i$ is an efficiency factor, which was taken identical to the one for thermal escape in the Yung et al. (1989) paper. The rest of the equation is similar to the thermal escape equation, except the factor which was taken off for considering it is hidden in the charge exchange rate.

The equation in Yung et al. (1989) paper has a negative sign that should be positive: Using the equation with that negative sign leads to negative escape fluxes. Using that equation, it happens that the charge exchange flux should increase with increasing exospheric temperature, which is not what is observed. It means that this simplified approach is not good enough for evaluating the charge exchange flux at Earth.
Shizgal and Lindenfeld (1982) developed a collisional model for computing the charge exchange-induced escape. The main difference with the previous approach is that the efficiency of charge exchange with respect to the temperatures is taken into account following Fitzpatrick and Shizgal (1975). It is shown that the charge exchange is, at Earth, the most efficient mechanism to remove H from the upper atmosphere during low solar activity (low exospheric temperature), while Jeans’ escape is the main mechanism during high solar activity. It is important to remember here that H escape is diffusion limited at Earth. In the following, A corresponds to the neutral atmosphere (O and H), \( n \) to the average density (of A, O, and H+) over the region of charge exchange, and \( \sigma \) corresponds to the energy independent hard sphere cross section, and:

\[
\lambda_{\text{CX}} = \frac{m_{\text{H}} v_{\text{esc}}^2}{2 k T_{\text{ion}}} \quad (47)
\]

\[
\dot{n} = \sigma_{H^+} \frac{[n_{H^+} n_A]}{n_{H^+} n_O} \Gamma(a) 1 + a \quad (48)
\]

\[
\tau_{\text{CX}} = \frac{T_{\text{exo}}}{T_{H^+}} - 1 \quad (49)
\]

\[
\Phi_{\text{CX(escape)}} = \dot{n} \sqrt{\frac{2 k T_{\text{esc}} e^{-\lambda_{\text{CX}}}}{\pi m_{\text{H}} \tau_{\text{CX}}}} \times \left[ 1 + \frac{\tau_{\text{CX}}}{\sqrt{1 + \tau_{\text{CX}} e^{-\lambda_{\text{CX}} \tau_{\text{CX}}}}} \right] \quad (50)
\]

This equation is valid for the escape of H at Earth from charge exchange. It supposes that (1) the H+ density varies slowly with altitude at the location where this process is the most efficient (from the exobase to 3,000 km), (2) the only species interacting are H, O, and H+ , and (3) the distributions are Maxwellian, with a fixed temperature in the altitude range.

At Earth, the charge exchange is the main mechanism to remove O+ from the ring current (Daglis et al., 1999). The exchange creates ENA that can be imaged to study the ring current evolution.

### 2.3.3. Charge Exchange With the Solar Wind

The charge exchange between the solar wind and the upper atmospheric species can enhance the escape rate through pickup like processes: a species M exchanges its charge with, for example, a proton from the solar wind H+\* + M → H+ + M+. The solar wind proton becomes an ENA and can create additional heating that increases the thermospheric temperature and therefore escape (Chassefière, 1996a). The created ion can escape thanks to pickup by the magnetic field.

At comets, charge transfer reactions primarily involve solar wind ions, H+ and He2+ and also multiply-charged minor species such as O6+, Si10+, or C5+ (Bodewits, 2007; Cravens, 1997; Simon Wedlund et al., 2016), with water molecules continuously outgassing upon sublimation from the nucleus. As the atmosphere of a comet is in expansion, charge-transfer reactions take place over a large region of space (of the order of several 10^6 km) and will have time to facilitate the absorption of the solar wind, converting fast ions into slow-moving ones. Charge transfer has recently been evidenced by the ESA/Rosetta ion spectrometers at comet 67P/Churyumov-Gerasimenko (67P), with the observation of H+ ions (Burch et al., 2015), and He+ fast ions (Nilsson et al., 2015). The latter charge-exchanged ions, originating from solar wind He2+ ions (composing about 4% of the bulk of the undisturbed solar wind), were present throughout the mission from a heliocentric distance ranging from 3.4 to 2 AU (Simon Wedlund et al., 2016; Simon Wedlund, Behar, Kallio, et al. 2019; Simon Wedlund, Bodewits, et al., 2019; Simon Wedlund, Behar, et al., 2019). The net effect of the charge transfer of He2+ solar wind ions with the neutral atmosphere of the comet (composed of molecules M) is the production of ENAs following the typical sequence of electron capture reactions (double charge transfer, and stripping reactions are ignored here for simplicity):

\[
\text{He}^2+ + \text{M} \rightarrow \text{He}^+ + \text{M}^+ \quad (51)
\]

\[
\text{He}^+ + \text{M} \rightarrow \text{He} + \text{M}^+. \quad (52)
\]
backscattered) ions and ENAs will coexist at any given altitude above the atmosphere, due to the nonuniformity of the magnetic field and to collisions, for an ENA, $X^+$, and its corresponding ion, $X^+$. In the following, $I$, the intensity, $\mathcal{P}$, the momentum transfer, $\mathcal{R}^{\text{CX}}$, the charge transfer, and, $\mathcal{Q}$, the local production, depend upon $(r, E, \mu)$. The transport of ENAs, denoted $X$, is as follows:

$$\frac{\partial I_X}{\partial t} = -I_X + \sum k_{\text{elas}}(E) n_k(z)$$

$$\frac{\partial I_{X^+}}{\partial t} = -I_{X^+} + \sum k_{\text{elas}}(E) n_2(z)$$

Numerical solutions of this system have historically made use of continuous slowing-down approximations (Decker et al., 1996), DSMC techniques (Basu et al., 2001; Shematovich et al., 2011), and a semianalytical exponential matrix solution (both with dissipative forces and angular redistributions Galand et al., 1997, 1998), including angular redistributions due to the nonuniformity of the magnetic field and to collisions, for an ENA, $X$, and its corresponding ion, $X^+$. In the following, $I$, the intensity, $\mathcal{P}$, the momentum transfer, $\mathcal{R}^{\text{CX}}$, the charge transfer, and, $\mathcal{Q}$, the local production, depend upon $(r, E, \mu)$. The transport of ENAs, denoted $X$, is as follows:

$$\frac{\partial I_X}{\partial t} = -I_X + \sum k_{\text{elas}}(E) n_k(z)$$

$$\frac{\partial I_{X^+}}{\partial t} = -I_{X^+} + \sum k_{\text{elas}}(E) n_2(z)$$

Similar equations can also be derived for the coupled ($H^+, H$) system. These processes lead to the almost total conversion of the solar wind into ENAs, potentially escaping or sputtering the nucleus, by the time the solar wind impinges within a few tens of kilometers from the comet’s surface, in the case of a highly outgassing nucleus (perihielion conditions). This total conversion depends on many parameters: outgassing rate, heliocentric distance, solar wind density, and speed (Simon Wedlund, Behar, Kallio, et al. 2019; Simon Wedlund, Bodewits, et al., 2019; Simon Wedlund, Behar, et al., 2019). The effect of minor solar wind species (multiply-charged heavy ions) can be seen in the production of X-rays through charge exchange emission with the cometary atmosphere (Cravens, 1997). The case of comets provides a unique opportunity to study charge-exchange processes within different and varying atmospheric environments.

The observation of escape from HD 209458 has been interpreted as increased by charge exchange processes between the solar wind and the hydrogen from the upper atmosphere of the planet (Holmström et al., 2008).

### 2.3.4. Charge Exchange With a Precipitating Particle

Particles precipitating in the atmosphere of planets can give rise, through charge exchange with the ambient neutral atmosphere, to the local production of ENAs. This is particularly significant at Earth in the case of protons of solar wind origin, first accelerated in the magnetosphere and then precipitating down the magnetic field lines in the polar regions. When protons are neutralized in collisions with neutrals (mainly oxygen atoms above 200-km altitude, $O_2$, and $N_2$ below), a process referred to as electron capture, the newly produced hydrogen ENAs, not being sensitive to the magnetic field, travels in straight trajectories, whose direction is related to the pitch angle distribution of the impinging protons, resulting in a horizontal spreading of the precipitating beam. Hydrogen ENAs, keeping most of the kinetic energy of the incoming proton, can in turn be ionized (electron stripping). Due to magnetic mirroring and angular redistributions stemming from collisions between the energetic species and the atmosphere, downwellng (or precipitating) and upwelling (or backscattered) ions and ENAs will coexist at any given altitude above the Earth region peak of the initial proton energy deposition (for a 10-keV initial proton peak will occur at 120-km altitude at Earth). The energy and angular degradation of a ($H^+, H$) beam in the atmosphere is usually formalized as a coupled system of two nonlinear Boltzmann transport equations (Galand et al., 1997, 1998), including angular redistributions due to the nonuniformity of the magnetic field and to collisions, for an ENA, $X$, and its corresponding ion, $X^+$. In the following, $I$, the intensity, $\mathcal{P}$, the momentum transfer, $\mathcal{R}^{\text{CX}}$, the charge transfer, and, $\mathcal{Q}$, the local production, depend upon $(r, E, \mu)$. The transport of ENAs, denoted $X$, is as follows:

$$\frac{\partial I_X}{\partial t} = -I_X + \sum k_{\text{elas}}(E) n_k(z)$$

$$\frac{\partial I_{X^+}}{\partial t} = -I_{X^+} + \sum k_{\text{elas}}(E) n_2(z)$$

Motivated by the MEX and MAVEN missions, there are an increasing number of studies of proton precipitation at Mars. Shematovich et al. (2011) have developed a DSMC model of the coupled ($H^+, H$) system in a (CO$_2$, N$_2$, and O) atmosphere and applied it to MEX ASPERA data in solar minimum conditions. They concluded that about 20% (10%) of the incoming particle (energy) flux was backscattered by the atmosphere and emphasized the role of the solar wind magnetic field pile-up region at altitudes above 100 km in increasing...
the backscattered flux by a factor up to 50%. Shematovich (2017) recently studied the production of suprathermal O atoms in Mars’ thermosphere via this process and concluded that a hot oxygen corona may form, creating an additional nonthermal escape flux of O that may become prevalent when extreme solar transient events, such as flares and coronal mass ejections (CMEs), take place. Finally, Halekas (2017) derived the ENA flux originating from the solar wind interaction with the Martian atmosphere from the observation of protons by MAVEN/Solar Wind Ion Analyzer (SWIA). From there, it was possible to retrieve the exospheric temperature of Mars (as well as the solar wind velocity).

At Jupiter, energetic precipitation involves protons (Bisikalo et al., 1996) and also singly or multiply-charged heavy ions such as \( \text{S}^{n+} \) and \( \text{O}^{n+} \) (with \( n \) the charge number) (Horanyi et al., 1988), colliding with \( \text{H} \) and \( \text{H}_2 \) (Waite & Lummerzheim, 2002). The high charged states of O at very high energies (above 200 keV/amu) are responsible for auroral X-ray emissions, as modeled in Cravens et al. (1995) and compared to X-ray observations of Jupiter. Such ion precipitation creating fast energetic atoms is also expected to play a role for satellites of Jupiter, and at Saturn, and its satellites.

### 2.3.5. Charge Exchange in the Ionosphere

This process is an hybrid between charge exchange and photochemical escape; it consists of having excess kinetic energy when a charge exchange is performed, such as \( \text{He}^{3+} + \text{N}_2 \to \text{He}^{(+9\text{eV})}+ \text{N}_2^+ \) (Shizgal & Arkos, 1996). This process has been suggested to address the problem of the He budget in the Earth’s thermosphere (Lie-Svendsen et al., 1992; Shizgal & Arkos, 1996).

#### 2.3.6. Key Parameters

The important parameters in the computation of the pickup/sputtering are the solar wind parameters, which can usually be found thanks to models of the interaction of the solar wind with the planet (Curry et al., 2013; Lee et al., 2017) and the cross sections for ionization and stopping power/elastic scattering. In addition, it is necessary to have good inelastic/interaction potential (Johnson, 1994) cross sections to be able to compute the \( \alpha \), \( \beta \) and \( p \) parameters in Equation 43. Finally, a particular attention should be given toward the nature of the model with respect to modifying the inputs of ion pickup models: For example, it is usually assumed that an exosphere is present in a MHD model, and mass loading will reduce the accuracy of the model. Hybrid modeling will better improve such models, as is done in a cometary environment (Simon Wedlund et al., 2017). A review of the comparative advantages/drawbacks of each type of solar wind models can be found in Ledvina et al. (2008).

For the majority of the recent work in pickup and sputtering, many cross sections are being used without being published, which is a major problem for the community. The state-of-the-art models for sputtering are now using a DSMC approach (Johnson et al., 2000).

For charge-exchange processes, in addition to the particle precipitation models and the solar wind models, it is important to have a good knowledge of the atmosphere composition and temperature, including the ion temperature.

#### 2.3.7. Questions

How much do these processes scale up with the solar wind density, speed, and orientation? How does the creation of an induced magnetic field influence these charge exchange processes?

#### 2.3.8. Observables

##### 2.3.8.1. Composition Change

The observation of the change in solar wind composition is a proof of charge exchange, for example, at comets (Simon Wedlund, Behar, Kallio, et al. 2019; Simon Wedlund, Bodewits, et al., 2019; Simon Wedlund, Behar, et al., 2019). At Mars, the charge exchange of solar wind protons at the bow shock leads to precipitation of H that can be observed by the effects on the chemistry and by the backscatter (Halekas, 2017), even if the H chemistry at Mars is complex (Chaffin et al., 2017) One more striking example of charge-exchange processes at Mars is the observation of heavier ions, such as \( \text{O}^+ \), that later lead to sputtering (Leblanc et al., 2015, 2018).

##### 2.3.8.2. Fractionation Due to Pickup/Sputtering

The fractionation due to pickup and sputtering is efficient because of its tendency to make the species at the top of the thermosphere escape. Since isotopes have a gravitational fractionation at these altitudes, the overall effect is to increase the number of heavier species in the atmosphere. This is known as a Rayleigh distillation (see section 3.4.2.2).
2.4. Ionospheric Outflow

Heating and energization of electrons and ions at a magnetized planet results in escape of ionospheric plasma, either onto open field lines where it joins the solar wind flow and is lost to interplanetary space or onto closed or reconnecting magnetic field lines where it becomes trapped in the magnetosphere and becomes subject to magnetospheric dynamics and loss processes. The escape of ionospheric plasma is often considered in the context of magnetospheric dynamics and as a competing source of magnetospheric plasma together with the solar wind. However, it also has a vital role in the context of atmospheric escape and evolution in that a charged particle has additional plasma physics processes acting on it, as compared to a neutral particle, which does not respond to the magnetic or electric field. These processes help reduce the gravitational potential barrier binding the charged particle to the planet.

The escape of ionized particles to space has several names in the literature, ion outflow, polar wind, bulk ion escape, polar outflow, etc. This leads to some confusion as sometimes authors are generically referring to escaping plasma, but other times they are talking about outflow energized by particular processes that vary in space and time, as shown in Figure 4. For instance, the “polar wind” typically refers to the supersonic outflow of ions from the polar ionosphere accelerated by ambipolar electric fields (Axford, 1968; Banks & Holzer, 1968). As the name implies, this polar wind is similar in concept to the solar wind, the supersonic expansion of the solar corona into space, proposed by Parker (1958) nearly a decade before. While outflows of polar wind were initially thought to contain only light species such as protons, the first quantitative observations of O⁺ in the polar wind by the Retarding Ion Mass Spectrometer on-board the Dynamics Explorer 1 (DE-1) demonstrated that heavy ions can be present as well in quite significant numbers. O⁺ accelerated by wave-particle interactions in the cusp is sometimes referred to as the “cleft ion fountain,” while the same process above the auroral region is occasionally referred to as an “auroral wind.” The variability in location, composition, and energy of outflowing ions at Earth has led to the variety of names that describe escape along magnetic field lines. In this section, we eschew these more specific terms instead will use the term ionospheric outflow or ion outflow with the more broad meaning of any population of plasma upflowing from the planet at high altitude.

When thinking about what drives ionospheric outflows, it is instructive to consider the types of energy input. These break down into two broad categories as outlined in Strangeway et al. (2005) (see Figure 5): (1) particle and (2) electromagnetic energy input from the magnetosphere. Both downward Poynting flux and soft electron precipitation from the magnetosphere were shown to correlate very well with outflow of ions observed by the Fast Auroral Snapshot (FAST) Explorer spacecraft. While correlation is not the same as causation, it so happens that there are a number of causal mechanisms associated with each type of energy input:

1. Particle: suprathermal electrons (photoelectrons, auroral electrons, secondary electrons,...) enhancing the ambipolar electric field and depositing energy to the thermal electron population.
2. Electrodynamical: Transverse heating of ions as a result of wave-particle interactions, ponderomotive forcing from Alfvén waves, field-aligned currents (FACs) driving Eᵋ, low-altitude frictional heating driving upwelling, centrifugal force due to field line convection, and curvature change and/or magnetic field rotation with the planet.
The varied timescales and spatial regions over which these processes act result in dynamic outflow that varies spatially. At lower altitudes, the influence of different drivers separates the upflowing plasma into what has been called Type 1 and Type 2 outflow (Wahlund et al., 1992), where Type 1 involves strong electric fields and Joule heating, and Type 2 involves particle precipitation and enhanced electron temperatures. At high altitudes, the escaping plasma also exhibits temporal and spatial variability. The polar region at Earth typically contains lower energy polar wind outflow, whereas additional energization, particularly from wave particle interactions, results in an energetic ion outflow and preferential acceleration of heavy ions in the auroral and cusp regions. Figure 5 shows the different pathways to ion outflow and some of the unknown.

2.4.1. Suprathermal Electron Effect

Suprathermal electrons refer to electrons whose mean energy is much greater than the thermal energy. The source of these electrons is from XUV light shining on the atmosphere creating photoelectrons, precipitating electrons of magnetospheric origin (auroral electrons), or secondary electrons formed by impact ionization of the neutral atmosphere. This population is known to alter the ion outflow solution through two main processes:

1. Formation of the self-consistent ambipolar electric field.
2. Coulomb collisions between the superthermal and thermal electrons raising $T_e$.

Relative to ions, suprathermal electrons are unbound by gravity and in absence of any other process would escape. However, this would lead to a net charge in the plasma violating the quasi-neutrality condition. Therefore, an electric field forms that retards the electrons and accelerates the ions, reducing the gravitational potential barrier. Another pathway through which these electrons influence the outflow is through the deposition of energy to the thermal electrons raising the electron temperature and eventually the ion temperature.

Photoelectrons, formed from ionization of the atmosphere by solar/stellar radiation, have been particularly well studied in the context of ionospheric outflows. There have a large number of theoretical studies (Khazanov et al., 1997; Lemaire, 1972; Su et al., 1998; Tam et al., 1995, 1998; Wilson et al., 1997) and observational studies (Kitamura et al., 2011; Lee et al., 1980; Peterson et al., 2008) showing that this population is critical to setting up the quiet time outflow solution.

2.4.2. Electrodynamic Energy Input

Waves also play an important role in the acceleration of plasma in the high-latitude, high-altitude polar region. They do this primarily through two mechanisms: the ponderomotive forces of Alfvén waves (Guglielmi et al., 1996; Khazanov et al., 1998, 2000, 2004; Li & Temerin, 1993) and wave heating (Barghouthi, 1997; Bouhram et al., 2003; Crew et al., 1990; Retterer et al., 1987; Waara et al., 2011).

Ponderomotive forcing due to low-frequency electromagnetic waves allows electromagnetic energy from the magnetosphere to transfer energy to the ionospheric plasma. It arises from a nonresonant interaction between the particle encountering different portions of the wave during different parts of the particle gyration. The ponderomotive forcing depends on the wave mode, propagation direction, frequency, and background fields. While there are several types of ponderomotive force derived in the literature, a useful description of the total field-aligned force $F_\parallel$ from Alfvén waves as given by Lundin and Guglielmi (2006) is

$$
F_\parallel = -\frac{m c^2}{2 B^2} \left[ E^2 \frac{\partial B}{\partial z} - \frac{1}{2} \frac{\partial}{\partial z} \left( \frac{1}{c_A^2} \frac{\partial}{\partial t} + \nu \right) E^2 \right],
$$

where $m$ is the mass, $c_A$ the Alfvén speed, and $\nu$ the collision frequency. $E$ the electric field and $B$ the magnetic field.
Although the upward ponderomotive acceleration of ions is not species dependent, it is countered by a downward force on electrons, resulting in a downward ambipolar field and a resulting species-dependent reduction to the acceleration (Miller et al., 1995).

In contrast, wave heating arises from the resonant interaction of particles with the portion of the turbulent wave spectrum that corresponds to the cyclotron motion of the particle. This preferentially heats the ions perpendicularly to the magnetic field. The mirror force converts this excess perpendicular energy into organized parallel motion. When modeling this interaction, the wave heating is often represented as a diffusion term on the right-hand side of the Boltzmann equation having a form like (Crew & Chang, 1985):

\[
\frac{1}{v_L} \frac{\partial}{\partial v_L} \left( v_L D_L \frac{\partial f}{\partial v_L} \right),
\]

where \( f \) is the velocity space distribution function, \( v_L \) is the perpendicular velocity, and \( D_L \) is a diffusion coefficient. The diffusion coefficient can be written approximately as (Crew et al., 1990)

\[
D_L = \left( \frac{\pi q^2}{2m} \right) |E_L|^2 (\Omega(l)),
\]

where \( |E_L|^2 \) is the electric field spectral density of left-hand-polarized waves and \( \Omega(l) \) is the gyrofrequency of an ion of mass \( m \) and charge \( q \) at position \( l \) along a field line. Clearly, this term acts to add energy to the ions transverse motion around the field increasing the first adiabatic invariant and enhancing the mirror force which accelerates the ion.

Resonant wave heating has a clear signature in the shape of the ion distribution function. When the wave heating is active, the distribution function becomes increasingly perpendicular and pancake shaped. The mirror force, which acts more strongly on particles with higher perpendicular velocity, causes the distribution to “fold” upward into a characteristic “V” shape. In three-dimensional velocity space this looks like a cone and hence the name “conic” distribution. The observation of a conic distribution is a clear signature of the presence of resonant wave-particle interactions. Figure 6, from Bouhram et al. (2004), shows three examples of this feature observed by different satellites.

### 2.4.3. FACs

FACs, (also known as “Birkeland currents”), driven by a planet’s magnetic interaction with the stellar wind, are another form of electromagnetic energy input that contributes to ionospheric escape. This process was looked at by Gombosi and Nagy (1989), who found that including a FAC causes the thermal electrons to respond by possibly enhancing the ambipolar electric field. More generally, a current conservation equation can be defined as follows (Glocer, 2016):

\[
n_e u_e + n_u u_u - \sum n_i u_i = -\frac{j}{en},
\]

which states that the current density, \( j \), must be equal to the difference between the flux of electrons (thermal and suprathermal) and the flux of ions. If a large current is driven into the ionosphere, then this condition requires other populations to react.

### 2.4.4. Joule Heating

Joule heating (see also section 3.2.3) refers to the frictional heating caused by the differential motion of ions being dragged through the neutral atmosphere. In this process, the planet’s magnetic field interaction with the stellar wind generates a cross polar cap potential, which sets up magnetospheric convection as well as strong convective flows in the \( E \) and \( F \) regions of the polar ionosphere. This convective flow is generated by the \( E \times B \) drift and is not felt directly by the neutral population. As a result there is a differential motion between the ions and the electrons. There are several presentations of Joule heating as described by Strangeway (2012), but fundamentally, the most direct way to model this process is as a frictional heating term. This term can be presented based on Burger’s fully linear approximation (Burgers, 1969) as

\[
J = \frac{\sum \rho_i v_i}{m_i + m_j} \left[ m_j (u_i - u_j)^2 \right]
\]
Gombosi and Killeen (1987) and Cannata et al. (1988) examined the role of Joule heating and found that transient upflows can result from this process. Centrifugal forces play a role both at Earth (Horwitz et al., 1994) and at Jupiter (Nagy et al., 1986), but the origin of the centrifugal forcing is different for the two planets. At Earth, the solar wind connected field lines convect across the high-altitude polar cap region, resulting in changes to the field line curvature that centrifugally accelerate the particles outward along the magnetic field. At Jupiter, solar wind-driven convection plays a less important role, but the rapid rotation of the planet results in outward acceleration at lower latitudes.

2.4.5. Escape Equations

There are several types of methods for modeling ionospheric outflows, but they can generally be divided into two categories: hydrodynamic models and kinetic models. In the case of hydrodynamic models usually a multimoment expansion of the Boltzmann equation for each ion species is undertaken. For magnetized planets, this is taken in the low $\beta$ limit where the magnetic field is strong. In this case the gyrotropic five-moment equations with heat flux along an expanding magnetic field are given by (Gombosi & Nagy, 1989)

$$\frac{\partial}{\partial t}(\rho_i) + \frac{\partial}{\partial r}(\rho_i u_i) = A S_i$$  \hspace{1cm} (60)

$$\frac{\partial}{\partial t}(\rho_i u_i) + \frac{\partial}{\partial r}(\rho_i u_i^2) + A \frac{\partial p}{\partial r} = A \rho_i \left( \frac{e}{m_i} E_i - g \right) + A \frac{\delta M_i}{\delta t} + A u_i S_i$$  \hspace{1cm} (61)

$$\frac{\partial}{\partial t} \left( \frac{1}{2} \rho_i u_i^2 + \frac{1}{\gamma_i - 1} \rho_i p_i \right) + \frac{\partial}{\partial r} \left( \frac{1}{2} A \rho_i u_i^2 + \frac{\gamma_i}{\gamma_i - 1} A u_i p_i \right)$$

$$= A \rho_i u_i \left( \frac{e}{m_i} E_i - g \right) + \frac{\partial}{\partial r} \left( A \kappa \frac{\delta T_i}{\delta r} \right) + A \frac{\delta E_i}{\delta t} + A u_i \frac{\delta M_i}{\delta t} + \frac{1}{2} A u_i^2 S_i.$$  \hspace{1cm} (62)

In this case $m$ refers to mass, $\rho$ is the mass density, $u$ refers to the velocity, $T$ to the temperature, $p$ to the pressure, $r$ to the distance along the field line, and $e$ is the charge of an electron. The subscriptions denote ion species or electron. Other terms include the expanding cross-sectional area of the flux tube ($A$), the heat conductivity ($\kappa$), the specific heat ratio $\gamma_i$ and Boltzmann’s constant ($k$). The electric field ($E_i$) is derived as an Ohm’s law from a the steady state electron momentum equation as

$$E_i = -\frac{1}{en_i} \left[ \frac{\partial}{\partial r} (p_e + \rho_e u_e^2) + A \frac{\delta \rho_e}{\delta r} u_e^2 \right] + \frac{1}{en_i} \frac{\partial}{\partial r} \left( \frac{m_e}{m_i} (u_e - u_i) + \frac{\Sigma}{\delta t} \right)$$  \hspace{1cm} (63)
The source on the right-hand side of Equations 60–62 represent the source due to ion production and loss (S), the source due to momentum transfer \( \left( \frac{\delta M}{\delta t} \right) \), and the source due to energy transfer \( \left( \frac{\delta E}{\delta t} \right) \), which includes Joule heating effects. There are several derivations of the collision-based source terms for momentum and energy transfer, but a common choice is determined by using Burgers’ fully linear approximation (Burgers, 1969). Specific expressions for these terms can be found in the textbook by Schunk and Nagy (2004). The equations outlined above are used in the Polar Wind Outflow Model (PWOM) (Glocer et al., 2009, 2013). However, other codes use different hydrodynamics expansions including higher moment approximations. For example, the model presented by Varney et al. (2014) uses the eight-moment approximation, while Barakat and Schunk (1982) use the 16-moment approximation. We do not elaborate further on these approach here but refer the interested reader to those papers.

Kinetic solutions to ionospheric outflow typically solve the Boltzmann equation in some approximation. In the steady state this equation is given by Khazanov et al. (1997) and Khazanov (2010):

\[
\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{1}{2} \mathbf{B} \cdot \frac{\partial B}{\partial \mathbf{r}} \frac{\partial f}{\partial \mathbf{A}} = -\left( \frac{e}{m_e} \mathbf{E} \right) \cdot \frac{\partial f}{\partial \mathbf{v}} + \frac{1}{v} \frac{\partial^2 f}{\partial \mu^2}.
\]

Solving this equation usually takes one of two forms. Khazanov et al. (1997) use a direct solution of the Vlasov equation along a field line, whereas Barakat and Schunk (2006) use a 3-D macroscopic particle-in-cell Monte Carlo technique. (Depending on the study, the time dependent or the static case is solved.)

2.4.6. Questions

The PWOM (Glocer et al., 2009) was applied to exoplanets (Airapetian et al., 2017) with the study of the atmospheric escape of a Earth-like planet at Proxima b (Garcia-Sage et al., 2017). These studies were able to show that large escape rates happen even in the presence of magnetic fields. This model has the advantage of taking into account the diffusion of ions, assuring that this escape is not limited by ionospheric modeling. The questions remaining for this process is how it evolves in the absence of a permanent magnetic field, that is, when there is only an induced magnetic field. Studies by Collinson et al. (2019) tend to indicate that ambipolar diffusion helps the ion escape at Venus, but it remains to be modeled exhaustively. The long-lived doubly charged ions, observed/predicted in several ionospheres (Gronoff et al., 2007; Lilensten et al., 2005; Simon et al., 2005; Thissen et al., 2011), are easier to lift and therefore to escape through these processes. Are the O\(^{2+}\) observed by ISEE (Horwitz, 1981) due to these processes, and are they a significant source of escape? The dependence upon the EUV flux in certain situations (Young et al., 1982) is consistent with such an hypothesis.

2.4.7. Observables

As explained, the solar wind is directly observed by plasma instruments. The dependence of the escape efficiency upon the q/m ratio means that lighter isotopes are easier to lift, therefore enhancing the gravitational distillation of the ions (which can be affected by the self-shielding effect section 2.2.6).

2.5. Other Ion Escape

While the main sources of losses could be linked to the previously cited ones, other ion escape mechanisms have been reported in the literature. They mainly come from the observation of “bulk” ion escape at Mars or Venus, during specific solar conditions (Halekas et al., 2016). This general denomination groups together escape that could come from very different processes while leading to loss of ions in an organized way. Overall, it is transfer of momentum from the solar wind to the ionosphere that makes these plasma escape, and it could be considered as ion pickup in a first approximation, but with more complex MHD effects. As explained in Terada et al. (2002), the main problem is that the production of ions above the ionopause is less than the escaping flux, which means that different processes have to diffuse these ions from the ionosphere to above the ionopause where they would be picked up.

This is this second class of processes that comes from complex interactions between the ionosphere and the IMF to help transfer ions to the top of the thermosphere, which are studied here. These ions are usually not energetic enough to escape, but their presence makes it easier for pickup and “bulk” escape. It is to be noted that, when conditions are extreme, it is possible to reach levels where the production/diffusion of ions is
limiting the escape. Numerous models of ion pickup do not take that situation into account, leading up to unrealistic ion escape fluxes (Egan et al., 2019).

In magnetospheres, similar problems arise: The different ion energizing processes lead to the creation of a plasmasphere, that is, ions trapped inside the magnetosphere, and these ions are removed either by falling back to the planet or leaving through different processes (Jackman et al., 2014; Seki et al., 2015).

### 2.5.1. Fluid Processes: Kelvin-Helmholtz and Other Instabilities

Kelvin-Helmholtz instabilities (KHIs) (Johnson et al., 2014) have been observed at Mercury (Sundberg et al., 2010), Venus (Lammer et al., 2006; Pope et al., 2009; Terada et al., 2002, and references therein: the observations of flux rope and detached plasma clouds are linked to KHI through modeling), Earth (Johnson et al., 2014), and Mars (Ruhunusiri et al., 2016). They occur at the interface between two fluids or plasmas having a velocity shear and lead to upward pressure gradients and the formation of a vortex. KHIs have effects in a variety of planetary processes. In the case that interest us, that is, the development of KHI at the interface between an ionosphere and the solar wind, it leads to the transfer of solar wind momentum to the ion and, ultimately, to their acceleration into space. (This is different for situations when magnetic fields through reconnection mean that a part of the initial magnetosphere is no more linked to the magnetic field of the planet and is ejected into space, along with the plasma it contains. Ejection of plasmoids is also associated to a return of plasma toward the planet. (This is different for situations when magnetic fields are weaker, such as Mars.) The equivalent of these plasmoids has been observed at Mars near the magnetospheres created by the crustal magnetic field; they are usually named “flux ropes” (Hara et al., 2017) and are linked with CME disturbances and other processes (Hara et al., 2017). They could be responsible for up to 10% of the present-day ion escape at Mars (Brain et al., 2010). Finally, plasmoid escape has also been observed at Venus, in the induced magnetic field, and looks like the Earth’s or Jupiter’s plasmoids (Zhang et al., 2012).

### 2.5.2. Pickup Processes: The Ion Plume

The ion plume of Mars was inferred from MEX observations and fully characterized by MAVEN (Dong et al., 2015; Liemohn et al., 2014). It originates from the interaction between the solar wind and the ionosphere of Mars, which creates an upward electric fields through $E = -U_{SW} \times B$. This process can be looked as a special case of ion-pickup since it is observed in such models. At Mars, this plume escape for $O^+$ is estimated to be 30% of the tailward escape, equivalent to 23% of the total ion escape (Dong et al., 2015).

### 2.5.3. Ambipolar Fluxes/Outflow Anomalies/Snowplow

The “cold” ions, that is, ions not energized at suprathermal temperature and coming from the ionosphere, are prominent in the plasmasphere (Kun et al., 2017). Several processes lead to the filling of that plasmasphere. Ion upwelling (Strangeway et al., 2005), which can be linked to the ambipolar electric field, but at levels that do not lead to escape, is one of these processes at Earth. For unmagnetized planets, ions are transported in the upper layers of the ionosphere by ambipolar electric field (Akbari et al., 2019; Collinson et al., 2019). The draping of the IMF could lead to additional induced field with respect to the processes described in section 2.4; this process is called the “snowplow” (Halekas et al., 2016). From there, the transfer of momentum from the solar wind to the ionosphere creates detached plasma clouds that are escaping. It is important to note that the $E \times B$ drift can help these ions escape as observed at Mars, where different escape rates are observed in the $+E$ and $-E$ hemispheres (Inui et al., 2019).

### 2.5.4. Plasmaspheric Ion Losses/Substorm Losses/Plasmoids/Flux Ropes

As previously shown, the charge exchange is a major ion loss process in magnetospheres. A fraction of ions can be lost from escape from the tail, but, as shown from the observations at Jupiter, this is a low percentage of the escape (Jackman et al., 2014). (At Jupiter, other processes have to be taken into account to balance the input of plasma from Io.) This escape from the tail works through the creation of a plasmoid: The pressureof the IMF elongates the planetary magnetic field, leading to the reconnection of that magnetic field. This reconnection means that a part of the initial magnetosphere is no more linked to the magnetic field of the planet and is ejected into space, along with the plasma it contains. Ejection of plasmoids is also associated to a return of plasma toward the planet. (This is different for situations when magnetic fields are weaker, such as Mars.) The equivalent of these plasmoids has been observed at Mars near the magnetospheres created by the crustal magnetic field; they are usually named “flux ropes” (Hara et al., 2017) and are linked with CME disturbances and other processes (Hara et al., 2017). They could be responsible for up to 10% of the present-day ion escape at Mars (Brain et al., 2010). Finally, plasmoid escape has also been observed at Venus, in the induced magnetic field, and looks like the Earth’s or Jupiter’s plasmoids (Zhang et al., 2012).

### 2.5.5. Questions

These ion escape processes are actively studied with missions such as MEX, VEX, and MAVEN, as well as numerical models. Most of the questions are linked to the actual amount of ions escaping due to these processes and how these evolve with the solar/stellar activity. From an observation point of view, it may be difficult to distinguish between processes from the a single point observation of the amount and location of
plasma escaping (Inui et al., 2019). In addition, some processes can be seen as generalization of other processes (e.g., the “snowplowing” is a generalization of the ion outflow observed in magnetospheres). These points led to the above definitions and organization of these ion escape processes.

2.6. Ion Return and Net Escape Rates

While the ionospheric outflow processes detailed above determine the escape of plasma from the ionosphere, a significant fraction of this plasma becomes trapped in Earth’s magnetosphere. Magnetospheric ions mostly consist in a mixture of H⁺ and O⁺ ions. Contrary to H⁺ ions, which can either originate from the solar wind or the ionosphere, O⁺ ions almost exclusively originate from the ionosphere and are used as tracers of ionospheric material in the magnetosphere. They have been observed by several spacecraft, including GOES 1 and 2 (Young et al., 1982), ISEE (Lennartsson, 1989; Lennartsson & Shelley, 1986), Van Allen Probes (Fernandes et al., 2017), GEOTAIL (Nosé et al., 2009; Ohtani et al., 2011), and Cluster (Kistler & Mouikis, 2016; Maggiolo & Kistler, 2014). All these observations show an increase the amount of O⁺ ions in the magnetosphere, and thus of ionospheric material, with increasing solar EUV/UV flux and geomagnetic activity, that is, with the amount of energy deposited din the ionosphere. Once in the magnetosphere, ionospheric material enters magnetospheric circulation patterns, which may ultimately result in loss to interplanetary space or return to the ionosphere. Seki et al. (2001) estimate the fraction of Earth’s oxygen lost to interplanetary space at about 1/10 of the ionospheric oxygen outflow during periods of low solar activity, based on the estimate of the O⁺ loss due to the main four escape routes for terrestrial ions: the escape of cold detached plasmaspheric particles through the magnetopause, of high-energy ring current/dayside plasma sheet particles through the magnetopause, of plasmasheet ions through antisunward flow in the nightside plasma sheet, and of terrestrial ion beams through the lobe/mantle. Note that the charge exchange loss of ring current ions was not considered by Seki et al. (2001). The outflow and loss rates are enhanced during high solar and geomagnetic activity but may not account for all magnetospheric loss mechanisms, particularly for low-energy ions that are difficult to observe. This estimate, then, should be considered a lower bound on escape, but the important point here is that not all outflowing ions escape from the magnetosphere-ionosphere system.

However, recent observations above the polar ionosphere in the magnetospheric lobes by the Cluster spacecraft provide evidence for a higher loss rate for ionospheric ions flowing through the lobe and mantle region. Slapak et al. (2017) showed that energetic ions (in the range of a few hundred to several thousands of eV) escaping from the cusp region through the magnetospheric lobes/mantle have a high probability of being lost to interplanetary space rather than being returned to the ionosphere. They even claim that over geological times a quantity of oxygen lost by the Earth’s atmosphere could be roughly equal to the amount of the present atmospheric oxygen content if the young Sun was actually more active than nowadays. Furthermore, the flux of precipitating ions as estimated by the DMSP satellites is only of the order of 10²⁴ (Newell et al., 2010): 1 to 2 orders of magnitude lower than the estimated flux of outflowing ionospheric ions. These new observations provide strong evidence against a high return rate of ionospheric ions and rather suggest that a significant fraction of ionospheric ions escaping from the ionosphere may actually be definitively lost into the interplanetary space.

3. Major Parameters and Concepts

In order to address the escape rate of an atmosphere and to retrieve its evolution with time, it has been demonstrated that several processes are in action. To evaluate whether or not they are negligible at a certain period in time, or to approximate the calculations, several concepts have been proposed, such as the energy-limited escape or the critical heating rate for hydrodynamic escape. The two major parameters in the different models are the energetic inputs, from the EUV-XUV fluxes to the electron precipitations, and the atmospheric structure and composition. Finally, it is very important to take into account the evolution with time, from the time dependence on small scales (typically sensitive to the solar/stellar activity) to the evolution of the atmospheric escape through eons, leading to isotopic fractionation, which is the main probe for the history of our solar system’s atmospheres (in the absence of better in situ measurements, e.g., trapped gases in rocks Jakosky, 1991).
3.1. Limiting Parameters

3.1.1. Critical Heating Rate

Present theory (section 2.1.1.2) incompletely describes transition from Jeans escape to hydrodynamic escape. Transonic models (Murray-Clay et al., 2009) have been used to describe rapid escape from exoplanets and from Pluto (Strobel, 2008a). However, Johnson et al. (2013a, 2013b) have recently discovered that this model for Pluto gave an incorrect upper atmospheric structure (Tucker et al., 2013b). This erroneous prediction of the upper atmospheric structure results from applying the Jeans expressions at the exobase (Chamberlain & Hunten, 1987) for uncertain boundary conditions at infinity (Tian et al., 2008) when simulating rapid escape using continuum gas dynamics. In this context, a sonic point is assumed to occur at an altitude \( r_a \), above which the density and temperature dependence can be simply characterized (Parker, 1964a, 1964b). The hydrodynamic, energy-limited (see section 3.1.2) escape rate, applied to exoplanet atmospheres (Lammer et al., 2009), is often assumed to imply that sonic boundary conditions are applicable (Erkaev et al., 2013). Johnson et al. (2013a, 2013b) used molecular kinetic simulations to show that this is not the case. Ignoring viscosity, Parker (1964a, 1964b) used the momentum and energy equations to describe escape when the dominant heat source is internal. This same model was applied to planetary atmospheres primarily heated at an altitude \( r_a \). For a Jeans parameters at \( r_0 \) (the lower altitude considered) as large as \( \lambda_0 \sim 40 \), such models were assumed to produce a transonic expansion, which is often referred to as a slow hydrodynamic escape (Strobel, 2008a, 2013b). However, rapid escape can occur for large Jeans parameters only when the Knudsen number is low, that is, when the collisional approximation cannot be assumed. Therefore, the gas does not go sonic in the collision-dominated region and the escape rate computed in the “slow hydrodynamic escape” paradigm is a few times larger than the Jeans rate (Volkov, Tucker, et al., 2011; Volkov, Johnson, et al., 2011). As we explained in section 2.1.3.1, the observations that led to the “slow hydrodynamic escape” hypothesis could be explained by alternative processes based on chemistry.

From there Johnson et al. (2013a, 2013b) have developed a criterion to check if a transonic solution will exists, that is, if we can approximate the escape by a hydrodynamic model. Assuming that \( r_0 < r_a < r_c \), which should be the case in hydrodynamic escape, it was found that the net heating rate \( Q_{\text{net}} \) should follow Equation 65:

\[
Q_{\text{net}} > Q_c \approx 4\pi r_a \frac{\gamma}{c_s^2} knm \sqrt{\frac{2U(r_a)}{m} U(r_0)} \tag{65}
\]

\[
U(r) = \frac{GM}{r}
\tag{66}
\]

\( K \) is the Knudsen number, and \( c_s \) is determined by the energy dependence of the total collision cross section, \( \sigma_c \).

If heat is primarily absorbed over a broad range of \( r \) below \( r_a \), we can use \( knm \sim 1 \) as an approximation. Here, it can be seen that \( Q_c \) does not explicitly depend on \( T_0 \), but on the sonic point only where a lower bound can be obtained by replacing \( r_a \) with \( r_a \), the mean absorption depth. This mean absorption depth is estimated from \( \sigma_a \), the absorption cross section. At threshold, the sonic point will approach \( r_a \), such that \( r_s \sim r_a \)

\[
1 + \left( \frac{\sigma_a}{c_s \sigma_c} \right) \lambda_{\text{ave}} \lambda_{\text{ave}} \sim (\lambda_{\text{ave}} + 2\gamma)/2, \text{which slightly increases } Q_c. \text{ Using Pluto as an example, UV/EUV absorption at } r_a \sim 1.5 \text{ times Pluto’s radius, } knm \sim 10^{-3}, \text{ and } r_a \sim r_a \sim r_0, \text{ Equation 65 gives } Q_c \sim 4.5 \times 10^{10} \text{ W for Pluto, which is well above the largest heating rate and shows that hydrodynamic escape should not be applied for the dwarf planet}. \text{ We compiled the values of } Q_c \text{ in Table 2.}
\]

3.1.2. Energy-Limited Escape—Radiation/Recombination-Limited Escape

The estimation of mass loss rate of exoplanets often assume an energy-limited escape (section 2.1.1.4). The basis of that assumption is that an exoplanet thermosphere is mainly composed of H, heated by ionization of H. From there, it is supposed that a large quantity of that heat is transformed into hydrodynamic escape. Therefore, one uses an efficiency coefficient \( \epsilon \) (sometimes \( \eta \)) for transforming EUV-XUV energy into escape. This led to Equation 14, with the standard efficiency coefficients found in the literature. Erkaev et al. (2007) show that this equation can be slightly modified to account for stellar gravity effects that affect close-in planets.
For giant planets close to very active stars, the radiation-recombination-limited escape is often used as a harsher limit to the energy-limited escape, because the \( \text{H}^+ \) can recombine, reducing some of the energy in the system (Linsky, 2019; Luger, 2017); this leads to an escape proportional to \( \sqrt{F_{\text{XUV}}} \) instead of \( F_{\text{XUV}} \). In the case of H atmospheres where the heating is only supposed to come from ionization, there is also a case where the escape is limited by the number of ionizing photons.

This approach has been developed to study close-in giant planets (Salz et al., 2016) and led to energy diagrams (Ehrenreich & Désert, 2011) to evaluate the mass loss from giant exoplanets. Unfortunately, it notably neglects the radiative cooling processes in the upper atmosphere of the planet; that is, it neglects the problem of the upper atmosphere temperature. (Note: Lopez, 2017 includes radiative cooling in an energy-limited diffusion approach.)

The main problem of the energy-limited escape approximation is that it is too often applied for rocky exoplanets while concealing these major limitations:

1. **The escape regime is not evaluated:** If the exoplanet is not in hydrodynamic escape, the energy-limited escape equation will give an overestimate of the escape.
2. **The atmospheric profile is not evaluated:** What is exactly the profile of the atmosphere and therefore the efficiency of the escape if it is truly in hydrodynamic regime?
3. **The atmospheric composition is not taken into account:** H is assumed to be the only species: This is related to the other problems; the presence of cooling species such as \( \text{CO}_2 \) may totally change the escape regime; diffusion-limited processes may prevent H to be present in large quantities in the thermosphere, etc.
4. **Only photo-ionization heating is taken into account:** Joule heating or particle precipitations can be large sources of heating for close-in exoplanets.
5. **Nonthermal processes are not addressed:** those can dramatically change the profile of the escaping species.

Energy-limited escape models can be interesting for studying H-rich rocky planets early in their histories, for which the escape of H may not have been diffusion limited but energy limited (Tian et al., 2005); however, energy-limited escape is less relevant to more comprehensive habitability studies.

### 3.1.3. The Diffusion-Limited Escape

Some escape processes can be very efficient, and limited by the amount of particle available for the escape, the bottleneck for the escape of these particles will then be the diffusion from the lower layers of the atmosphere to the upper atmosphere. Typically, the escape of H at Earth is diffusion limited. It follows the Equation 15. At Titan, like Earth, \( \text{H}_2 \) escape is determined by the limiting flux through the homopause deep in the lower thermosphere (Bell et al., 2014; Cui et al., 2008; Strobel, 2012). However, there is currently a discrepancy between the densities of \( \text{H}_2 \) measured in situ by INMS and those produced by modeling studies (Bell et al., 2014; Cui et al., 2008; Magee et al., 2009; Strobel, 2002). Despite this discrepancy, all modeling studies to date have indicated that the \( \text{H}_2 \) upwelling into the lower thermosphere, combined with additional \( \text{H}_2 \) produced in the thermosphere, sets the eventual planetary escape flux of \( \text{H}_2 \). A more complete theory of diffusion-limited escape, including the cases where the diffusing species has a nonnegligible mass with respect to the main species can be found in Hunten (1973).

### 3.2. Energetic Inputs

#### 3.2.1. The EUV/XUV Flux

The EUV-XUV flux modifies the temperature of the exosphere and the exobase altitude. It therefore changes the concentration of particles above the exobase. It is also responsible for the creation of hot atoms through photochemical processes. At the Earth, the EUV-XUV flux varies substantially as a function of solar activity. When the variability of the solar irradiance is rather low for the visible and the IR, with less than 0.1% and 1% from minimum to maximum respectively, the solar irradiance variability in the XUV /EUV can be more than doubled with a direct impact on the upper atmosphere (Haigh, 2007). This variability is of two different origins: One depends on sporadic explosive events such as flares with timescales from minutes to hours, while the second one is linked to the full Sun disk activity with longer cycles, from days to years. The latter one is then related to the appearance and disappearance of active regions on the solar disk, which causes then the variability on a 27-day solar rotation scale, associated with a 13.5-day modulation from the center-to-limb variation. The long-term monitoring of the solar EUV flux, however, is a difficult task, mainly because of the heavy degradation experienced by the solar instruments that are in orbit (BenMoussa et al.,
Before 2002 with the launch of the TIMED satellite (Woods et al., 2005), measurements of the solar EUV flux variability were rather scarce. This has led to the development of several empirical approaches for reconstructing the solar XUV/EUV part of the spectrum.

A common approach lies with using solar proxies such as the radio measurements at 10.7 cm (F10.7) (Tapping & Detracey, 1990) and the MgII core-to-wing index (Heath & Schlesinger, 1986). Many models are then using a linear combinations involving these proxies and their 81-day running means or even non-linear combinations (Hinteregger, 1981; Lean et al., 2003; Richards et al., 2006). However, no single index can properly reconstruct the solar XUV/EUV irradiance at all timescales (Dudok de Wit et al., 2009). Moreover, some widely used proxies, such as F10.7, are not really suited for the XUV/EUV lines reconstruction, whose originated from the solar corona. The F10.7 index is, however, used as the solely index to estimate the solar variability within thermospheric and ionospheric models. For the solar minimum in 2008, when the thermospheric density dropped by 28%, the F10.7 only decreased by 4% (Emmert et al., 2010), outlining then the limitations of the F10.7 index for ionospheric studies (Solomon et al., 2010). More appropriate solar proxies have been recently suggested such as the radio measurements at 3 and 30 cm, which are directly linked to chromospheric and corona emissions (Dudok de Wit & Brusselsma, 2017).

A different approach considers that the solar spectrum is a linear combination of reference spectra that coming from different regions of the solar disk. Those regions are attributed to the quiet Sun, coronal holes, and active regions and can be disentangled using solar images or solar magnetograms. Their respective contrast can be obtained by an empirical approach (Worden et al., 1998) or using the differential emission measure (Kretzschmar et al., 2004). A few terms are normally needed to reconstruct the solar irradiance in the XUV/EUV spectral range (Ambard et al., 2008). This strongly outlines that the spectral variability is highly coherent through the spectrum, but this only for timescales that exceed the dynamic time of solar flares, since the solar atmosphere is strongly structured by the magnetic field. The solar spectrum in the XUV/EUV can then be reconstructed from measurements of a few correctly chosen passbands (Cessateur et al., 2011, 2012). For the short-term spectral variability, a specific model has been developed, the Flare Irradiance Spectrum Model (FISM) (Chamberlin et al., 2008), based on Thermosphere Ionosphere Mesosphere Energetics and Dynamics’ (TIMED)/Solar EUV Experiment (SEE) and Solar Dynamics Observatory (SDO) data.

The effects of the solar XUV/EUV variability on Earth’s upper atmosphere have been quantified with empirical models (Bowman et al., 2008), which specify the exospheric temperatures as a function of indices of EUV radiation at different wavelengths (Tobiska et al., 2008). At Mars, Luhmann et al. (1992) computed the influence of the EUV flux on the escape processes. It is complicated by the fact that the solar wind pressure is also included in the calculations: The EUV flux increases; therefore, the density of hot oxygen above the exobase increases (and the altitude of the exobase increases). Therefore, the escape of hot oxygen increases, and the density of pickup ions increases as well, and so the sputtering and the sputtered atoms. These nonlinear effects lead to the large variations in the escape rates as computed in Figure 3. More recent modeling and data show that the actual increase is less important than that previous simulations (Lillis et al., 2015). The correlation of MEX observations of ion escape at Mars with the EUV flux shows that it is difficult to draw a direct relation between the two (Ramstad et al., 2015) in the 7-year span these observations took place. However, the nonlinearity of the dependence, and the fact that negligible escape processes can become very important for extreme EUV-XUV flux, such as in the conditions in the beginning of the solar system, is still valid.

### 3.2.2. The Electron Flux

#### 3.2.2.1. The Auroral-Like Electron Flux

The energetic electron flux at high latitudes is produced as a result of the interaction of the solar wind and IMF with the magnetic field and magnetosphere of the planet, which in turn drives ionospheric electric fields and currents. Upward currents may contain a significant, downward energy flux from electrons (Fuller-Rowell & Evans, 1987). The energy flux at the Earth typically ranges from under 1 up to 20 GW (Newell et al., 2010).

As Mars only has a limited magnetosphere, there is no significant energy deposited by the aurora as discovered by MEX in 2005 (Bertaux et al., 2005). This conclusion can be challenged by the observations of global aurora during solar events (Schneider et al., 2018). Diffuse electron (Clancy et al., 2017) and proton aurora
(Deighan et al., 2018) observed by MAVEN may carry significant amounts of energy, but the total flux still needs to be estimated.

On the other hand, both Jupiter and Saturn do have large internal magnetic fields and correspondingly large magnetospheres, so there is considerable power in their aurorae. As there are no direct measurements available, much of what is known about the outer planets’ aurorae has been obtained from UV measurements, at first on the Voyager flyby of Jupiter (Broadfoot et al., 1979). Most recent UV observations are from the Hubble Space Telescope (HST). In a review of such observations, (Grodent, 2014) indicated that the auroral emissions at Jupiter and Saturn are on the order of 1 and 0.1 TW, respectively. Uranus and Neptune are much weaker, at 1 GW or less, and observations are sparse. Of course, the power of the emissions is less than the kinetic energy that is deposited. The Voyager UV measurements at Jupiter has implied a power injection on the level of 12 TW (Broadfoot et al., 1981), and Gérard et al. (2014) stated that the auroral precipitation at Jupiter has a power on the order of 10 to 50 TW. As this level of heating is much greater than that from solar radiation, the aurora has a significant contribution to the thermal properties of the upper atmosphere.

3.2.2. The Suprathermal Electrons
Suprathermal electrons are electrons with energy higher than the typical electron in an ionosphere: When looking at the flux of electrons in function of energy, the suprathermal electrons are responsible for the departure of the curve from a Maxwellian at high energy. These electrons come mainly from the precipitation of electrons from outside of the ionosphere, from local creation (typically photoionization—hence the name of photoelectrons), but also from other ionization, including from suprathermal electron impact). Electric potential drops can accelerate electrons to suprathermal energies, but they occur outside the ionosphere and are responsible for some magnetospheric precipitation at Earth. To understand the effect of the suprathermal electrons, it is necessary to compute their transport in an atmosphere. Codes such as Aeroplanets and PWOM do that.

The basis of these codes is to compute the flux of electrons by solving their transport equation. The existence of codes not based on a Monte Carlo scheme, such as Aeroplanets, allow to fast compute large quantities of conditions and to perform sensitivity analysis (Gronoff, Simon Wedlund, Mertens, & Lillis, 2012; Gronoff, Simon Wedlund, Mertens, Barthélemy, et al., 2012). We refer to these papers for the equations to solve in the ionosphere/thermosphere and for the uncertainties encountered.

3.2.3. The Electromagnetic Energy
The Joule heating is the heating created by the resistance of the thermosphere to the electric current due the ionospheric plasma (Vasyliunas & Song, 2005). It is computed by evaluating the electric field and the conductivities.

Joule heating in the polar ionosphere has a significant effect on the exospheric temperatures and hence the amount of outflow (section 2.4.4). At the Earth the total Joule heating is normally in the range of a few hundred GW but in extreme events can range from 1 TW (Lu et al., 1998) up to 5 TW while increasing the mean temperature of thermosphere by up to 500°K (Weimer et al., 2011). At the same time, the additional heating tends to increase the amount of nitric oxide in the thermosphere, which acts to accelerate the rate at which it cools down to the equilibrium temperature set by the solar EUV radiation (Weimer et al., 2015). Wilson et al. (2006) had found that Joule heating is most typically about 3 times the energy from precipitating particles, with the ratio varying from 2 to 7 in the different events that were studied.

At other planets there are no direct measurements of the electromagnetic energy input into their ionosphere and thermosphere, so at present it can only be estimated. At Jupiter, Strobel (2002) estimated the Joule and auroral particle heating to be about 1,000 times larger than at the Earth for typical conditions, which would be on the order of 500 TW.

The generation of currents and electromagnetic energy at Jupiter may be dominated by processes much different from at the Earth, as the interaction of the solar wind and IMF are weaker. It is thought that the planet’s rotation and magnetic field provide a significant contribution to the energy sources of the heating processes (Eviatar & Barbosa, 1984; Waite & Lummerzheim, 2002).

Due to the lack of observations of the electromagnetic fields at other planets, most of what is known is derived from computer simulations, such as the Jupiter Thermospheric General Circulation Model (JTGCM), which addresses global temperatures, three-component neutral winds, and neutral-ion species.
distributions (Bougher et al., 2005). In a case study with auroral forcing plus ion drag, Bougher et al. (2005) calculated exospheric temperatures at auroral latitudes ranging from 1200 to 1300 K, which match available multispectral observations. The levels of Joule heating are in the range of 70 to 140 mW/m² in the auroral ovals, while the auroral particles produce 2 to 8 mW/m². With different model parameters higher levels of the Joule heating can be produced and exospheric temperatures above 3000 K may be achieved. Other numerical studies have been done, too numerous to mention here. The main point is that Joule heating can significantly modify the heat budget of the thermosphere in the Jovian gas giant, and similar processes would be expected at similar exoplanets. As there are many assumptions and approximations made in the modeling process, more work needs to be done to more accurately calculate the contribution of Joule heating to the exospheric temperatures and the resulting effects on the outflow, particularly the contributions from the solar wind dynamo.

3.2.4. The Cross Sections and the Computation of Ionization

Elastic and inelastic cross sections are at the core of the computation of the energy transfer from particle precipitation to the atmosphere. To that extent cross sections for ionization, excitation, and dissociation are necessary tools for all the computations. Several efforts have been made to gather cross sections. The most comprehensive one has been recently developed with the study of upper atmospheres in mind, called AtMoCIAD. Its advantage is the inclusion of error bars, which allows the computation of the propagation of the experimental or theoretical uncertainties (Gronoff, Simon Wedlund, Mertens, & Lillis, 2012; Gronoff, Simon Wedlund, Mertens, Barthélemy, et al., 2012) but also the inclusion of all kinds of particles (photons, electrons, protons, hydrogen, ...) colliding with atoms or molecules.

The precise knowledge of all types of cross section can improve the computation of the different conditions at different planets. A consistent set of cross sections allows to perform comparative planetology studies. An example of such a computation can be seen in Figure 7.

Other cross sections such as charge-exchange cross sections are of importance for escape studies. The database maintained by the Atomic and Molecular Collisions Group of the Department of Physics and
Astronomy at Rice University (Houston, USA, http://www.ruf.rice.edu/~atmol/) is among the most populated with species of interest for space science studies (Lindsay & Stebbings, 2005).

3.3. Atmospheric Structure

Addressing the atmosphere structure is one of the more complex parts of the study of upper atmosphere. Model should address both ionospheric problems, such as the precipitation of particles and Joule heating, as well as fluid problems like the heat transport, winds, or radiative problems such as the CO$_2$ 15-μm cooling (Johnstone et al., 2018, and references therein).

3.3.1. Vertical Mixing and Photochemical Modeling of Atmospheres

Most of the planets in our solar system have a substantial atmosphere, with the exception of Mercury, which has a very tenuous atmosphere. Several moons in the solar system also have atmospheres. An atmospheric gas can be made up of a variety of chemical species that were distributed unevenly at the time of the formation of the solar system. A basic relationship between fundamental quantities governing the gas distribution of chemical species is the ideal gas law, \( p = n(z)kT \), which becomes increasingly less valid for pressures greater than 1 bar, after which Van der Waals equation of state should be used (Parkinson, 2002).

Knowledge of the photochemical and chemical processes governing the transformation of a particular atmospheric species into another can be used to calculate the distribution of each species considered throughout the atmosphere. Since a particular atmospheric constituent might be the source of one or more other constituents, this calculation requires the simultaneous solution of a series of coupled continuity equations, one for each atmospheric species considered; namely,

\[
\frac{\partial n_i}{\partial t} + \nabla \cdot \phi_i = P_i - L_i n_i, \tag{67}
\]

where \( \phi_i \) is the flux of a particular species and \( t \) is time. The species number density is given by \( n_i \), \( P_i \) is the chemical production rate, and \( L_i \) is the loss frequency at altitude \( z \) and time \( t \) (Chamberlain & Hunten, 1987, see).

The solution of Equation 67 yields the distribution of the species that are being studied. This solution is obtained by considering the various photochemical and chemical production and loss terms in addition to the effects of composition, eddy diffusion, temperature, mixing ratio, and the solar flux on the various constituents distribution. This method of solution is described for one dimension in the sections that follow.

3.3.1.1. One- Dimensional General Method of Solution

The vertical distribution of a minor constituent in a planetary atmosphere is governed by the one-dimensional continuity equation for each species, \( i \):

\[
\frac{\partial n_i}{\partial t} + \frac{\partial \phi_i}{\partial z} = P_i - L_i n_i, \tag{68}
\]

where the vertical flux, \( \phi_i \), can be approximated by

\[
\phi_i = \phi_i^K + \phi_i^D. \tag{69}
\]

The eddy flux, \( \phi_i^K \),

\[
\phi_i^K = -K \left( \frac{\partial n_i}{\partial z} + \frac{1}{H_w} + \frac{1}{T} \frac{\partial T}{\partial z} \right) n_i \tag{70}
\]

represents the vertical flux that parameterizes macroscopic motions, such as the large-scale circulation and gravity waves, and \( \phi_i^D \)

\[
\phi_i^D = -D_i \left( \frac{\partial n_i}{\partial z} + \frac{1 + \alpha_i}{T} \frac{\partial T}{\partial z} + \frac{n_i}{H_i} \right) \tag{71}
\]

is the vertical flux carried by molecular diffusion. The species number density is given by \( n_i \), \( P_i \) is the chemical production rate (cm$^{-3}$ s$^{-1}$) and \( L_i \) is the loss frequency (s$^{-1}$) at altitude \( z \) and time \( t \) (e.g.,
One of the fundamental properties of a planetary atmosphere is the amount of mechanical mixing forced by eddy diffusion. The molecular diffusion coefficients, \( D_i \), are taken from Chamberlain & Hunten (1987) where applicable using the formula

\[
D_i = \frac{b_i}{n_{bg}} = \frac{AT_i}{n_{bg}}
\]

where \( b \) is the binary collision parameter (expressed in terms of the coefficients \( A \) and \( s \)) and the subscript “bg” denotes background. \( H_i \) and \( H_{av} \) are, respectively, the constituent and background atmospheric pressure scale heights; that is, \( H_i = \frac{kT}{M_i g} \) and \( H_{av} = \frac{kT}{M_{av} g} \), where \( M_i \) and \( M_{av} \) are, respectively, the molecular weights of the constituent and the atmosphere. In these calculations we have neglected the effects of the thermal diffusion factor, \( \alpha_z \), as its inclusion contributed less than 1% to a given species column in test runs.

Eddy mixing tends to homogenize the atmosphere such that, where there are no effects due to chemistry, all species would be distributed according to the mean atmospheric pressure scale height. Molecular diffusion tends to separate constituents by their individual molecular weights. The atmospheric level at which the molecular diffusion coefficient is equal to the eddy diffusion coefficient is defined as the homopause for the \( i \)th constituent. Above this altitude, molecular diffusion dominates and the time constant for reaching diffusive equilibrium is given by \( \tau_D = \frac{H_i^2}{D_i} \) (Chapman et al., 1990; Colegrove et al., 1966). Below the homopause, eddy diffusion dominates and the long lived species are “mixed,” and the mixing time constant is analogously expressed as \( \tau_K = \frac{H_{av}^2}{K} \).

Equation 68 is solved using a finite central difference approximation for the vertical derivatives and the species densities are solved semi-implicitly in time using a simple tridiagonal solver. For these applications we have assumed a steady state exists and so have driven the solution so that \( \frac{1}{P} \frac{\partial n_i}{\partial t} = 0 \). Examples of such models and details are given in Parkinson (2002) and Yung and DeMore (1982).

### 3.3.1.2. Eddy Diffusion Coefficient, \( K(z) \)

One of the fundamental properties of a planetary atmosphere is the amount of mechanical mixing forced by large-scale circulation, gravity waves and other processes. In a one-dimensional model, this mixing is often characterized by the eddy diffusion coefficient, which we will denote by \( K, K_z \), or \( K(z) \). The value of \( K(z) \) in the vicinity of the homopause, \( K_\ast \), is critical in determining the onset of the importance of molecular diffusion. Estimates of \( K_\ast \) for the outer planets have been obtained by various means: for example, analyses of the H Lyman-\( \alpha \) albedo (Wallace and Hunten, 1973; Atreya, 1982; Ben Jaffel et al., 1993; Ben Jaffel et al., 1994), the fall-off in hydrocarbon profiles, as measured against an H\( _2 \) background, using solar and stellar occultation data (Atreya et al., 1981; Festou & Atreya, 1982; Romani et al., 1993), the He 584 Å albedo (McConnell et al., 1981; Parkinson et al., 1998; Sandel et al., 1982; Vervack et al., 1995), and the CH\(_4\) fluorescence (Drossart et al., 1999).

### 3.3.1.3. Thermospheric-Ionospheric Simulations

The modeling of a thermosphere-ionosphere is slightly different than the deeper layers of the atmosphere: A density profile has to be taken into account for each different neutral species, since they follow their own scale height. Suprathermal species can exist, such O in the upper atmosphere of Mars, resulting from \( O_2^+ \) dissociation. For the ionized species, a different temperature has to be computed (and it changes with the species in the most complicated simulations). Finally, electron temperatures have to be addressed. The full description of these models is outside the scope of this paper. We refer the reader to the following studies and their included references (Gough et al., 2005; Johnstone et al., 2018).

### 3.3.1.4. Importance of the 3-D Modeling

Three-dimensional models (3-D) models provide a broad characterization of the whole atmosphere that couple chemistry, dynamics, and energy balance. These numerical tools, while not capable of including the details of their one-dimensional (1-D) counterparts, can capture the effects of global dynamics, diurnal chemistry, and the resulting energy balance. It has been shown that the approximations made with 1-D modeling are not able to fully reflect the reality of a planetary climate. For example, the presence of clouds, ice sheets, oceans, etc. has large effect able to change a nonhabitable planet into one (Way et al., 2016, 2018). For thermospheres-ionospheres, the 3-D effects of transport and cooling lead to different results as well, which may change our view of an exoplanet.
3.3.2. Exospheric Temperature

The exospheric temperature, \( T_{exo} \), is one of the most important parameters in the study of nonhydrodynamic atmospheric escape. It is the temperature at the base of the exosphere. Its effects on atmospheric escape are numerous. First, a higher \( T_{exo} \) means a higher thermal escape. Second, with a warmer thermosphere, the exobase increases with altitude thereby increasing the exobase surface. This in turn implies a higher total escape from the planet and a greater cross section to nonthermal escape. Third, a high \( T_{exo} \) means that non-thermal processes can be more efficient.

The exospheric temperature depends upon (a) the UV flux (photon heating), (b) the chemical heating, (c) the electromagnetic energy (Joule heating), and (d) precipitation (auroral heating) in the atmosphere. The rate of cooling, primarily by IR radiation, depends upon the composition and the adiabatic expansion. The equilibrium between the heating and cooling factors gives the temperature. Since wind and UV heating are important factors, major dayside-nightside exospheric temperature differences can occur. Full 3-D models such as the Global Ionosphere-Thermosphere Model (GITM) (Ridley et al., 2006) are therefore necessary to obtain a correct value for the exospheric temperature. A 1-D approximation of the temperature can be made, but, in the case of the study of nonthermal escape, it may become a major problem. This is because the day-night asymmetry from the escape processes is correlated with the asymmetry from the exospheric temperature, which could lead to severe errors in the determination of the magnitude of the escape.

The exospheric temperature is determined by the equilibrium between heating and cooling. Since these processes are altitude dependent, it is often necessary to determine the structure of the thermosphere and compute the exospheric temperature from the basic equations. Empirical models exist, for example, for Earth (Weimer et al., 2011). In the planets of the solar system, heating is dominated by (1) photoexcitation and cooling by (2) thermal conduction (González-Galindo et al., 2009). The photoexcitation/photodissociation heating is due to the kinetic energy left in these processes: The difference between the threshold \( E_{th,k} \) of the \( k \)th reaction on a species, \( s \), and the energy, \( E \), of the photon is transformed into heat. When the flux of photon per unit energy is \( \Phi(E) \), we have

\[
Q_{UV,k} = \int_{E_{th,k}}^{\infty} (E - E_{th,k}) n_k \sigma_k(E) \Phi(E) dE.
\]  

The thermal conduction is solved through the following equation (González-Galindo et al., 2009):

\[
\frac{\partial T}{\partial t} = \frac{1}{\rho c_p} \frac{k}{\partial z} \left( \frac{\partial T}{\partial z} \right)
\]

(73)

\[
k = AT^{0.69}
\]

(74)

With \( \rho \) being the density (kg/m\(^3\)), \( c_p \) the heat capacity, and \( A \) the weighted average of the thermal conductivities.

Two major parameters are to be carefully determined when estimating the exospheric temperature and are the most complicated to address to date: (3) the chemical heating/cooling and (4) the radiative cooling.

The chemical heating, due to the exothermic reactions, and cooling, due to the endothermic reactions, follow the ionization and dissociation by precipitating particles (including photons). Computing this contribution requires to carefully evaluate the chemical reactions chains and their energies. Those are atmospheric-composition dependent and can be quite complex and not well understood (e.g., Titan).

The radiative cooling is mainly due to the de-excitation of molecular species in a rotational or vibrational state. Simple approximations of that cooling can be made if the cooling species is in low quantity in the atmosphere and if it is excited only by thermal processes, that is, if it is in a LTE and if the emission line (or band) is optically thin. More complex cases exist in the atmospheres (such as non-LTE processes that are known to happen in auroral regions and optically thick cases), which require precise radiative transfer calculations (Mertens et al., 2008, 2009). In addition, very complex cases such as state inversion and MASER can be obtained, such as those occurring at Mars and Venus at 10 \( \mu \)m (Mumma, 1993) and probably at some
exoplanets (Cosmovici & Pogrebenko, 2018). Finally, some of the radiative species can be obtained by chemical reactions when the system is out of equilibrium, for example, NO cooling at Earth (Weimer et al., 2015). For the extrapolation of solar system planets’ situation to other stellar systems, it is important to validate such approximations.

Other important parameters have to be considered depending on the cases studies: (5) near-infrared (NIR) heating, important in the case of CO2-rich planets, (6) dynamic cooling—from winds or expansion, and (7) heating from gravity waves dissipation (Hargreaves, 1992).

### 3.3.3. The Exobase Altitude

The exobase is the altitude at which the scale height is equal to the mean free path of a thermalized particle (at $T_{\text{exo}}$). Above this altitude, the mean free path is greater than the scale height, and a particle with sufficient energy is likely to escape without any collision.

One can approximate the density in the thermosphere by $n(z) = n_0 \times e^{-z/z_{\text{exo}}}$ (nb: this is valid for a thermosphere with one constituent; if multiconstituent a H will have to be defined for each of those, but the exobase is usually defined for the main constituent). At the exobase, we have $n_{\text{exo}} = H / \sigma$ with $\sigma$ being the collision cross section between the main molecules. If we suppose an isothermal thermosphere, that is, H does not vary with altitude, it is possible to easily retrieve the exobase altitude: $z_{\text{exo}} = z_0 - H ln \left( \frac{n_{\text{exo}}}{n_0} \right)$. For multicomponent atmospheres and varying temperature, the evaluation becomes more complex since H and $\sigma$ (and therefore $n_{\text{exo}}$) vary with altitude.

### 3.4. Time Dependence and Creation of Observable Markers

Once the main processes leading to atmospheric escape are known, the study of their influence in time requires evaluating the evolution of the stellar forcing parameters. If possible, the study of the isotopic ratio in the planetary atmosphere will be a major input for validating the calculations and estimating the influence of other processes such as outgassing, etc.

#### 3.4.1. Evolution in Time of the Stellar Forcing Parameters

Stellar rotation drives the magnetic activity responsible for UV to X-ray emission from Sun-like stars through a dynamo mechanism thought to be seated near the bottom of the stellar convection zone. In turn, this magnetic activity influences rotation itself through angular momentum loss to a magnetized wind that leads to a gradual slow down of the rate of spin.

Stars are born with a natural spread in their rotation periods and these initially evolve quite rapidly with time due to changes in moment of inertia as stars contract on to the main sequence. This initial rotational evolution then involves spin-up, rather than spin-down. All newly formed Sun-like stars are thought to possess a residual disk of gas, called a protoplanetary disk, within which planets form. In the early, so called “T Tauri,” phase of evolution (named after the representative prototype) lasting a few million years, the protoplanetary disk is expected to prevent them from spin-up through a mechanism known as disk locking (Rebull et al., 2002, 2004). While the detailed physics behind this is still poorly understood, the underpinning of the idea is that there is angular momentum exchange between the star and the disk modulated by magnetic fields that connect them—in essence, the disk applies a magnetic brake. After anything from a few Myr up to 10 Myr, the disk gets dispersed and stars then freely spin-up as a consequence of contraction. Once on the main sequence, contraction has stopped and magnetic braking through the stellar wind results in an efficient spin-down process.

Magnetic braking is determined by the magnetic fields on their surfaces (Kawaler, 1988; Weber & Davis, 1967). This self-regulating mechanism results in the rotation period evolving with time following the Skumanich law for spin-down $P_{\text{rot}} \propto t^{1/2}$ (Skumanich, 1972). This is the foundation of gyrochronology (Meibom et al., 2015), a very powerful tool that enables the conversion of rotation periods into stellar ages. Studies of the rotation periods of stars in young open clusters have revealed a bimodal distribution, recently attributed to different magnetic evolutionary paths of stars with different initial rotation periods (Garraffo et al., 2018). Stars that start off spinning faster will have smaller Rossby numbers, and this is expected to result in a more complex geometry of the surface magnetic fields. This, in turn, has the effect of closing otherwise open field lines, preventing the stellar wind to escape removing angular momentum. As a
consequence, stars with short initial rotation periods will remain rotating fast for longer than their initial slow rotators counterparts (see Figure 8 for an illustration of the effect of different initial periods in the spin evolution of a 1\,M\odot star). The period of time for which the initially fast rotators will remain rotating fast is larger the lower the stellar mass is. Eventually, at an age that depends on the stellar mass (~600 Myr for solar mass stars), initial conditions have been erased and all stars follow the Skumanich law, making gyrochronology fairly reliable. However, the activity history of these stars can be quite different depending on their initial rotation history and that can potentially make a difference in the survivability of their planets’ atmospheres and habitability.

The establishment of rotation (or more correctly, differential rotation) as the driver for the magnetic dynamo activity that gives rise to UV, EUV, and X-ray emission that drive planetary atmospheric ionization and loss processes can be traced back to the 1960s when it was noticed that Ca II H and K emission fluxes of stars declined linearly with stellar rotation velocity. The magnetic nature of stellar coronae was essentially established a decade later by the Einstein observatory, and the realization that X-ray luminosity was highly correlated with stellar rotation (Pallavicini et al., 1981; Vaiana, 1981; Walter et al., 1980). Some fraction of the magnetic energy created within the star by dynamo action and subject to buoyant rise is dissipated at the stellar surface and converted into particle acceleration and plasma heating. Although none of these processes are fully understood, the dependence of activity diagnostics and stellar UV and X-ray fluxes on rotation shows a very simple empirical relation in terms of a magnetic “Rossby” number illustrated in Figure 9. The Rossby number in this case is the ratio of the rotation period and convective turnover time near the base of the convection zone, \( \text{Ro} = P_{\text{rot}}/\tau_{\text{conv}} \) (see also Noyes et al., 1984).

Figure 9 shows stellar X-ray luminosities normalized to the total stellar bolometric output, \( L_{\text{X}}/L_{\text{bol}} \), as a function of the Rossby number for late-type stars ranging from spectral type F down to mid-M, including fully convective M dwarfs. At slower rotation rates, \( L_{\text{X}}/L_{\text{bol}} \sim \text{Ro}^2 \), where Wright et al. (2018) find \( \beta = -2.3 \), up until a threshold at which point X-ray emission saturates, \( L_{\text{X}}/L_{\text{bol}} \sim 10^{-3} \), close to a Rossby number \( \text{Ro} = 0.13 \). This saturation behavior was already apparent from data obtained by the Einstein observatory (Micela et al., 1985; Vilhu, 1984), although its origin is still debated. It is likely that it represents saturation of the dynamo itself (see, e.g., the discussion in Wright et al., 2011, and Blackman & Thomas, 2014). The rotation period at which saturation sets in increases for decreasing stellar mass. For a solar mass main sequence star, X-ray emission saturates at a ~1.25-day period, while it can be more than 100 days for an early M dwarf. (Wright et al., 2011). This means that lower mass stars are expected to be saturated and therefore more active and UV and X-ray bright than higher mass stars for much longer.

The X-ray luminosities and cumulative X-ray doses for a 1\,M\odot star as a function of age for the different rotation histories shown in Figure 8 based on the rotation-activity relations of Wright et al. (2018) are illustrated in Figure 10.

### 3.4.2. Isotopic Fractionations and Retrieving the History of Planetary Atmospheres

The study of isotopes is the major tool to study the history of planetary systems. In planetary atmospheres, it allows to create an history of the escape. Unfortunately, it is usually a ill-posed problem, and hypotheses are required, such as an atmosphere with basically the same composition over eons, and it allows to retrieve the fraction lost to space, without any
Figure 10. The X-ray and EUV luminosities, \(L_X\) and \(L_{EUV}\), for a solar mass star as a function of time. Shown are the luminosity trajectories for three different rotation evolution tracks, together with observed X-ray luminosities for single stars in open clusters. Upper limits are indicated by inverted triangles. Solid horizontal lines indicate 10th, 50th, and 90th percentiles of the observed distributions of inverted triangles. Solid horizontal lines indicate 10th, 50th, and 90th luminosities for single stars in open clusters. Upper limits are indicated by different rotation evolution tracks, together with observed X-star as a function of time. Shown are the luminosity trajectories for three.

3.4.2. Theory for Jeans/Nonthermal Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).

3.4.2.1. Theory for Hydrodynamic Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).

3.4.2.2. Theory for Jeans/Nonthermal Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).

3.4.2.1. Theory for Hydrodynamic Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).

3.4.2.2. Theory for Jeans/Nonthermal Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).

3.4.2.1. Theory for Hydrodynamic Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).

3.4.2.2. Theory for Jeans/Nonthermal Fractionation

When discussing fractionation with respect to an escaping atmosphere, the Rayleigh fractionation/distillation law and its notation are often used (Jakosky, 1994; Johnson et al., 2000; Mandt et al., 2009). In this nomenclature, \(R\), the ratio of two species—often isotopes—is the main parameter.

First, it is important to make a confusion between the observed ratio of species \(b\) and \(a\), \(R(\tau) = n_b(\tau)/n_a(\tau)\), and the total ratio \(R = N_b(z)/N_a(z) = N_b(z_\text{homopause})/N_a(z_\text{homopause})\). The observed ratio is fractionated in altitude above the homopause. The main hypothesis of the Rayleigh distillation law is that any species is lost proportionally to its total amount: \(dN_i = k_i N_i\). Therefore, we have

\[
\frac{dN_b}{N_b} = \frac{k_b N_b}{k_a N_a} = \frac{f N_b}{N_a}
\]

Solving this equation allows to evaluate the total amount of \(b\) that escaped. The most interesting conclusion that we can make from this equation, while having no knowledge of \(f(t)\), is that the escape of \(N_b\) stops at the time \(t_2\) such that \(f(t_2) = \frac{m_b - m_a}{m_b^0 - m_a} = \frac{m_b}{m_b^0}\).
3.4.2.3. Outstanding Problems Related to Fractionation

The observation and modeling of fractionation highlighted major events in the evolution of planetary atmospheres. The D/H ratio observed at Venus, $>1.6 \times 10^{-2}$ (Donahue et al., 1982; Marcq et al., 2018), suggests that the loss of water at Venus may be consistent with the loss of Earth’s ocean (Shizgal & Arkos, 1996).

The enrichment of $^{15}$N over $^{14}$N at Mars can be explained by nonthermal processes (Shizgal & Arkos, 1996).

Since Ar does not react chemically, Jakosky et al. (2017) used the $^{38}$Ar/$^{36}$Ar ratio to determine that Mars lost 66% of its atmosphere to space. It is to be noted that the location and timing at which an isotopic ratio is measured can have an effect: Livengood et al. (2020) have shown that surface adsorption at Mars performs isotopic fractionation that can be highlighted by daily variations (temperature changes the amount of gas adsorbed).

At Earth, the fractionation of noble gases has been explained by hydrodynamic escape processes (Shizgal & Arkos, 1996), except for xenon. Xenon is depleted by 1 order of magnitude relative to other noble gases and other volatile elements when normalized to the chondritic composition (e.g., Marty, 2012) and is largely enriched in its heavy isotopes relatively to solar or chondritic xenon. This peculiarity of xenon compared to other noble gases is known as the “xenon paradox.” The specific electronic structure of xenon makes it the most reactive element among noble gases with the lowest ionization potential (12.13 eV or 102.23 nm) and an extended photoabsorption cross section covering part of the vacuum ultraviolet (VUV) spectrum (up to about 150 nm). From this consideration and because of the difficulty to explain Xe depletion and fractionation with other mechanisms, the escape of Xe$^+$ driven by H$^+$ ion escape is considered as a plausible explanation of the xenon paradox (Zahnle et al., 2013). On the other hand, Hébrard and Marty (2014) proposed a scenario combining the trapping of heavy xenon isotopes in haze with an efficient escape of Xe$^+$ ions, which are both consistent with the Xenon depletion and fractionation. The time history of Xe isotope fractionation has been investigated in detail by Avice et al. (2018) and showed that it started evolving at least 3.5 Ga ago until it reached the modern-like atmospheric Xe composition at around 2.1 Ga ago. They concluded that termination of the isotopic fractionation of Xe may coincide with the end of the hydrogen escape, which has previously been suggested to explain the progressive oxygenation of the Earth’s atmosphere (Zahnle et al., 2013). However, such significant escape of Xe$^+$ ions with no associated loss of other noble gases is challenging due to the large mass of Xe$^+$ and its associated large gravitational binding energy ($\sim 85$ eV). The work of Parai and Mukhopadhyay (2018) shows that xenon can be trapped in the crust from the oceans but is basing its atmospheric fractionation on the work of Pepin (1991) who does not consider that xenon can be trapped back into the Earth. It may be that the xenon paradox could be solved by escape physics but by crustal absorption. This is in agreement with the organic haze scavenging of Xe hypothesis, investigated in Avice et al. (2018), that solves the paradox without requiring atmospheric escape. An interesting point is the requirement for atmospheric hazes in that hypothesis: Specific atmospheric conditions, similar to the present Titan (McKay et al., 2001), are required to create those, and therefore Xe isotopic fractionation may be an indicator of the atmospheric conditions of the Archaean Earth. Such a solution has the potential to reconcile the conclusions of Parai and Mukhopadhyay (2018), that is, no plate tectonics before 2.5 Gyr ago or extremely dry tectonics, with the observation of ancient plate tectonics with excess water before 3.3 Gyr by Sobolev et al. (2019).

4. Escape at Solar System’s Planets and Bodies

The atmospheric escape in the solar system shaped most of the planets and dwarf planets’ atmospheres, as well as those of some satellites. Mercury may have had a protoatmosphere above its magma ocean, like the Moon (Greenwood et al., 2018), but, unlike the Moon, no sample from the surface are available and a detailed study of the history of Mercury’s atmosphere lacks too much experimental evidence. Mars, Venus, the Earth, and Titan have on the contrary a large quantity of data showing an atmosphere that has been transformed by escape. Currently, H and He are the most important species escaping for these objects. We are also observing other escape processes, such as O, and we try to understand the pathways of escape of CO$_2$ at Mars, so we can understand the evolution of its atmosphere in time. Nonthermal escape processes at Jupiter and Saturn are known to fill a part of their plasmaspheres and lead to some minor escape. Overall, the giant planets are too big for efficient escape to take part and change drastically their atmospheric evolution. Uranus and Neptune are similar in that their mass prevents a lot of escape. In
addition, the ice giants have only been visited by the Voyager probes and have not had a Galileo or Cassini-like mission allowing study of their atmospheres as comprehensively as Jupiter or Saturn. A recent work by DiBraccio and Gershman (2019) has shown that, as for Jupiter and Saturn, plasmoids have been observed at Uranus. However, a large quantity of these plasmoids’ loss can come from the satellites of the giant planets. Finally, relatively small loss are suspected to come from polar wind (Glocer et al., 2007).

Understanding the current escape processes allows to perform some interpolation back in time, thanks to a better understanding of the conditions (especially the solar forcing in time) and of the important parameters for each escape process. Such work necessary to understand how the atmosphere could evolve into an habitable one. To have a control point to these interpolations, it is necessary to know about the isotope ratio. Unfortunately, surface processes (volcanism and adsorption) and other events such as comet falls or cosmic ray spallation processes (e.g., Pavlov et al., 2014) can affect these results and possible concurrent models leads to the currently observed state.

4.1. The Solar Forcing in Time

The solar magnetic activity forcing changed substantially over the lifetime of the Sun, as discussed in the general stellar context in section 3.4.1. In its early stages, the rotation of the young Sun was faster than its present rotation, with rotation periods of only few days compared to its present 27-day period. During this time, the Sun’s ratio of X-ray to bolometric luminosity declined by a factor of about 1,000. Thus, in general, the most important aspect of the evolution of solar forcing is that the EUV and X-ray fluxes were much higher during the early solar system than today, with related consequences for earlier planetary atmospheric escape rates. This time evolution is shown in Figure 10.

In addition to the solar radiation, the solar wind also plays a role in planetary atmospheric escape. However, the change of the solar wind in time is much less well defined than the solar radiation, since it is almost impossible to measure the signatures of weak winds of solar analogs. In general, the magnetic activity paradigm suggests that a more active Sun should produce a stronger solar wind. However, no clear evidence of that assumption has been discovered so far. Scaling laws have been developed based on observations of the neutral Hydrogen absorption line generated at the edge of stellar astrospheres, where the stellar wind collides with the interstellar medium (ISM) (Wood, 2006; Wood et al., 2014). Some modeling work was done to characterize the winds of Sun-like stars (e.g., Cohen & Drake, 2014; Cohen et al., 2010). In both cases, the winds of young stars were not found to be dramatically stronger than older stars, and many of the observed systems were found to deviate from the scaling low and present weaker winds than expected.

Understanding the young Sun is important in the context of solving the FYS Paradox (section 7.1.1). However, it is important to note that if the hypothesis on the atmosphere composition and pressure is off, no conclusion can be made only from the star parameters.

4.2. Coupling With the World Below

The escaping region of an atmosphere does not exist in isolation. Once the primordial atmosphere of a telluric planet has eroded, the hydrogen and other light elements that escape ultimately come from the interior of the planet and pass through the lower/middle atmosphere to reach the region of escape. Just to understand hydrogen escape, it is crucial to understand the basic processes that control: (1) the flux of hydrogen-bearing species (e.g., water, methane, H₂) from their source regions at the surface throughout the atmosphere and (2) the flux of hydrogen-bearing species to the atmosphere.

Accurate modeling of hydrogen-bearing species are particularly important for understanding potentially habitable exoplanets. First, methane is a potential biosignature. Second, the chemistry of these species is often connected with the chemistry of O₂, O₃, CO₂, and CO through the HOₓ reactions in the atmosphere and through analogous reactions in the interior (e.g., Kasting et al., 1993). Third, sufficient hydrogen escape can modify the redox state of the atmosphere, surface, and interior (e.g., Catling & Claire, 2005; Kasting et al., 1993).

4.2.1. The Present: A Focus on Hydrogen-Bearing Species

The vertical temperature structure of the atmosphere is a critical control on the ability of condensable species to move upward from the warm troposphere to where it can escape. The stratospheric water trap on present-day Earth is the classical example of such a control. Transport across such barriers may be
accomplished diffusively or dynamically by means of atmospheric moist convection. The effectiveness of the cold trap for water may depend on the presence of other hydrogen-bearing species, such as methane.

4.2.1.1. Earth

On present-day Earth, the dominant sources of hydrogen-bearing species are evaporation from the Earth’s oceans (H$_2$O) and anthropogenic sources of methane. Nonanthropogenic, biogenic sources of methane remain significant and probably greatly exceed geological sources (Dlugokencky et al., 2011), though Etiope and Klusman (2002) argue that natural geological sources may be currently accounted to anthropogenic emissions in error. Atmospheric water poorly mixes into the middle atmosphere. There is a strong contrast between water vapor mixing ratios typical of the troposphere (≈1,000 ppmm) and water vapor mixing ratios near the Earth’s mesopause (≈5 ppmm), where the photodissociation of water by solar radiation at Lyman α wavelengths takes place (Roell, 2012). A parallel contrast exists between water vapor concentrations near the surface and in the upper troposphere (≈10,000 ppmm vs. ≈100 ppmm in the tropics: Sun & Lindzen, 1993). The region of rapid fall-off in water vapor mixing ratio is known as the hygropause.

The contrast in humidity between the troposphere and the mesosphere results from the large-scale temperature structure of the atmosphere, in which the atmospheric temperature minimum is at the tropopause and lower stratosphere. Any excess water beyond the point of saturation will condense to liquid and ice, which may precipitate. Thus, moist air is freeze-dried to the equilibrium water vapor concentration at the ambient temperature. At the temperature minimum of the tropopause and the lower stratosphere, an “atmospheric cold trap” forms. The contrast in humidity between the surface and the upper troposphere partly arises from the same mechanism. Therefore, relatively slow vertical mixing of water vapor by large-scale processes such as the Hadley cell or synoptic-scale systems will be set by the vertical thermal structure of the atmosphere, which radiative-convective models can estimate approximately (Pavlov et al., 2000).

Mesoscale processes also have some impact on water vapor transport. Strong vertical motions in buoyant moist convection can transport ice to higher altitudes, evading “cold trap” effects. In some cases, moist convection can “overshoot” the tropopause, injecting large amounts of water ice into the stratosphere over an areally limited region. If this water ice sublimates in the stratosphere, the stratosphere is hydrated locally by the same order as the background stratospheric water vapor concentration (Grosvenor et al., 2007; Liu et al., 2010). If overshooting moist convection were more intense and/or more efficient at transporting water ice to the stratosphere, the contrast between tropospheric and mesospheric water vapor concentrations could be reduced. (We assume that the increase in water vapor due to overshooting convection is greater than the decrease in water vapor due to mixing resulting from downdrafts of overshooting convection.)

Methane does not condense at atmospheric temperatures, which reduces the surface-mesosphere contrast considerably (1.8 ppm vs. 0.1 ppm) (Summers et al., 1997). It is slowly dissociated in the stratosphere and mesosphere, so its composition in the upper atmosphere will be controlled by the relative balance between chemistry and vertical transport in the stratosphere and mesosphere as well as the intensity of stratospheric-tropospheric exchange.

4.2.1.2. Mars

On present-day Mars, the vertical structure of water vapor differs greatly from that of the Earth. The main source and sink of water vapor is sublimation from and condensation on the polar caps. A seasonally varying hygropause is apparent at a characteristic height of 40 km above the surface in the tropics (Clancy et al., 2017; Heavens et al., 2018). Yet detached water vapor layers are frequently observed as high as 80–90 km above the surface (Maltagliati et al., 2013). This structure partly reflects differences in the temperature structure. The tropical middle atmosphere is not separated from the lower atmosphere by a strong thermal inversion analogous to the stratosphere (except perhaps in global dust storms, when the entire atmosphere is effectively inflated by the heating of dust). Thus, to first order, the depth of the Hadley cell sets the hygropause height (Richardson & Wilson, 2002). Detached water vapor layers originate from mesoscale transport processes, such as injection within dust plumes in Mars’s well-known dust storms (Fedorova et al., 2018; Heavens et al., 2011, 2015, 2018; Maltagliati et al., 2013; Spiga et al., 2013) or associated with topographically driven circulations, with or without dust storm activity (Heavens et al., 2015, 2018; Michaels et al., 2006; Rafkin et al., 2002). In regional and global dust storms, convective transport of water to the middle atmosphere within dusty air can be so strong that we cannot really speak of detached water vapor layers;
the mean hygropause of the planet can rise to 80 km (a change mostly caused by ascent in the tropical hygropause) (Fedorova et al., 2018; Heavens et al., 2018).

While a variety of observations suggest the presence of atmospheric methane, sufficiently little is known about it to make discussion of its surface sources and transport to the upper atmosphere entirely speculative (Formisano et al., 2004; Mumma et al., 2009; Webster et al., 2015).

4.2.1.3. Venus

On present-day Venus, the main sources of water to the atmosphere are believed to be cometary and meteoritic impacts and volcanic outgassing in uncertain proportions (Taylor & Grinspoon, 2009). The possibility of ongoing volcanic outgassing has been bolstered by observations of temporal and spatial variability in atmospheric SO2 and transient NIR emission from a prominent rift zone (Marcq et al., 2013; Shalygin et al., 2015).

Estimated lower atmospheric (5–45 km) water vapor concentrations from spectroscopic observations range from 25–50 ppmv with typical uncertainties at the 20% level (Chamberlain et al., 2013). There is an outlier estimate of 200 ppmv in the 30- to 45-km altitude range (Bell et al., 1991). While this value is consistent with some in situ measurements by entry probes, the entry probe data are mutually inconsistent and generally mistrusted (Meadows & Crisp, 1996). Current observations are unable to probe water vapor concentrations within 5 km of the surface, but it is speculated that water vapor might be depleted near the surface because of reactions with surface rocks (Chamberlain et al., 2013; Fegley, 2003).

Water vapor concentrations above the troposphere are 3–11 ppmv near the top of the sulfuric acid cloud deck at 60- to 70-km altitude and likely decrease to 1 ppmv at 100 km (Fedorova et al., 2008, 2016). Water vapor at these altitudes would be vulnerable to photochemical loss processes. Water vapor in the middle atmosphere is most abundant near the equator, a phenomenon that suggests convective transport of water vapor from the lower atmosphere into the middle atmosphere (Fedorova et al., 2016). Otherwise, the water vapor distribution in the middle atmosphere appears quite sensitive to the altitude of the cloud deck, suggesting that the sulfuric acid cloud deck is an effective hygropause for Venus due to the formation of sulfuric acid from H2O and SO2.

Venus's atmosphere does not contain measurable amounts of methane at present measurement sensitivities (Taylor & Grinspoon, 2009). Early in situ measurements by Pioneer Venus suggested atmospheric methane concentrations were up to 6,000 ppmv, but these measurements likely were contaminated by reactions within the measurement apparatus itself (Donahue & Hodges, 1993). Yet some methane input from meteoritic and cometary sources is possible. If the mantle of Venus has remained sufficiently reducing, a source of methane from volcanic outgassing is possible as well.

4.2.1.4. Titan

The principal hydrogen-bearing species in the present-day atmosphere of Titan is CH4. The total amount in the atmosphere as vapor exceeds the amount present on the surface as liquid by at least a factor of 2 (Lorenz et al., 2008). Most hydrogen-bearing species in Titan's atmosphere, such as H2 and various organic compounds, are likely derived from photochemical reactions involving CH4 (Krasnopolsky, 2014; Owen & Niemann, 2009; Strobel, 2012), though an H2 source derived from serpentinization also has been proposed (e.g., Atreya et al., 2006). An exception is H2O, which is likely supplied to Titan by ablation of micrometeorites and/or plume material from Enceladus (Coustenis et al., 1998; Dobrijevic et al., 2014; Lara et al., 1996). The ultimate source of methane on Titan is believed to be episodic outgassing from Titan's deep interior (Lorenz et al., 1997; Tobie et al., 2009; Wong et al., 2015).

CH4 concentrations near the surface are ≈50,000 ppmv, decrease to ≈15,000 ppmv in Titan's stratosphere (above 32 km) (Niemann et al., 2005), and remain uniformly mixed at that up to altitudes near the homopause region ≈850–1,000 km where diffusive separation causes the relative fraction of methane to increase with altitude up to the exobase (Bell et al., 2014; Johnson et al., 2010; Yelle et al., 2008). The major barrier to transport is an atmospheric cold trap occasionally broken by deep convective clouds of CH4 (Griffith, 2009).

4.2.2. The Past: Coupling, Unusual Escape Regimes, and Current Evidence for Atmospheric Mass and Upper Atmospheric Composition

4.2.2.1. Earth

In the course of the Earth's early history, the upper mantle was gradually oxidized by means of coupled pathways: (1) Reductants in the upper mantle were emitted into the atmosphere by volcanic processes, were
transported into the upper atmosphere by processes at mesoscale to planetary scales, and escaped the Earth system by a mixture of physical and chemical processes in the upper atmosphere, resulting in an unbalanced loss of reducing power from the upper mantle; and (2) weathered (hydrated) or even oxidized crust was recycled into the mantle, where water reacted with mantle material to form hydrogen, a light, readily escaping reductant (Holland, 1984). This period of mantle oxidation closed when the mantle was sufficiently oxidized that water-hydrogen conversion in the mantle ceased (Kasting et al., 1993).

These processes are mediated by escape itself. Water that is vertically transported into the upper atmosphere photodissociates, leading to the production of H2 and O2 abiotically (Lewis et al., 1983). The H2 is highly vulnerable to escape, transporting reducing power out of the system, while the O2 may mix down into the lower atmosphere and oxidize the crust. Simultaneously, precipitation of atmospheric water to the surface hydrates the crust.

Planets like present-day Earth exchange water with the mantle in the course of plate tectonic processes such as subduction, which are relatively efficient. If Venus’ surface were wetter, some exchange would take place during putative resurfacing events (Strom et al., 1994). Similar speculations on might be made about Mars’ volcanic activity. However, crustal recycling on both Mars and Venus are thought to be much weaker than the Earth and, on average, weaker in the past (Taylor & McLennan, 2009).

However, there are signs from the extant record of early Earth history that plate tectonics may not be the upper limit for crustal recycling rates on the Earth. Instead, the Earth may have experienced a “heat-pipe” phase (Moore & Webb, 2013). In this phase, persistent mafic to ultramafic volcanism regularly resurfaced the Earth. Both crustal material and surface water are cycled back into the mantle through repeated eruption and burial of older flows. The heat pipes were associated with greater eruptive volumes of volcanic material as well as faster crustal recycling than plate tectonics. Observations of the other terrestrial bodies in the solar system are also consistent with heat pipe operation in their early phases (Moore et al., 2017).

Therefore, the upper mantle of the Earth began in a far more reduced state than today and may have degassed far more intensely than today. An important consequence of the reduced state of the Earth’s early mantle (without enhanced degassing) would have been higher proportional degassing of carbon from the interior in the form of CH4 as opposed to CO2. In addition, formation of H2 from H2O in the mantle would have resulted in significant emission of H2 to the atmosphere (Kasting et al., 1993). High concentrations of atmospheric H2 would have interrupted the OH radical pathway for CH4 oxidation. Both gases have a demonstrable greenhouse effect (Pavlov et al., 2000; Wordsworth & Pierrehumbert, 2013a, which enhances near-surface water vapor abundance. And absorption of visible/NIR radiation by CH4 strongly heats the lower stratosphere (Pavlov et al., 2000). Based on radiative-convective simulations with variable CH4 values, Pavlov et al. (2000) argued that there would be a direct relationship between a more reduced mantle, a warmer “cold trap,” deeper vertical mixing of water vapor, weaker contrast in water vapor concentrations across the hygropause, enhanced water vapor photoysis, and oxygen production in the upper atmosphere.

As in the present day, CH4, unlike H2O, would not condense at Earth atmosphere temperatures. Therefore, under a variety of “cold trap” conditions, CH4 from the putative reduced mantle source would diffuse or advect beyond the hygropause to altitudes at which it will photolyze to produce H2 (but not O2).

Speculations that H2O and CH4 would react to form an organic haze (e.g., C2H2 and C2H4), which would oppose any CH4 or H2 greenhouse effect, challenge this picture (Haqq-Misra et al., 2008; Pavlov et al., 2001). However, recent simulations of this aerosol suggest that it would be optically thin in the visible but optically thick in the UV. The haze would have little effect on greenhouse warming but shield the atmosphere below it from photolysis (Wolf & Toon, 2010) (and may affect the Xenon isotopic ratio, section 3.4.2.3).

At the same time, Pavlov et al.’s (2000) simulations greatly simplify quantitative treatment of chemistry, transport, and hydrogen escape. As is noted, “[a figure which shows the relationship between CH4 flux and atmospheric concentration] is somewhat deceptive in that it implies that atmospheric CH4 concentrations can be calculated by specifying the surface CH4 flux.” One example of a complication is that higher water vapor abundance in the atmosphere will reduce the atmospheric concentration of CH4 by providing a demonstrable greenhouse effect (Pavlov et al., 2000; Wordsworth & Pierrehumbert, 2013a), which enhances the warming.
Another distinct type of hydrogen escape regime might occur if the entire surface were glaciated, as is speculated to have occurred during portions of Paleoproterozoic and Neoproterozoic time on the Earth. This regime has been invoked to explain mysterious rises in atmospheric $p_{O_2}$ (O$_2$ partial pressure) during the deglaciation from Snowball events.

$p_{O_2}$ climbed to 10% or even close to present-level $p_{O_2}$ in the aftermath of the Paleoproterozoic Snowball (Barley et al., 2005). However, the connection between an increase in atmospheric oxygen and the Paleoproterozoic glaciation is disputed (Hoffman, 2013). It is entirely possible that the glaciations preceded the rise in oxygen by $\approx$100 million years. $p_{O_2}$ then dropped to 1–3% before rising again to 5–18% in the Neoproterozoic, a time when the connection between deglaciation and the oxygen rise is better established (Canfield et al., 2007, 2008; Fike et al., 2006; Halverson et al., 2009; Sahoo et al., 2012; Scott et al., 2008). This higher level of $p_{O_2}$ coincided with the first appearance of metazoans in the rock record at around 600 Ma (Canfield et al., 2007) and with the end of the Cryogenian era of Snowball Earth glaciations.

Simple models of atmospheric chemistry suggest that the unusually cold conditions of an entirely ice-covered Earth would favor the production of H$_2$O$_2$ in the atmosphere (Liang et al., 2006). As in present-day Antarctica, this H$_2$O$_2$ would be incorporated into ice. During deglaciation, this H$_2$O$_2$ would enter an ocean rich with Fe and Mn, poisoning existing anaerobic organisms while creating the selection pressure for the development of oxygenic photosynthesis (Kopp et al., 2005; Liang et al., 2006). Oxygen limitation thereafter then would be due to oxidation and precipitation of Fe and Mn outcompeting oxygen production by nitrogen-limited early photosynthetic autotrophs (Liang et al., 2006). The dependence of this mechanism on the composition and emission rate of mantle effluents (and thus geological activity/upper mantle oxidation state) is unknown. And this question does figure in interpreting and extrapolating from the Snowballs, because the Earth’s mantle was likely more oxidized during Neoproterozoic time than Paleoproterozoic time.

An additional variable to consider for the past is the identity and abundance of the principal atmospheric constituent (N$_2$ at present). Models consistent with abundant liquid water that assume a FYS either assume higher atmospheric pressure from gases such as N$_2$ (e.g., Goldblatt et al., 2009) or assume higher concentrations of greenhouse gases such as CH$_4$, H$_2$, and CO$_2$, whose ability to warm climate is strongly dependent on pressure broadening (e.g., Kasting et al., 1984). Higher atmospheric pressure also can reduce surface temperature as a result of increased molecular scattering of incoming solar radiation (Goldblatt et al., 2009; Poulsen et al., 2015). However, the sign of the net effect is unclear. Radiative-convective model simulations suggest that increased N$_2$ or O$_2$ will result in net positive radiative forcing even at $p_{CO_2}$ much less than at present (Goldblatt et al., 2009; Payne et al., 2016). Simulations with a GCM that included clouds suggest that the net radiative forcing can be negative as a result of cloud feedback effects at higher atmospheric pressure (Poulsen et al., 2015).

Data from the geological record about past atmospheric pressure has wide uncertainties but may argue against the Earth’s atmosphere being much thicker in the Archean. A recent study of gas bubbles in an Archean (2.7 Ga) lava flow near paleosea level by Som et al. (2016) suggests that the Earth’s atmospheric pressure was no more than 50% of present and most likely $\approx$25% of present at that time. Raindrop-based reconstructions also have been attempted. Som et al. (2012) suggested an upper bound for atmospheric density of approximately twice present, but Kavanagh and Goldblatt (2015) argued that raindrop size was more sensitive to rainfall rate than atmospheric pressure and suggested an upper bound for atmospheric density of approximately 11 times present. It is to be noted that the work of Airapetian et al. (2016) and its extension (Gronoff et al., in preparation) consider an alteration of the atmospheric chemistry by SEP events to create N$_2$O, which increase the temperature of the early Earth even for atmospheric pressure lower than 0.5 bars.

Modeling suggests that if were possible to keep liquid water stable in a low pressure N$_2$ atmosphere (200 hPa), water transport by moist convection to the middle and upper atmosphere would be extremely efficient, resulting in high rates of water photolysis (Kleinböhl et al., 2018). The resulting atmosphere evolves to a state in which abiotic oxygen dominates the atmosphere, unless there is a strong sink of oxygen at the surface (Kleinböhl et al., 2018). Such a mechanism could explain bursts of oxygenation coincident with the formation of banded iron formations, but the model relies on a one-dimensional parameterization of moist
convective adjustment (Kasting, 1988) that requires testing in a framework that more explicitly resolves the physical processes.

### 4.2.2.2. The Moon

The exosphere of the Moon is interesting in several ways: (1) It is easier to experiment on it: We can study the decay of artificial gases released on it by lunar lander in function of the solar activity (Vondrak, 1974, 1992; Vondrak et al., 1974); (2) it has the same solar wind conditions as the one measured for space weather at Earth, and therefore, studies such as the impact of CME on it are easier (Killen et al., 2012); and (3) we have samples from the Moon, and we can study the possibilities of ancient atmosphere from it.

The hypothesis of a secondary atmosphere due to volcanic activity at the moon has been proposed in Needham and Kring (2017) based on the analysis of samples from the Apollo mission. It is possible that an an atmosphere with up to a few mb at the surface was created and stable for thousands of years. In Aleinov et al. (2019), a study of the thermal escape was made, showing the limitations of the creation of such an atmosphere, as well as the climatic conditions an atmosphere would have had. These conditions are interesting since they show the transport of volatiles to the poles. it would be possible to find some clues of that atmosphere in samples from the poles.

### 4.2.2.3. Mars

When Mars had an intrinsic magnetic field early in its history, its hydrogen-bearing species fluxes to the upper atmosphere likely occupied a phase space that could be described by the early Earth or even present-day Earth phase spaces (Alho et al., 2015). Transition to the regime observed today may have depended on the timing of magnetic field loss. This transition has so far been modeled as a primarily CO₂ atmosphere condensing to form at least one permanent ice cap (Soto et al., 2015, and references therein). The principal unknown about the last billion years or so is how fluctuations in Mars' obliquity have changed the location of surface and subsurface ice reservoirs, which could affect the water cycle, the total atmospheric mass, and the dust cycle (Fastook et al., 2008; Madeleine et al., 2009). A lot of questions have also been asked about the effect of the magnetic field in the loss of the atmosphere. Since observation shows that similar amount of heavy ions are lost above magnetic fields at the current Mars than above nonmagnetized parts (Sakai et al., 2018), it may be that it influence has been greatly exaggerated in previous studies.

Like Earth, there are some constraints on past atmospheric mass for Mars. The atmosphere filters the impact crater population by ablating the lower end of the bolide size distribution (Jakosky et al., 2017). On this basis, Kite et al. (2014) proposed that Martian paleopressure was never higher than \( \approx 3 \) bar (and likely much less). A higher palopressure would have led to a collapse of the atmosphere. From meteoritic observation constraints, and considering that some isotopic reservoirs can be replenished by meteoritic/cometic falls, Kurokawa et al. (2018) slightly modified the history presented in Jakosky et al. (2017) and suggested a minimum paleopressure of 0.5 bar. Jakosky et al. (2018) suggested that Mars lost more than 0.8 bar of CO₂ or the equivalent of 28 m of water.

An interesting point at Mars is the observation of solar wind H deposition in the thermosphere (Halekas et al., 2015); this deposition follows a charge-exchange process and could have led to changes in D/H ratio if large enough in the early solar system; however, it is probable that this deposition would have been countered by hydrodynamic escape.

### 4.2.2.4. Venus

Venus, at some point during its history, likely occupied an additional phase space with respect to coupling between the surface and the exosphere: that of the runaway greenhouse (Ingersoll, 1969). However, this regime is somewhat analogous to the elimination of the “cold trap” by absorption of visible/NIR radiation by CH₄. The twist is that it is the IR greenhouse effect of H₂O that breaks the cold trap.

The effect can be conceptualized semiquantitatively. Consider a layer of the atmosphere at which vertical mixing from the surface is relatively efficient. Now raise the surface temperature by some amount by introducing a higher amount of solar insulation. To first order, the relationship between water vapor concentration and temperature should be exponential, following the Clausius-Clapeyron relation that defines the saturation curve. In the Earth's atmosphere, however, it is observed that the effects of vertical mixing and pseudo-adiabatic precipitation processes reduces the sensitivity of mean \( p_{\text{H}_2\text{O}} \) to surface temperature in the lower troposphere (Held & Soden, 2006) but may enhance it in the upper troposphere (Gettelman & Fu, 2008). Thus, water vapor concentration will increase exponentially in response to the increase in
surface temperature. The layer’s temperature likely will increase as well in response to the increase in surface temperature. (This is easiest to visualize at the surface itself.)

At the same time, the increase in water vapor will increase the IR opacity of the layer, reducing outgoing longwave radiation from the layer (and below the layer). Yet the increase in the layer’s temperature will result in increased outgoing longwave radiation according to Stefan-Boltzmann’s law. At low temperatures and water concentrations, it is easy to see that the principal change in outgoing longwave radiation will be due to the increase in layer temperature. However, as temperatures increase, the exponential dependence of water vapor on temperature eventually will overcome the quartic dependence of outgoing longwave radiation on temperature. Thus, for any sufficiently abundant IR absorber condensing and evaporating, there is some critical point at which outgoing longwave radiation in the layer will decrease rather than increase with surface temperature, initiating a runaway positive feedback loop. Warming of the troposphere eventually results in its expansion and enhancement of vertical transport in the middle and upper atmosphere. For water, this runaway loop is slowed by UV hydrolysis of water in the middle and upper atmosphere and stopped by exhaustion of the surface reservoir, a process that Ingersoll (1969) argued had occurred on Venus (rather than Earth or Mars) as a result of the former’s higher insulation.

This escape regime has been simulated by Kasting and Pollack (1983), Kumar et al. (1983), and Chassefière (1996a, 1996b). None of these simulations challenge the basic mechanism but emphasize (1) that hydrolysis rates will be dependent on the oxidation state of the atmosphere and buffering by chemical reactions in the crust and (2) that the EUV flux of the Sun (a major unknown early in its lifetime) is the principal control on the rate of escape. Another interesting conclusion is that the present D/H ratio in Venus’ atmosphere must be a consequence of a period of reduced escape rates that closed the runaway greenhouse phase. In the ideal runaway greenhouse escape regime for Venus, D would have been stripped off as easily as H (Kasting & Pollack, 1983, 1996a). The question of when that runaway escape happened is difficult as it was suggested that Venus could have been able to sustain liquid water up to a approximately Gyr ago (Way et al., 2016).

4.2.2.5. Titan

The large size of the atmospheric reservoir of methane in comparison with the surface reservoir of methane and methane’s photochemical products (Lorenz et al., 2008) strongly suggests that a methane-rich atmosphere for Titan has been a relatively unusual condition during Titan’s history (Lorenz et al., 1997). Once a sufficient amount of time has passed, photochemistry will refine methane to organic compounds that will form surface deposits of liquid and solid higher order hydrocarbons. The resulting atmosphere will lose the portion of its greenhouse effect driven by pressure broadening of methane, and Titan will lose its stratosphere (Wong et al., 2015). Any hydrogen escape presumably will be restricted to photochemical loss of water derived from micrometeorite ablation, etc.

Yet the presence of CH$_4$ in Titan’s atmosphere likewise implies occasional, episodic release of methane into the atmosphere by volcanism (Tobie et al., 2009). Depending on the exact nature of this volcanism, Titan could have experienced a more intense hydrogen escape regime in the past.

5. Escape at Exoplanets

Since their first detections around stars in the middle to late 1990s (Mayor & Queloz, 1995), a particular interest has been set to the atmospheric escape of exoplanets. In particular, the intense heating and radiation at close-in orbit planets, such as the planets orbiting M dwarfs in the HZ, or the giants close to their host stars (the so-called ‘hot-Jupiters’ that we name close-in giant in the following since the nature of, notably, their atmospheric escape cannot be considered as Jupiter-like), may lead to very high atmospheric mass loss rate and potentially a complete evaporation of the planetary atmosphere (in addition to potential atmospheric stripping by the stellar wind) (e.g., Cohen et al., 2015; Lammer et al., 2003).

5.1. Current Observations and Modeling

5.1.1. Close-In Giants

Observations of the close-in giant planet HD 209458 have revealed absorption in the Lyman α line, which associated with the existence of neutral Hydrogen (H I) at or above the estimated Rosch lobe (Vidal-Madjar et al., 2003, 2004). The fractional difference of in- and out-of-transit flux was wavelength dependent, with much of the flux decrease occurring at wavelengths more than 100 km s$^{-1}$ from line center. Though the atom-photon cross section is much larger at line center, and one would expect much larger transit depths
there, interstellar absorption and geocoronal emission contaminate wavelengths \(<50 \text{ km s}^{-1}\) from line center, and only measurements further from line center may be trusted. Later observations have indicated the existence of heavier atoms, such as Ca II and O I at this altitude (Ehrenreich et al., 2008; Linsky et al., 2010). These observations suggest that the planet has an inflated atmosphere with a high mass loss rate of the order of \(10^7 \text{ kg s}^{-1}\), and an escape velocity of the order of \(100 \text{ km s}^{-1}\).

The current paradigm assumes that close-in giants lose mass from their atmospheres due to hydrodynamic escape (Ben-Jaffel, 2007, 2008; Linsky et al., 2010; Vidal-Madjar et al., 2008). However, it is not obvious what is the mass loss rate and the escape speed, what is the altitude of the observations, and what is the overall structure of the inflated atmosphere. It has also been suggested that due to the fast orbital motion, the extended atmosphere may have a comet-like tail (Cohen et al., 2011; Linsky et al., 2010).

A number of models have been developed to study atmospheric escape from close-in giants (Adams, 2011; Baraffe et al., 2004; García Muñoz, 2007; Koskinen et al., 2014; Lecavelier Des Etangs, 2007; Murray-Clay et al., 2009; Penz et al., 2008; Schneider et al., 2007; Stone & Proga, 2009; Tian, 2009; Trammell et al., 2011; Yelle, 2004), where most of the models assumed that the intense hydrodynamic escape is due to photoevaporation by the intense stellar radiation. The models listed above (partial list) vary in the equations they solve, their assumptions about the energy sources and distributions, their complexity, and the way they are solved. The mass loss rate obtained by these models covers few orders of magnitude. Therefore, despite of the vast modeling effort, the nature of atmospheric escape from these close-in giants is not fully understood yet. The efforts by Tanaka et al. (2014, 2015) to model atmospheric escape from close-in giants have to be noticed since the model is based on MHD wave heating leading to ionospheric outflow. This is a case of the more general ionospheric outflow described in section 2.4.

In the case of HD 209458b, the implies the existence of a large “corona” or “cloud” of atomic hydrogen. This cloud must be optically thick to Lyman-\(\alpha\) at wavelengths \(>100 \text{ km s}^{-1}\) from line center out to several (optical continuum) planetary radii, approaching the planet’s Hill radius, beyond which stellar tides dominate over the planet’s gravity. There are two models to account for this large hydrogen density at such high altitudes.

The first model (Yelle, 2004) is that the absorption is due to thermal particles in the planet’s upper atmosphere. Photoelectric heating from hydrogen ionization, balanced by slow adiabatic expansion, raises the temperature to \(T \sim 10^4 \text{ K}\). The resulting large scale height implies a slow outward decrease of the density and hence large density at high altitude. In this model, the thermal speed of the atoms is \(v_{\text{th}} \sim 10 \text{ km s}^{-1}\) and absorption at \(>100 \text{ km s}^{-1}\) implies a large column of hydrogen is needed to overcome the small cross section at \(>10 \text{ Doppler widths from line center}\).

The second model (Holmstrøm et al., 2008) relies on fast hydrogen atoms (ENA), which must move at speeds comparable to the line width. The large atomic speeds imply that vastly smaller columns are needed to attain optical depth unity. The thermal hydrogen speeds, and bulk velocity in hydrodynamic escape, are expected to be only \(~10 \text{ km s}^{-1}\). The production of fast hydrogen atoms is through charge exchange with \(v_{\text{th}} = v_{\text{bulk}} \sim 100 \text{ km s}^{-1}\) stellar wind protons. There are variants of this model in which atoms are ballistically fired outward from the planet and interact with the stellar wind (Holmstrøm et al., 2008) and also models in which the mean free paths of the atoms are small, and the interaction occurs in a hydrodynamic mixing layer (Tremblin & Chiang, 2013).

These models are in a sense not independent but rather focus on two separate aspects of the same problem, since both thermal and nonthermal hydrogen may contribute to the absorption. In particular, the density of hydrogen atoms which may interact with the stellar wind (Model 2) is set by the outer limit of the upper atmosphere (Model 1). It has to be noted that, even if the models are complementary, the conclusion drawn from each other separately are not totally the same: The ENA model is consistent with a much smaller escape than the thermal escape model.

### 5.1.2. Rocky Planets

In the case of atmospheric escape from terrestrial/rocky planets, some modeling work has been done (Cohen et al., 2015; Dong et al., 2017; Gao et al., 2015a; Kislyakova et al., 2014; Tian, 2009; Wordsworth & Pierrehumbert, 2013b), but no reliable observations have been obtained so far, mainly due to the large size of the telescopes needed for the measurements (Gronoff et al., 2014), except in the case of extremely close-in rocky planets such as the disintegrating planet KIC 12557548b (Rappaport et al., 2012).
Major efforts went to model the planets in the HZ of Proxima Centauri b and Trappist 1. The work of Garcia-Sage et al. (2017) shows that a Earth-like planet at the location of these planets would suffer an enhanced ion escape, leading to the loss of the equivalent of the Earth's oceans over a billion years; the location of many of the planets inside the Alfvén surface (section 5.2) further prevents the existence of a sustainable atmosphere. It means that, to sustain habitability in the sense of liquid water existing at the surface, such planets would require a large amount of volatiles in their initial inventory and that they should not lose them in the active young years of their host star. To that extent, work has been done to look at the hydrodynamic escape of planets in the HZ of their active stars showing that even N₂ would be hydrodynamic (Johnstone et al., 2019). This theoretical work has been confirmed by the recent work of Kreidberg et al. (2019) that was able to show, using National Aeronautics and Space Administration (NASA)/Spitzer observations, that the exoplanet LHS 3844b has no thick atmosphere: Such an atmosphere would have been able to reduce the temperature difference between the nightside and the dayside of the planet compared to the observations. The conclusion of that problem is that, while there are the easiest targets for detecting habitable exoatmospheres with instruments such as the James Webb Space Telescope (JWST), planets in the HZ of red dwarfs may not be able to sustain them and therefore would be the worst target.

5.2. The Stellar Wind and the Alfvén Surface

The classical HZ of stars fainter than the Sun resides closer to the star. In particular, the HZ of M dwarf stars is located at planetary orbits of less than 0.1 AU. While the size of M dwarf stars is about 0.1–0.3 \( R_\odot \), their magnetic fields seemed to be overall stronger than the field of K, and G stars (Reiners & Basri, 2007). As a result, their Alfvén surface, at which the stellar wind exceeds the Alfvén speed and open the coronal field lines into the interplanetary space, is more extended than that of the Sun.

Since the Alfvén surface can be a measure to the boundary between the stellar corona and the interplanetary space (filled with the fully developed stellar wind), planets residing within the Alfvén surface could be considered to be inside the corona. In addition to the extreme temperatures that can exceed a million degrees, planets in this regime will orbit in densities and magnetic fields that can reach 2–4 orders of magnitude higher than those at Earth (Garraffo et al., 2016b) and, as a result, experience extreme dynamic and magnetic pressures at the planetary orbit. Such space environment conditions (Garraffo et al., 2017) may lead to an Alfvén wings (Io-like) topology of the planetary magnetosphere, at which a significant fraction of the planetary field is open to the stellar wind. As a result, the planetary atmosphere may be exposed to intense heating due to the incoming stellar wind energy in the form direct particle precipitation and Alfvén wave energy that is transmitted by the stellar wind. Additionally, since the planets reside in the sub-Alfvénic corona, some of the Alfvén wave heating that is deposited in the corona may be transferred to the planet. While essentially no work has been done on these processes in exoplanets, the scenario described above may suggest that it is unlikely that these planets are habitable. Therefore, the Alfvén surface might serve as an inner limit at which the HZ can be placed for a given stellar system. The result of the simulations of the distance of the Alfvén surface from its parent star as a function of the average magnetic field of that star is given in Figure 11. The spread in distance for a given magnetic field strength arises from possible differences in the geometric distribution of the magnetic field on the stellar surface. The Alfvén surface is smaller for complex field structures (i.e., higher order in the multipolar expansion) (Garraffo et al., 2016a).

6. Developments Needed in Measurements and Modeling of Atmospheric Escape

In the context of astrobiology, developments are needed in modeling, observations, and laboratory measurements for being able to observe and characterize rocky exoplanets’ atmospheres and for reconstructing the history of the atmospheres of the planets in the solar system. We review here some of the planned and suggested improvements. For a more comprehensive list of suggestion, white papers submitted to the National Academy of Science (2019) have a comprehensive list. The JWST, being the astrophysics flagship of the
2020s, is shaping the direction of astrobiology research, and therefore, many currently proposed studies and developments are linked to its targets of choice, the exoplanets around M dwarfs.

6.1. Measurements/Observations

The detection of a technosignature, currently searched in radio waves, would be the ultimate proof of a life that developed outside Earth (Wright, 2019). The development of observatories for such endeavor is outside the scope of this study since a mere detection would not provide data on how life developed there, under which conditions, and how it started. To answer these questions, better measurements are needed to understand the Sun and other Stars, laboratory measurements and sophisticated measurements are needed for better interpreting our planetary observations, and better instruments are needed for looking at exoplanets.

Space missions such as CHEOPS (Broeg et al., 2014) and Transiting Exoplanet Survey Satellit (TESS) (Ricker et al., 2015) are expected to find thousands of transiting planets with many terrestrial-like planets including hundreds of super-Earths over the next few years (Fridlund et al., 2016). With the ability to discover transiting exoplanets, efforts are being pursued for spectroscopic observations of exoplanetary atmospheres. The National Academy of Sciences (NAS) “Exoplanet Science Strategy” report mandated by the U.S. Congress recommends a direct-imaging telescope as a follow-on to the Wide Field InfraRed Survey Telescope (WFIRST) mission scheduled to fly in the mid-2020s after the launch of the JWST. Large amounts of time are being dedicated on the Hubble and Spitzer space telescopes, as well as major ground-based telescopes (e.g., Very Large Telescope [VLT], Keck, Gemini, Magellan, and Canada-France-Hawaii Telescope [CFHT]) (Madhusudhan et al., 2016). JWST and European-Extremely Large Telescope (E-ELT) will revolutionize exoplanetary spectroscopy. M dwarfs are prime targets for the detection and characterization of terrestrial exoplanets by the JWST, as they are abundant in the solar neighborhood and their small radii allow for greater transit signals from Earth-sized exoplanets (e.g., Quintana et al., 2014). Future spectra with JWST would be of unprecedented precision and resolution which will enable us to derive precise chemical abundances for transiting exoplanets. Our modeling can help constrain future observations in making precise determinations of detectable key species abundances, distributions, and understanding of their processes. The NAS report also recommends that the U.S. National Science Foundation (NSF) invest more in the future Giant Magellan Telescope (GMT) and proposed Thirty Meter Telescope (TMT) now being built in Hawaii and Chile, respectively. These telescopes would provide more focused study of exoplanets by using spectroscopy to seek out signs of free oxygen in their atmospheres. Such a study would be well suited for the goals of the astrobiology community.

For F and G stars, much of the observing is likely to initially be in the IR so the lower bound to some of the observations are likely to be near the tropopause (outside of H2O bands). For Venus-like exoplanets, this does not present a problem, but for Earth-like exoplanets it may be for F type stars. However, for all K- and M-type stars, processes of interest governing distribution of key species in atmospheres for these types of detections may be directly observed in the mesosphere and thermosphere. Additionally, there are spectral regions in the UV (and possibly in the visible to IR) where O2 and O3 in the mesosphere/upper stratosphere would dominate remote measurements. Those spectral regions would be an ideal target for future observations.

6.1.1. Solar and Stellar Measurements

The capabilities of the JWST mean that the search for biosignature with that observatory is mainly limited to M dwarfs. The increased risk of atmospheric escape due to Joule heating in the classical HZ around these stars (Garraffo et al., 2016b) means that it is extremely important to study the activity of these stars. The best way to perform this is with UV instruments such as Space Telescope Imaging Spectrograph (STIS) on the HST (France et al., 2013). The HST being on its end of life, with no repair mission planned, a mission dedicated to the EUV-XUV measurements on close stars (such measurements would be less affected by interstellar H absorption) would be critical to support the modeling and observations of planetary atmospheres that could harbor life.

The recent observation of a CME around another star (Argiroffi et al., 2019) shows the possibilities of such measurements. However, they should be improved to have a better idea of the fluxes of particles at other stars and to validate the semiempirical laws linking flares and CME (Moschou et al., 2019). See also section 3.4.1.
6.1.2. For Planets

As already emphasized, the knowledge of accurate cross sections is of critical importance for the precise evaluation of escape processes. For charge transfer, including double charge exchange, electron capture, and stripping, it has been customary to study these processes in the laboratory in two main directions: (i) under the wide umbrella of nuclear research, radiation dosimetry, and the effect of radiation on living tissues (Nikjoo et al., 2012), especially in water and carbon, and (ii) in astrophysics and heliophysics studies, especially rather recently with respect to X-ray production (Dennerl, 2010; Wargelin et al., 2008). In (i), one of the goals is to calculate the stopping power of particles in matter using so-called track-structure Monte Carlo models. In (ii), the applications are numerous, from ISM to cometary X-ray emissions. This has resulted in a rather well-understood behavior of charge-transfer cross sections at energies typically above 10-keV/amu impactor energy and peaking in the MeV range (Uehara & Nikjoo, 2002). At low energies, from a few tens of eV/amu to 10 keV/amu, which are the typical energies for solar wind charge exchange and in planetary ionospheres, the information is usually fragmented and one is often forced to extrapolate, more or less arbitrarily, the shape of the cross sections, leading to high uncertainties (Simon Wedlund, Behar, Kallio, et al. 2019).

Over the two last decades, experimental physicists have punctually studied aspects of solar wind charge exchange. Several international groups have specialized on different aspects (Dennerl, 2010), for example, the UV spectroscopy group at the University of Groningen (Netherlands) for cometary environments (Bodewits, 2007; Bodewits et al., 2004; Juhász, 2004), motivating studies of impacts of fast solar wind-like ions with several neutral species of planetary atmosphere relevance. The examples of H2O, CH4, CO, and CO2 are particularly relevant: Greenwood et al. (2000, 2004) and Bodewits et al. (2006) have recently measured with good accuracy charge-transfer cross sections of protons and helium ions on H2O, CH4, CO, and CO2 for astrophysics applications. However, these cross sections were not measured in the very low energy range (below 50 eV for helium ions, below 1.5 keV for protons). Moreover, certain electron capture and stripping reaction cross sections have yet to be altogether measured by any group. For example, the stripping reaction involving hydrogen or helium fast atoms and H2O ((H,He) + H2O → (H+,He+) + H2O + e−) has never been measured below 20-keV/amu energy; it may prove an important sink for the produced ENAs and hence may play a role in the escape of such particles into space.

A few online databases exist for several charge transfer cross sections. Despite ongoing work made to create online database and recommended sets of cross sections (e.g., Lindsay & Stebbings, 2005), it is left for a supplementary critical review of charge-transfer cross sections in planetary and cometary atmospheres to list all of the available cross sections, their energy range, evaluate their uncertainties and the gaps in our present knowledge, and provide a final recommendation that can be used in models and data analysis. Such a specific review is outside the scope of the present article, and we will here only point out one direction that experimental physics teams are encouraged to further study, that is, the solar wind charge transfer with a neutral atmosphere.

Further studies will have to choose colliding species, such as

- Neutrals of interest (nonexhaustive list by increasing atomic/molecular weight): H, H2, He, C, N, O, CH4, OH, H2O, Ne, N2, CO, O2, S, Ar, and CO2.
- Impactor of interest: H, H+, He, He+, He2+, and high ion states of O and Fe.

and to consider the two following points:

- Study more systematically all sources and sinks for the ion-ENA system: single and multiple electron capture, single and multiple electron stripping, and impact ionization by fast atoms and ions.
- Measure new energy-dependent cross sections and uniformly extend current cross section measurements to energies ranging from threshold to 20-keV/amu energy, most relevant for solar wind studies.

Other developments are needed for planets, such as Cassini-like missions to the Ice Giants, for a better understanding of the development of these atmosphere and their satellite. Such missions would give more insight into the evolution of our solar system. Improved instrumentation could be of use around the Earth to better discriminate the escaping species: It is currently extremely difficult to know if it is an O or a N that is leaving the atmosphere.
From that point of view, it is also important to investigate the X-Ray emission of planets, exoplanets, and comets: The X-ray halo, created by charge exchange between the exosphere and the incoming solar wind, gives some insight to the composition of escaping species (Dennerl, 2007, 2010; Krasnopolsky et al., 1997). This is of interest both for planets and exoplanets since detecting such a halo in another stellar system would give some direct insight in the composition of the exoplanetary exosphere.

6.1.3. For Exoplanets

Atmospheric escape from exoplanets can be constrained by observations of components that affect the escape. This includes compositional observations (i.e., transmission spectra) and direct observations of escaping material. These observations will pose challenges in the upcoming decades.

It is also necessary to observe magnetic fields of exoplanets as they may play a key role in the atmospheric escape. One promising option is to obtain information about exoplanets magnetic fields via observed signature star-planet interaction. These signatures include induced chromospheric activity (e.g., Cauley et al., 2019; Shkolnik et al., 2008) or modulation of coronal radio emissions (Cohen et al., 2018). The direct detection of exoplanets magnetic fields (via radio observations of auroral emissions Zarka, 2007) has recently been reported (Vedantham et al., 2020). This has only confirmed the existence of the magnetosphere: More work is needed to be able to estimate, for example, the magnetic moment from these observations. The modeling of the interaction of the stellar wind with the planetary magnetosphere of HD 209458b led to the estimation of its magnetic moment (Kislyakova et al., 2014) from the observation of Ly-α. Giant space UV telescopes would be choice instruments to study the upper atmosphere of rocky exoplanets. Gronoff et al. (2014) proposed a technique to detect hydrodynamic escape of CO$_2$ or O$_2$-rich planets using such laboratories. For the detection of biosignatures, a review of techniques and developments needed can be found in Fujii et al. (2018). The generalization of the detection technique used by Kreidberg et al. (2019) to detect the absence of an atmosphere around LHS 3844b is also needed to look at the best target for future telescopes.

6.2. Modeling

To get a comprehensive view of the escape of planetary atmospheres, models have to be developed to take into account all the energetic inputs and all the processes leading to the escape. The outputs of such models have to be compared with observations. Problems lies with inputs parameters for the model (cross sections and observation of, e.g., solar flux), the estimation of the uncertainties, and also with the neglected parameters. It is often the case that our instrumentation gives a very detailed view of the conditions on a planet; however, the uncertainties in the input parameters of the models make it challenging to interpret (Sánchez-Cano et al., 2018). The estimation of model uncertainties from the different input parameters can be an arduous task (Gronoff, Simon Wedlund, Mertens & Lillis 2012; Gronoff, Simon Wedlund, Mertens, Barthélemy, et al. 2012) and becomes problematic once free parameters are needed, which is often the case in our models of atmospheres, stellar wind, stellar wind interactions, etc. It is possible to begin solving the problem by careful comparison with solar system observations, then with extrapolation. On the other hand, the instrumentation may not be sensitive enough to observe interesting phenomenon in exoplanetary atmospheres or to provide significant model constrains.

Future modeling should also include the dynamical response of the planet’s atmosphere to dynamic drivers, stellar evolution scale changes of atmospheric escape, and self-consistent coupling between the external drivers and the different regions of the atmosphere.

6.2.1. Modeling of Solar and Stellar Environments

Global models for the solar corona have been developed since the late 1960s by solving the MHD equations. The models are driven by data of the photospheric radial magnetic field in combination with the potential field method (Altschuler & Newkirk, 1969). In recent years, more self-consistent models have been developed for the solar corona and solar wind (e.g., Downs et al., 2016; Lionello et al., 2014; van der Holst et al., 2014). These models incorporate coronal heating and wind acceleration in the form of large-scale heating and momentum terms. These large-scale terms are parameterized and tuned to match solar observations, and the models have been successful in reproducing the observed density and temperature structure of the solar corona, and the observed structure of the solar wind.

The limited availability of observations of photospheric magnetic field of selected stars using the Zeeman-Doppler Imaging technique (Semel, 1980) has led to a growing global modeling in stellar coronae and stellar winds of Sun-like stars (e.g., Cohen et al., 2010; Garraffo et al., 2016b; Vidotto et al., 2011).
However, since the stellar winds of solar analogs cannot be directly measured, the results of these studies are poorly constrained. Therefore, a better modeling work is needed to constrain the magnitude of the stellar wind, and the coronal structure and temperature for different stars as these parameters define the stellar environments at which exoplanets reside in. In particular, the scaling of the global heating and acceleration parameters needs further investigation and quantification to better understand how these processes scaled with stellar type.

6.2.2. Modeling Atmospheric Escape From Exoplanets

The current modeling tools for planetary atmospheric escape are built on and tune to known, measurable atmospheres within the solar system. These tools have already been used to study escape from exoplanetary atmospheres with no significant constraints of the results. A number of features, which are different from solar system bodies, have already been identified to be crucial for exoplanetary atmospheric escape, especially in the case of close-orbit planets. However, these features need more self-consistent modeling in order to be better defined and quantified.

The first notable feature is that atmospheric escape from close-orbit planets may be extremely high, to the point that atmosphere could be completely lost. This is due to extremely high dynamic pressure of the stellar wind near these planets (e.g., Dong et al., 2017; García-Sage et al., 2017; Garraffo et al., 2016b), the strong orbital variations of the stellar wind conditions, and potential strong heating of the upper atmosphere (e.g., Cohen et al., 2014, 2018). A more detailed model is required to quantify the exact energy deposition between the wind and the planetary atmosphere, as current models focus on the stellar wind-magnetosphere interaction, without detailed modeling of the energy and mass transfer to and from the atmosphere itself.

The second notable feature is the impact on the planetary upper atmosphere and ionosphere. Current models provided estimation about the Joule heating assuming specific, constant atmospheric conductance. Since the conductance is the key to determine the heating, further self-consistent modeling is needed to estimate the ionospheric conductance. In particular, these calculations are needed for the case where the EUV and X-ray stellar radiation are much higher than the Earth case and for different atmospheric composition.

Finally, close-orbit exoplanets may reside within the Alfvénic point inside the stellar corona. Therefore, a direct star-planet interaction is expected to occur. In order to investigate the impact of such a direct interaction between the stellar corona and the planet, a self-consistent modeling that couples the corona and the planetary atmosphere domains is needed.

An example of a code in development to address some of these problems could be Institut d’Astrophysique de Paris Particle-In-Cell- Electromagnetic Global 3D Code (IAPIC), a particle-in-cell electromagnetic 3-D global code, used (Baraka, 2016; Baraka & Ben-Jaffel, 2010; Ben-Jaffel & Ballester, 2013, 2014) to produce the magnetosphere (XZ plane) of an Earth-like planet. Both plasma density and field lines are shown in Figure 12. It is interesting to see that the PIC simulations naturally recover the FAC (streams of particles appearing between cusps and current sheet in the figure) that drive particles precipitation from the magnetosphere into the polar regions, producing auroral emissions. IAPIC can provide both the angular and energy distributions of the impinging magnetospheric particles into the ionosphere. Charge separation is obtained in the code so that kinetic effects could be obtained while conserving charge (Villasenor & Buneman, 1992). These electrons and ions enter the upper atmosphere to trigger ion-chemistry, heating, and winds. Their fluxes should be used as input in existing ionospheric models to evaluate new species produced and atmospheric inflation due to the extra heating deposit in the auroral region of any exoplanet. The simulation, shown in Figure 12, was carried out with these code parameters for a grid size of 0.1κRₚ and an ion-electron mass ratio of \( \frac{m_i}{m_e} = 100 \).

6.2.3. Modeling Exoplanetary Magnetic Field Observations

Following the previous sections, it seems like exoplanetary magnetic fields may play a crucial role in the evolution and sustainability of exoplanets atmospheres. However, these planetary field currently cannot be detected and observed.

Modeling of star-planet interaction suggest that this interaction can potentially generate observable signatures that can help to quantify the planetary magnetic field (e.g., the broadening of Ly-α in Kislyakova et al., 2014, and the soft X-ray emission in Kislyakova et al., 2015). However, it is clear that a deep understanding of...
the stellar background field is needed for this purpose (Cohen et al., 2011; Matsakos et al., 2015; Shkolnik et al., 2008; Strugarek et al., 2014). In that context, more detailed modeling work on the background stellar environment can support the interpretation of star-planet interaction observations.

A number of attempts were made to estimate auroral radio emissions from exoplanets (Burkhart & Loeb, 2017; Grießmeier et al., 2007; Lazio & Farrell, 2007; Lynch et al., 2018; Nichols & Milan, 2016; See et al., 2015; Turnpenney et al., 2018; Vidotto et al., 2015; Zarka, 2007). Most of these studies have concluded that such auroral emissions are not detectable with the current radio telescopes. An alternative approach has been recently proposed by Cohen et al. (2018), who proposed to look for planetary modulations of the ambient coronal radio emission (exoplanet radio transit), instead of looking for the planet as a radio source. Using idealized models for the planetary and stellar fields, Cohen et al. (2018) have shown that observing the ambient coronal radio intensity, as well as the planetary modulation are more feasible. To better quantify and estimate the latter method, further, more detailed modeling work is needed, targeting specific planetary systems.

7. Discussion and Conclusion

The problem of atmosphere escape is more complex than just the estimation of Jeans’ escape or energy-limited escape. Several outstanding questions in planetary science can be linked to atmosphere escape, from the problem of the FYS to the question of detection of astrobiological signatures. These problems are leading to a roadmap of future investigations.

7.1. Problems Still to Be Resolved

7.1.1. The FYS Paradox

The FYS paradox was introduced by Sagan and Mullen (1972). The paradox states that based on stellar evolution models, during its earlier stages, the Sun’s luminosity was about 30% lower than its current luminosity. As a result, the surface equilibrium temperature of the Earth would be below the freezing point of water. However, many types of geological evidence for the existence of liquid water were found both at Earth and Mars. Therefore, we need to introduce some heating process, which increased the average surface temperature of the Earth above 0°C.

The most prominent solution to the paradox is the existence of greenhouse gases in the atmosphere, which lock the IR radiation and lead to a global warming of the Earth’s surface (see, e.g., Kasting, 1993). An enormous amount of work has been done on this topic in what came to be the science field of “Global Climate Change” (Feulner, 2012). It has been suggested tropical cirrus clouds could also enhance the greenhouse effect, being either the main explanation to the FYS paradox or a complementary source to a greenhouse gas (Rondanelli & Lindzen, 2010). Goldblatt and Zahnle (2011) further develop the more general discussion on the effect of clouds on the climate, noting that Rondanelli and Lindzen (2010) base the cirrus solution to the FYS on the “iris theory,” which stipulates that the cirrus coverage should increase if the surface temperatures decrease, which is quite controversial for the current Earth (see, e.g., the comparison of the theory with observations by Chambers et al., 2002) but may be applicable on other atmospheres. Urata and Toon (2013) applied the theory to the Early Mars, but the work of Ramirez and Kasting (2017) shows that there is no room for error when considering cirrus clouds for warming: A large cirrus cloud coverage, greater than 70%, should be present. This explanation is therefore unlikely to be applicable to Mars since such a cloud coverage is not realistic: Cirrus cloud formation is limited by the parts of the atmosphere that are undersaturated in water. Overall, cloud warming has not been proven to be the solution to the FYS paradox, but it shows the importance of addressing the problem of cloud formation and humidity transport and therefore shows the importance of looking at the climate using 3-D GCMs. On the other hand, solutions to the FYS paradox that involve external factors, that is, particle precipitations or a heavier Sun, are attractive since they could solve the problem both at Mars and the Earth. Some theories suggested that cosmic rays may affect the

Figure 12. Field-aligned current and prototype kinetic simulation of exoplanet. These simulations are leading to a better understanding of the dynamic within an exoplanetary magnetosphere, which is needed for computing both energetic inputs in the atmosphere and loss of plasma from the magnetosphere.
cloud condensation in the Earth’s atmosphere, with an overall cooling effect when there is fewer galactic cosmic ray (GCR) (i.e., when the solar activity is higher) (Shaviv, 2005; Svensmark & Friis-Christensen, 1997). Thus, a significant reduction in the cosmic ray flux may increase the surface temperature of the Earth. These models are controversial both in the cosmic rays ability to affect the cloud condensation (Kirkby et al., 2011) and the heating efficiency of the process. The results of the CLOUD experiment at CERN tend to indicate that the present-day cloud formation cannot be effectively affected by GCR flux (Dunne et al., 2016; Pierce, 2017). The aerosols responsible for cloud condensation come mainly from the ground, and setting the GCR flux at 0 would only reduce the cloud coverage by ≈10%. It is to be noted that this experiment does indicate that GCR can affect aerosol/particle formation (Kirkby et al., 2016; Tröstl et al., 2016), which is something that is also observed on other atmospheres such as Titan (Dobrijevic et al., 2016; Loison et al., 2015). The contentious point is whether or not GCR variation has an influence on the current climate or not. Studies such as Lanci et al. (2020) do not find historical evidence of GCR control of climate, while Svensmark et al. (2017) advocate for a strong control of the climate by GCR in the past, with increased GCR leading to aerosols that act as cloud condensation nuclei (CCN), that lead to the formation of clouds, and ultimately heating. It has to be noted that aerosols do not always act as CCN but can act as coolant (e.g., Trainer et al., 2006b). The work of Airapetian et al. (2016) suggests, on the contrary, that the increased particle flux from the SEP events, more frequent for the Young Sun, led to the creation of greenhouse gases, that could help resolve the paradox. This interaction between ionizing radiation and climate is fundamentally different since it requires extreme radiation rates (compared to present values) to get a significant increase in greenhouse gases. In addition, it does not suppose that CCN is the limiting factor in the creation of clouds: The atmosphere can be undersaturated (Ramirez & Kasting, 2017).

In the context of astrophysics, a solution for the paradox can be found if one can show that the solar mass was about 10% larger than its current mass. This requires the young Sun’s mass loss rate to be very high, with the ability to keep this high mass loss rate for rather long time. As mentioned in section 4.1, it is unlikely that the mass loss to the ambient solar wind can be sufficient. However, it is possible that due to high activity levels at early stellar stages, the Sun lost large amount of its mass via CMEs, although present estimates also indicate this mechanism is insufficient (Drake et al., 2013).

### 7.1.2. Impact of Planetary Magnetic Field

Whether a planet is magnetized or not impacts ionospheric ion outflow processes and the return rate of ions outflowing from the ionosphere. It was believed that a planetary magnetic field was shielding its atmosphere from escape, until observations, reported by Barabash (2010), showed that the escape rate at Earth is higher than at Venus and Mars. This was further discussed by Strangeway et al. (2010) and explored more in details in Brain et al. (2013). This latter work also discussed the influence of magnetic field on incoming gases that could also have an effect on climate. Tarduno et al. (2014) and Ehlmann et al. (2016) looked at the effect of magnetospheric escape respectively for the early Earth and exoplanets.

Gunell et al. (2018) compiled the effect of the planetary magnetic moment on ionospheric ions outflow rate in the current solar system and compared it to other sources of escape. Considering that the observed ionospheric ion outflow rate on Earth, Mars, and Venus is of the same order of magnitude (10^{25} s^{-1}) while only Earth has a strong magnetic field, the authors made empirical models of ion outflow for three hypothetical planets with atmospheric conditions similar to the Earth, Mars, and Venus but with a variable magnetic moment. They show that for each of those planets the mass escape rate, including both oxygen and hydrogen is similar in the unmagnetized range and for high magnetizations. In-between, they identify two maxima where outflow is enhanced by a factor of 2–5, one corresponding to polar cap escape and dominant for hydrogen and another corresponding to cusp escape. The presence of a large magnetosphere around a planet actually diverts part of the stellar wind energy and protects the atmosphere from sputtering and ion pickup. The induced magnetospheres of the unmagnetized planets also provide protection from sputtering and ion pickup but to a lesser extent. However, magnetospheres are much bigger objects than the planets themselves. The presence of a magnetosphere increases the size of the interaction region between the stellar wind and the planet and thus the amount of energy, which can potentially be transferred into the ionosphere. For instance, the cross section of the Earth magnetosphere with the solar wind is about 100 times higher than the cross section of the Earth itself with the solar wind. Consequently, the amount of energy transferred from the stellar wind to the ionosphere of magnetized planet is not necessarily lower than for unmagnetized planets (Brain et al., 2013).
Large-scale magnetospheres enable polar cap and cusp escape, which increases the escape rate. Two outflow processes are enhanced by the presence of a magnetosphere. The first is the polar wind, which corresponds to a thermal ion outflow on the open magnetic field lines at high latitudes, above the polar caps. It maximizes for moderate magnetic moments when the size of the polar cap is maximum. The second corresponds to outflow from the cusp region where a significant amount of the stellar wind energy is deposited. This energy deposition increases with the size of the magnetosphere (i.e., with its cross section with the stellar wind) but is limited by the amount of ions available at the ionospheric level.

The effect of magnetospheres on the ion return rate is discussed in section 2.5. In that case the protective effect of the magnetosphere is not related to the outflow itself but to the trapping of outflowing ionospheric ions, even those with high energies well above the gravitational binding energy, which was thought to result in a significant return rate into the atmosphere. However, recent observations in the Earth magnetosphere question the validity of this protective effect. Indeed, that the measured flux of precipitating ions in the ionosphere is well below the measured flux of outflowing ionospheric ions and the escape route above the polar ionosphere, where polar cap and cusp outflows occur, seems to preferentially lead to a direct ion loss to interplanetary space rather than to a return in the atmosphere.

On the other hand, the thermospheres of Mars and Venus are called cryospheres because of the cooling effect of CO₂. Their thermospheric temperature is extremely low, effectively shielding the atmospheres from several escape processes: The Earth’s atmosphere without a magnetic field is believed to be escaping more efficiently. In addition, the main species in the ionosphere of Venus is O₂⁺, while it is O⁺ at Earth (Mendillo, 2019), and this may affect the amount of ions able to escape via nonthermal processes (Mars is a special case since above ≈200 km O⁺ is in majority while it is O₂⁺ below). A recent study by Wei et al. (2014) shows that using an escape model developed for Mars, but with Earth's upper atmosphere, increases greatly the O escape. This study also suggests that there are some correlations between lower content of O₂ at Earth and magnetic field inversions. Such a study could be criticized on the basis that the magnetic field does not seem to disappear during inversions (e.g., Nowaczyk et al., 2012) or on the basis that the fluxes of O₂ at Earth are dominated by the biosphere and the oxidation of the crust. The fluxes of oxygen in Wei et al. (2014) are indeed of the same order of magnitude as the current oxygen losses in the crust (Catling, 2014). A reduction of carbon burial—which is a life-controlled process leading to net O₂ fluxes—could explain the loss of oxygen without the need for a fast process. It is also in disagreement with the observation of higher ion escape near magnetic anomalies at Mars (Inui et al., 2019; Sakai et al., 2018).

To summarize, while the presence of a magnetosphere has a clear impact on ionospheric outflow, recent developments in the study of the coupling between stellar wind, magnetospheres, and ionospheres challenge the idea of a protective effect of magnetospheres on atmospheric erosion. It could simply be that the question is poorly asked and that a better question is “what kind of atmospheres require a magnetic field to be sustainable in a given set of solar/stellar activity conditions.” In any case, recent studies such as Brain et al. (2013), Gunell et al. (2018), Airapetian et al. (2017), and Garcia-Sage et al. (2017), as well as the case of Mercury, show that an intrinsic magnetic field does not totally protect an atmosphere. A contrario, the case of Venus shows that a magnetic field absence does not prevent sustaining a dense atmosphere.

### 7.1.3. Impacts of Stellar Dynamics

In a large fraction of the studies of escape through time, the stellar parameters, that is, the EUV-XUV fluxes and the stellar wind, are considered as slowly varying with the epoch. The impact of the frequency of intermittent stellar events such as flares, CMEs, and SEPs on the escape is seldom taken into account. This is a major problem for studies of close-in exoplanets since these events can extremely affect the atmosphere as the observations of the variations of escape rate at Mars due to a CME has shown (Jakosky et al., 2015). MAVEN is currently showing that extreme solar events have a very important role in the loss of atmosphere at Mars (Jakosky et al., 2018; Mayyasi et al., 2018). As an example, the increase in the exospheric temperature due to a flare has been observed (Elrod et al., 2018), along with change in the upper atmospheric ion and neutral composition (Thiemann et al., 2018) and an increase of ≈20% of the escape (Lee et al., 2018).

The work of Garcia-Sage et al. (2017) has shown that the EUX-XUV flux can lead to extreme absorption at rocky exoplanets around M dwarfs; however, it does not answer the question of how much active a G-star an Earth-like planet could survive.
7.2. The Role of Nonatmospheric/Stellar Processes

While it is not generally explicit in the discussion above, the mass of the planet that is experiencing escape is a critical factor. Closely related is its radius and therefore its density. The planets of the solar system are there to show that the mass is the first factor to consider when estimating if a body will have an atmosphere or not; the energy received/distance to the Sun being the second factor. Mass is still challenging to retrieve, especially the mass of small planets, whose signal in radial velocities can be hidden by the natural variations of the star. Once mass and distances are considered, it may be possible that interesting effects come from close-in exoplanets, such as the Roche limit of the star reaching for the planet’s atmosphere. Overall, it should not be forgotten that the inventory of volatiles, which has be estimated from the density of the exoplanet, will define the lifetime of an habitable world with large escape rates.

7.3. The Future of Research on Escape Processes

The study of planetary atmosphere habitability and evolution has, as shown here, three main directions.

1. Escape modeling efforts. The approach of this review has been reductionist; we have sought to isolate the individual escape processes and identify simple ad hoc models that can determine whether or not a specific escape process is important. Yet a better approach to escape would be to create so-called “grand-ensemble models” that are able to examine the interactions between the different processes without a priori exclusion of processes. An already invoked example comes from Chassefière (1997), in which comprehensive treatment of multiple types of charge exchange predicted an increased exospheric temperature and therefore, indirectly, higher thermal escape. Improved models will allow the evaluation of critical parameters to help people work on the deeper parts of the atmosphere to estimate which species are escaping and at which rate, in order to prevent poor estimates based on energy-limited escape (that do not take diffusion limitation into account). An additional consideration in modeling is to devise a standard procedure for asynchronously coupling climate, chemistry, and escape models at exoplanets to study the evolution of climate and composition in tandem with stellar evolution.

2. Laboratory work. A major limiting factor of escape models is the quality of the input parameters, such as chemical reaction rates and cross sections. Laboratory experiments and model-laboratory data comparisons such as that of Simon Wedlund et al. (2011) are needed to refine the accuracy of the physicochemistry parameters and, in turn, may help identify the needs of the community for the study of specific processes. Laboratory data are also crucial to retrieve parameters from observational data.

3. Observation work. Observations efforts are limited and currently concentrated to what is believed to be the “best known targets” for habitability. Unfortunately, our instrumentation is not optimized for detecting habitability signatures on these targets. Future observations should be designed not just to characterize the bulk properties of the atmosphere but also to consider known or potential observables affecting atmospheric stability. One advantage of doing so is that processes like thermal escape mostly take place above cloud and haze layers and so may not be as challenging as observing lower in the atmosphere. These targets have broader characteristics than current one and could benefit from the whole range of existing instrumentation to answer questions leading to constrain the conditions required for habitability. Techniques should be improved to better understand the stellar environment of exoplanets, such as the observation of the stellar winds potentially impacting exoplanets as well as stellar variability in general, which has a strong potential impact on transit observations of planetary atmospheres (Wakeford et al., 2018). As characterizing the variability of a host star typically requires less sensitive instrumentation than detecting a telluric planet orbiting it (particularly for warmer stars), it may be worth prioritizing observations of variability of types other than those around which telluric planets currently can be observed.

Overall, the challenge is thus to couple a grand-ensemble escape model with a complex planetary atmosphere model, itself coupled with a planetary interior model. From there it would be possible to obtain the whole story of an empirical planetary atmosphere in time. The uncertainties in each of these submodels have to be correctly evaluated, so that it is possible to address the overall model validity. This is why a validation strategy is also of utmost importance for this kind of work.

On the stellar part, the challenge will be to determine the activity history of a star, both from the slowly evolving parameters, such as luminosity, and the discrete events such as flares. From there, it would be possible
Figure 13. Overview of the escape processes, along with an example of a planet where they are major/important. The Earth’s main escape processes are Jeans’, charge-exchange, and polar wind. At Venus, interaction with the solar wind; at Mars, photochemical escape and ion escape; at some exoplanets/comets hydrodynamic escape.

to evaluate how a given star stresses an atmosphere over time. Finally, it will be necessary to develop observation missions dedicated to study the UV flux of stars to validate the model of activity in time.

7.4. Effects of Escape on Biosignatures

The escape processes reviewed above have significant influence on the composition of the upper atmosphere and acting over geologic time can affect the bulk composition of the atmosphere, surface, and interior. The consequences of atmospheric escape for our search for life via chemical biosignatures in the atmosphere and on the surface must therefore be considered (Airapetian et al., 2016; Des Marais et al., 2002; Domagal-Goldman et al., 2014). The alteration of planetary chemistry by escape can result in both false positive and false negative biosignature errors if it is not accounted for (Gao et al., 2015b).

False positive biosignatures that can be produced by interactions at the top of the atmosphere include oxygen and oxidized species such as NO as well as organics such as the haze materials produced through UV photolysis at Titan and elsewhere. The preferential loss of hydrogen from water is one way for oxygen to arise from escape-related abiogenic processes. The processing of sufficient water to influence the bulk oxidation state of the surface materials is likely the cause of high oxidation in the Martian surface (Lammer et al., 2003). False negative biosignatures would result from the masking of true biogenic molecules by escape processes, either through rapid modification by particle or photon radiation or through chemical interaction with, for example, photolysis-produced oxygen.

Biosignatures related to disequilibrium chemistry (Krissansen-Totton et al., 2018) must contend with non-LTE behavior in the upper atmosphere and the potential for disequilibrium signatures to be
transferred from the upper to the lower atmosphere. Even biosignatures that are based on time variations (Olson et al., 2018) need to contend with seasonal variations in star-planet interaction caused by a tilted magnetic field axis which could produce either false positive or false negative results.

This discussion is not intended to be exhaustive or definitive, but instead, we seek to highlight the importance of understanding the impact of the stellar environment on the production, destruction, or masking of putative chemical biosignatures. In general, although space weather processes involve small fractions of the planetary mass, they can, like biological reactions, be quite selective in their reactants and products and over time can yield significant signals that must be differentiated from biological ones.

7.5. Final Thoughts

We have reviewed the different escape processes considered so far in the literature and summarized in Figure 13. Understanding these processes, and also ones that are still to be discovered, makes it possible to understand how a planetary atmosphere evolves. This is, however, not enough to understand the whole history of an atmosphere: Change in the atmosphere composition, change in the stellar activity, and change of the outgassing from the planetary interior are examples of processes that affect the development of an atmosphere and can lead to very different pathways. To that extent, life is one of the major modifiers of Earth’s atmosphere. It would be easy to consider an atmosphere that is out of equilibrium or that is very difficult to model/understand by our current means, as harboring life; this is the idea behind the concept of biosignatures. However, the detection of a Titan-like atmosphere outside our solar system may lead to life detection claims that are not (at least to date) consistent with Titan’s observations. On the other hand, the atmosphere of an early Earth may be detected but considered as sterile.

Since this paper has been showing that habitability is a dynamic process and that the habitability of a planet is the result of its history, and not just of its location with respect to its star, it is important to extend that notion to biosignatures. In this respect, it would be preferable to announce the detection of molecules relevant to prebiotic chemistry instead of directly announce biosignatures, so that no extraordinary claim is made without extraordinary evidence.

**Acronyms**

CME Coronal mass ejection  
DSMC Direct simulation Monte Carlo  
ENA Energetic neutral atom  
ESA European Space Agency  
EUV Extreme ultraviolet  
FAC Field-aligned currents  
FYS Faint Young Sun (typically for speaking about the Faint Young Sun paradox)  
GCM Global Circulation Model  
HZ Habitable zone  
KHI Kelvin-Helmholtz instability  
MAVEN NASA/Mars Atmosphere and Volatile and EvolutioN mission  
MEX ESA/Mars Express  
MHD Magnetohydrodynamic  
NASA National Aeronautics and Space Administration  
SEP Solar Energetic Particle  
UV Ultraviolet  
VEX ESA/Venus Express  
XUV X-ultraviolet

**Notation**

\( k \) the Boltzman constant  
\( m \) the average molecular mass  
\( g \) the gravitational acceleration (typically dependent upon the altitude)
Data Availability Statement

Data were not used nor created for this research.

References

Adams, F. C. (2011). Magnetically controlled outflows from hot Jupiters. The Astrophysical Journal, 730, 27. https://doi.org/10.1088/0004-637X/730/1/27

Airapetian, V. S., Barnes, R., Cohen, O., Collins, G. A., Danchi, W. C., Dong, C. F., et al. (2019). Impact of space weather on climate and habitability of terrestrial-type exoplanets. International Journal of Astrobiology, 1–59. https://doi.org/10.1017/S1473550419000132

Airapetian, V. S., Glocer, A., Gronoff, G., Hébrard, E., & Danchi, W. (2016). Prebiotic chemistry and atmospheric warming of early Earth by an active young Sun. Nature Geoscience, 9, 452–455. https://doi.org/10.1038/ngeo2719

Akbari, H., Andersson, L., Peterson, W. K., Espley, J., Benna, M., & Ergun, R. (2019). Ambipolar electric fields. The Astrophysical Journal Letters, 886(1), L3.

Akbari, H., Andersson, L., Peterson, W. K., Espley, J., Benna, M., & Ergun, R. (2019). Ambipolar electric field in the Martian ionosphere: MAVEN measurements. Journal of Geophysical Research: Space Physics, 124, 4518–4524. https://doi.org/10.1029/2018JA026325

Alesini, I., Way, M. J., Harman, C., Tasgarian, K., Wolf, E. T., & Gronoff, G. (2019). Modeling a transient secondary paleolunar atmosphere: 3-D simulations and analysis. Geophysical Research Letters, 46, 5107–5116. https://doi.org/10.1029/2019GL082494

Alho, M., McKenna-Lawlor, S., & Kallio, E. (2015). Paleo Mars energetic particle precipitation. Planetary and Space Science, 119, 103–110.

Altschuler, M. D., & Newkirk, G. (1969). Magnetic fields and the structure of the solar corona. I. Methods of calculating coronal fields. Solar Physics, 9, 131–149. https://doi.org/10.1007/BF00145734

Ambillard, P.-O., Moussaoui, S., Dudok de Wit, T., Aboudarham, J., Kretzschmar, M., Liliensten, J., & Auchère, F. (2008). The EUV Sun as the superposition of elementary Suns. Astronomy and Astrophysics, 487, L13–L16. https://doi.org/10.1051/0004-6361:200805958

Anglada-Escude, G., Amado, P. J., Barnes, J., Berdtkas, Z. M., Butler, R. P., Coleman, G. A. L., et al. (2016). A terrestrial planet candidate in a temperate orbit around Proxima Centauri. Nature, 536, 437–440. https://doi.org/10.1038/nature19106

Argiroff, C., Reale, F., Drake, J. J., Ciavarella, A., Testa, P., Bonitto, R., et al. (2019). A stellar flare- coronal mass ejection event revealed by X-ray plasma motions. Nature Astronomy, 3, 742–748.

Arney, G., Domagal-Goldman, S. D., Meadows, V. S., Wolf, E. T., Schwieterman, E., Charnay, B., et al. (2016). The pale orange dot: The spectrum and habitability of hazy Archean Earth. Astrobiology, 16(11), 873–899. https://doi.org/10.1089/ast.2015.1422

Arney, G., & Kane, S. (2016). Venus as an analog for hot Earths. ArXiv e-prints.

Atreya, S. K., Adams, E. Y., Niemann, H. B., Demick-Montelara, J. E., Owen, T. C., Fulchignoni, M., et al. (2006). Titan’s methane cycle. Planetary and Space Science, 54, 1177–1187. https://doi.org/10.1016/j.pss.2006.05.028

Atreya, S. K., Donahue, T. M., & Festou, M. (1981). Jupiter—Structure and composition of the upper atmosphere. The Astrophysical Journal Letters, 247, L43–L47. https://doi.org/10.1086/183586

Avie, G., Marty, B., Burgess, R., Hofmann, A., Philippot, P., Zahnle, K., & Zakharov, D. (2018). Evolution of atmospheric xenon and other noble gases inferred from Archean to Paleoproterozoic rocks. Geochimica et Cosmochimica Acta, 232, 82–100. https://doi.org/10.1016/j.gca.2018.04.018
Gronoff et al. 69 of 77

Journal of Geophysical Research: Space Physics

10.1029/2019JA027639

Lanci, L., Galeotti, S., Grimani, C., & Huher, M. (2020). Evidence against a long-term control on Earth climate by galactic cosmic ray flux. *Global and Planetary Change*, 185, 103.095. https://doi.org/10.1016/j.gloplacha.2019.103.095

Laneville, M., Kameya, M., & Cleaves, H. J. (2018). Earth without life: A systems model of a global abiotic nitrogen cycle. *Astrobiology*, 18(7), 897–914. https://doi.org/10.1089/ast.2017.1700

Lara, L. M., Lellouch, E., López-Moreno, J. J., & Rodríguez, R. (1996). Vertical distribution of Titan's atmospheric neutral constituents. *Journal of Geophysical Research*, 101, 23,261–23,263. https://doi.org/10.1029/96JE02036

Lazio, T. J. W., & Farrell, W. M. (2007). Magnetospheric emissions from the planet orbiting γ Bootis: A multipeoch search. *The Astrophysical Journal*, 668, 1182–1188. https://doi.org/10.1086/519730

Lean, J. L., Warren, H. P., Mariska, J. T., & Bishop, J. (2003). A new model of solar EUV irradiance variability 2. Comparisons with empirical models and observations and implications for space weather. *Journal of Geophysical Research*, 108, 1059. https://doi.org/10.1029/2001JA009238

Leblanc, F., & Johnson, R. E. (2001). Sputtering of the Martian atmosphere by solar wind pick-up ions. *Planetary and Space Science*, 49(6), 645–656. https://doi.org/10.1016/S0032-0633(01)0003-4

Leblanc, F., & Johnson, R. E. (2002). Role of the molecular species in pickup ion sputtering of the Martian atmosphere. *Journal of Geophysical Research*, 107(E2), 5–1. https://doi.org/10.1029/2000JE001473

Leblanc, F., Martinez, A., Chaufray, J. Y., Modolo, R., Hara, T., Luhmann, J., et al. (2018). On Mars' atmospheric sputtering after MAVEN's first Martian year of measurements. *Geophysical Research Letters*, 45, 4685–4691. https://doi.org/10.1002/2018GL077199

Leblanc, F., Modolo, R., Curry, S., Luhmann, J., Lillis, R., Chaufray, J. Y., et al. (2015). Mars heavy ion precipitating flux as measured by Mars Atmosphere and Volatile EvolutioN. *Geophysical Research Letters*, 42, 9135–9141. https://doi.org/10.1002/2015GL066170

Lecavelier Des Etangs, A. (2007). A diagram to determine the evaporation status of extrasolar planets. *Journal of Geophysical Research*, 112, 1–14. https://doi.org/10.1029/2007JA012144

Lee, C. O., Jakosky, B. M., Luhmann, J. G., Brain, D. A., Mays, M. L., Hassler, D. M., et al. (2018). Observations and impacts of the 10 September 2017 solar events at Mars: An overview and synthesis of the initial results. *Journal of Geophysical Research: Space Physics*, 123, 8871–8885. https://doi.org/10.1002/2018GL079162

Lionello, R., Velli, M., Downs, C., Linker, J. A., & Mikić, Z. (2014). Application of a solar wind model driven by turbulence dissipation to a 2D magnetic field configuration. *The Astrophysical Journal*, 796, 111. https://doi.org/10.1088/0004-637X/796/2/111
