From star-formation to recombination: expanding our view of the radio recombination line universe

Emig, K.L.

1 | Introduction

Galaxies are gravitationally bound systems of stars, gas, dust, and a central supermassive black hole, all of which reside in a dark matter potential. A major question guides current astrophysical research: what regulates the evolution of galaxies? The characteristics of the interstellar medium (ISM), the formation of stars and their feedback on the ISM, events driven by the central black hole, accretion of intergalactic gas, galaxy mergers and their environment, and the dark matter potential in which a galaxy resides may all influence the lifecycle of a galaxy.

The gaseous and dusty components amongst the stars of a galaxy are referred to as the interstellar medium. In the ISM of a galaxy, clouds with characteristic phases typically form (see Section 1.1). The basic lifecycle of interstellar clouds is illustrated in Figure 1.1. Cold, dense self-gravitating molecular clouds collapse to form new stars. As stars evolve, stellar ejecta influence the structure, physical properties, and chemical composition of the ISM. Gas and dust that has been heated, stirred up, and enriched, might cool and condense into clouds, setting the conditions for new star formation.

The origin and evolution of galaxies are closely tied to the cyclic feedback processes between stars and the ISM. Stellar feedback will disrupt molecular clouds thereby stopping star formation (negative feedback) while at the same time gas is compressed in shells, promoting gravitational instabilities (positive feedback) (Elmegreen & Lada 1977; Hopkins et al. 2014). Phases of the ISM reflect radiative and mechanical energy input by massive stars into their environment (Field et al. 1969; Weaver et al. 1977; McCray & Snow, T. P. 1979; Wolfire et al. 1995). Stellar radiation dissociates, ionizes, and heats gas, and together with the cooling law of atomic gas, sets the temperature and controls the cold and warm phase structures of the ISM. Mechanical energy — input by protostellar outflows, stellar winds, and supernovae — stirs up the medium, controls the pressure of the gas and transfers material from one phase to another. Numerical simulations reveal the active interplay of radiative and mechanical feedback by massive stars (Walch et al. 2012; Kim et al. 2013; Dale et al. 2014; Walch & Naab 2015; Walch et al. 2015; Kim et al. 2018; Haid et al. 2018).

The aim of this thesis is to explore the characteristics of the interstellar medium, on global (galactic) scales down to sub-cloud (pc) scales. We explore new methods to investigate the ISM, through radio recombination line observations, and develop the tools and strategies needed to process new low-frequency observations. We also infer the presence of massive stars and characterize their influence on the ISM, by determining gas physical conditions. In subsequent chapters, we specifically address

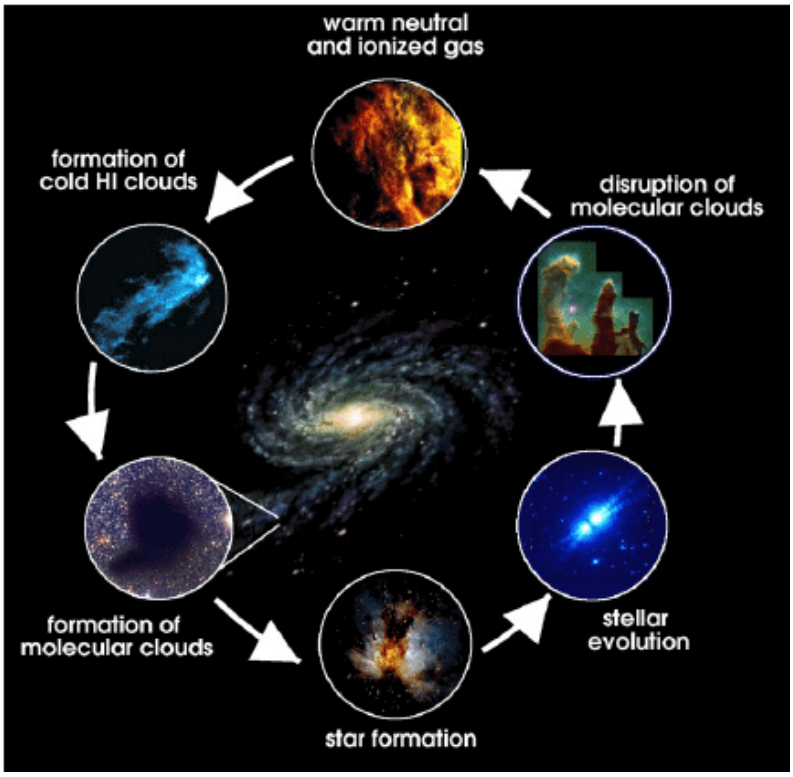

Figure 1.1: The lifecycle of interstellar clouds (Groppi et al. 2009). Starting from the central bottom panel, stars form in dense molecular clouds. As these stars evolve, they develop winds enriching the ISM with nucleosynthetic products. The radiative and mechanical feedback from stars and supernovae disrupt molecular clouds, heating and rarefying the medium. As massive stars die out, warm gas eventually cools. In diffuse clouds, dust grains and a complex chemical network create molecular species, leading to the formation of cold and dense molecular clouds.

the questions:

- How does low-density ionized gas affect the evolution of the massive, galactic star-forming region, Cygnus X? Are the same fingerprints present in surveys of low-density ionized gas in our Galaxy?
- What are the properties of star formation (star clusters) in the central starburst of the galaxy NGC 4945?
- Can the ISM be explored outside of the local universe through radio recombination line observations? What are the ISM properties of a dwarf-like galaxy at $z = 1.1$?
- What techniques are best suited to detect faint radio recombination lines (at an unknown redshift) in extragalactic sources?

In the sections of this chapter, we frame these concepts for the reader, guiding how each of these questions furthers our understanding of processes influencing the lifecycle of a galaxy. We describe phases of the ISM and its basic tracers (Section 1.1), star formation in star clusters (Section 1.2), radio recombination lines as a method to observe gas in the ISM, stars and stellar feedback (Section 1.3), observatories which have enabled the investigations reported in this thesis (Section 1.4), a brief outline of the contents of this thesis in Section 1.5, and a future outlook which builds on the advancements facilitated by this thesis (Section 1.6).

1.1 Phases of the Interstellar Medium

The ISM contains a number of phases — diffuse clouds, warm intercloud gas, and hot intercloud gas — characterized by different temperatures, densities, and ionization fractions and that exist in pressure equilibrium (Tielens 2005). We show a schematic of the ISM in a typical Galactic region in Figure 1.2. We first briefly describe the characteristics of these phases, and in additional sections, we elaborate on cold diffuse clouds, warm ionized media and extragalactic tracers of these phases.

Hydrogen in its atomic state, traced by the H I 21 cm line (van de Hulst 1945), constitutes a large reservoir of gas ($M \sim 5 \times 10^9$ M$_\odot$) in the ISM. It exists in two stable phases primarily as a result of cooling by the fine structure lines of [C II] and [O I] and heating by FUV photons ($6 - 13.6$ eV). Diffuse ($n \sim 50$ cm^{-3}) H I clouds — the cold neutral medium (CNM) — contain about 40% of the H I mass with temperatures of 70 K (Heiles & Troland 2003). The remainder of H I is in the warm neutral medium (WNM) — a phase with the largest filling factor ($\sim$0.4) in the disk — with typical densities of $n \sim 0.5$ cm^{-3} and temperatures of $T \sim 8000$ K, although about half of the WNM lies in the thermally unstable regime with 500 — 5000 K (for a review see Dickey & Lockman 1990; Kalberla & Kerp 2009). When hydrogen is ionized, the gas heats up and temperatures stabilize near $T \sim 10^4$ K as a result of cooling from collisionally excited lines such as [O III] 4959/5007 Å and [N II] 6548/6553 Å, and low-density ionized gas ($n \sim 0.1$ cm^{-3}) — the warm ionized medium (WIM) — fills a considerable volume (filling factor of 0.25) of the disk (for

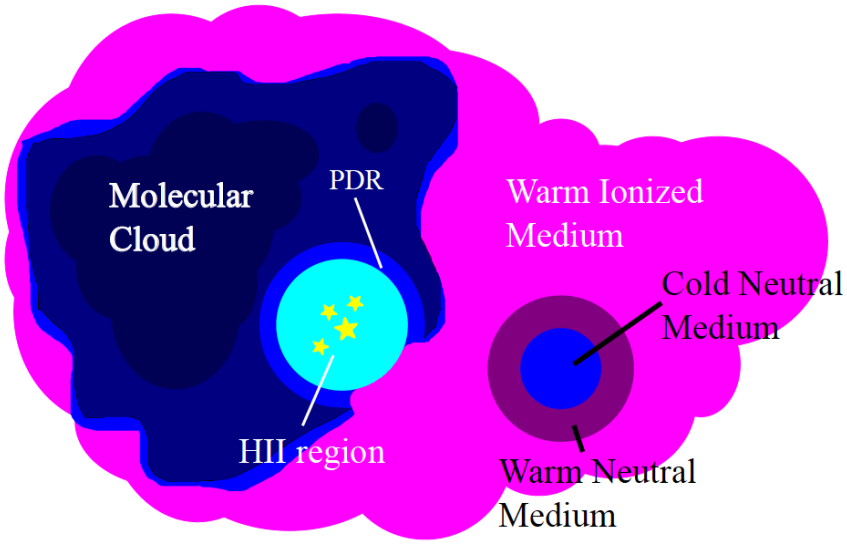

Figure 1.2: A schematic of the classic structure of the interstellar medium in a normal galaxy (Salgado Cambiazo 2015). Phases of the ISM, including the cold neutral medium, warm neutral medium, and warm ionized medium, are depicted. Also shown is a giant molecular cloud in which stars have formed and have created an ionized H II region. Notice the photodissociation region (PDR) at the ionized gas and molecular cloud interface, as well as the cold neutral material on the surfaces of molecular clouds.

a review see Haffner et al. 2009). The hot ($T \sim 10^6$ K) intercloud medium (HIM) of tenuous ($n \sim 10^{-3}$ cm^{-3}) gas has been revealed through X-ray and UV observations. Within the disk it originates through shocks driven by stellar winds and supernovae, and in some cases, it vents out of the disk. Hot gas fills the volume of the halo. Star formation activity is limited to the cold ($T \sim 10$ K) denser ($n \sim 200$ cm^{-3}) molecular (H_2) clouds traced in bulk by CO, i.e., the ^{12}CO(1-0) transition at 2.6 mm. While molecular clouds are in pressure equilibrium with the ISM, they are self gravitating, and thus deep inside the pressure can be much higher.

1.1.1 The Cold, Diffuse ISM

Since warm hydrogen dominates the emission of H I 21 cm observations, the CNM can only be probed when conditions lend themselves to self- and continuum absorption studies (e.g., Dickey & Lockman 1990). Temperatures of cold H I range from 40 – 100 K, weighted by column density the median temperature is 70 K (Heiles & Troland 2003; Dickey et al. 2009). Densities of $n \sim 60$ cm^{-3} (Gibson et al. 2000; Heiles & Troland 2003) have been inferred from a representative pressure of $P/k \approx 3800$ K cm^{-3} (Jenkins & Tripp 2001; Jenkins et al. 2011). The CNM is distributed within 150 pc of the plane (e.g., Kalberla & Kerp 2009). While the classic CNM cloud size is a few pc, observations have revealed a considerable fraction with sheet-like and filament morphology and some regions with globular substructure on sub-pc (< 0.1 pc) scales (Heiles 1997; Gibson et al. 2000; Heiles & Troland 2003). These sizes contrast markedly with the 600 pc cloud sizes of the WNM (e.g., Kalberla & Kerp 2009).

The 21 cm line does not constrain the physical conditions of the gas (e.g., temperature, density) well, making it hard to identify the full extent and physical conditions of the CNM. Observation are limited to select lines of sight where H I can be observed in absorption and/or at high (less confusing) galactic latitude. Optical and UV absorption line studies — which are good probes of the physical conditions — are limited to pinhole experiments towards bright, nearby stars with only low column densities of gas. While the CNM shines in [C II] 158 μm emission, so do denser regions of PDRs and ionized gas (Heiles 1994), making it difficult to elucidate the cold, diffuse component.

A key missing link in the galactic lifecycle is thus the interrelationship of neutral gas and molecular gas. The traditional three-phase model of the ISM (Mckee & Ostriker 1977) cannot reconcile the significant presence of luke-warm ($T \sim 500 - 5000$ K), thermally unstable H I gas in the ISM (Heiles & Troland 2003). In addition, the presence of "translucent" clouds has long been recognized (van Dishoeck & Black 1988) and its general importance has been driven home by γ-ray and far-IR studies, indicating that much of the molecular gas mass is in a form not probed by CO (so-called CO-dark molecular gas) (Grenier et al. 2005; Planck Collaboration et al. 2011). For H I diffuse clouds, hydrogen is atomic and carbon is ionized (C+). In CO-dark molecular gas, hydrogen is in H_2 but C+ rather than CO is the dominant form of carbon (Visser et al. 2009; Wolfire et al. 2010). The major mass constituents of our Galaxy of cold H I and CO-dark phases have so far eluded detailed characterization.

1.1.2 Warm Ionized Media

The warm ionized medium, with a typical density of 0.1 cm^{-3}, occupies a large volume (filling factor of 0.25) in the disk of our Galaxy, is characterized by large ($z \sim 1$ kpc) scale heights, and contains about $M_{\rm ion} \sim 10^9$ $M_\odot$(for a review see Haffner et al. 2009). Within the plane of the Galaxy, ionized gas is found in (i) compact ($R <$ a few pc) H II regions ($n > 10^3$ cm^{-3}) associated with O stars in the early stages of their interaction with the ISM when they are still embedded in (i.e., interacting with) their birth environment, (ii) the diffuser ionized gas due to stars drifting away from the association and/or when large (wind) bubbles have broken open and vented their ionized gas (and hot plasma) into the environment, and (iii) an ionized medium surrounding neutral clouds. Mezger (1978) first emphasized the importance of an ionized gas phase that fills a substantial portion of the ISM in the plane of the Milky Way with a density of $\sim$5 cm^{-3} and a temperature of 7000 K. Recent studies of the [N II] fine-structure lines indicate that an even denser phase of ionized gas ($\sim$50 cm^{-3}) is also prominent in the ISM of the Milky Way as well as other galaxies (Goldsmith et al. 2015; Pineda et al. 2019).

Massive stars of type O and earlier than B3 produce EUV photons ($E > 13.6$ eV) capable of photoionizing gas. Although the ionization of the WIM involves less than 1% of the produced ionizing photons, it does require that the ISM is structured in a way (also considering the motion of stars) that enables EUV photons to escape their immediate environment. Outstanding open questions in studies of diffuse ionized gas include (a) possible heating sources of the WIM by non-ionizing means (photoelectric heating from dust grains, dissipation of turbulence, shocks and cooling hot gas), (b) the distribution, morphology and physical properties of diffuse ionized gas in the midplane, and (c) how many ionizing photons escape their immediate environments as well as the disk and galaxy. In addition, the characteristics of the ionized gas phases, their distribution in the galaxy, and their relationship to each other and to compact HII regions and their ionizing stars is not well understood.

Optical lines (e.g., of Hα, [N II], [O II] and [S II]) and pulsar dispersion measures (DM) have been major workhorses in characterizing the WIM. Somewhat denser gas has been uncovered through thermal radio continuum and high-frequency radio recombination lines (for more details, see Section 1.3). However, they have not (yet) provided answers to remaining open questions. Particularly close to the Galactic plane, optical lines are heavily attenuated, and the DM only provides an average density along the line of sight, without much information about inhomogeneities. Through these means, physical properties cannot be determined (outside of classic H II regions) that would quantify the direct impact of mechanical and radiative processes from (massive) stars.

1.1.3 Extragalactic Tracers of the ISM

Looking outside the Milky Way, in nearby galaxies H I 21 cm observations probe the majority of the warm neutral component. FIR atomic fine structure lines of [O I] 63 μm and [C II] 158 μm trace neutral components, but [C II] also arises from cold neutral and warm ionized phases. Molecular gas content can be observed in bulk

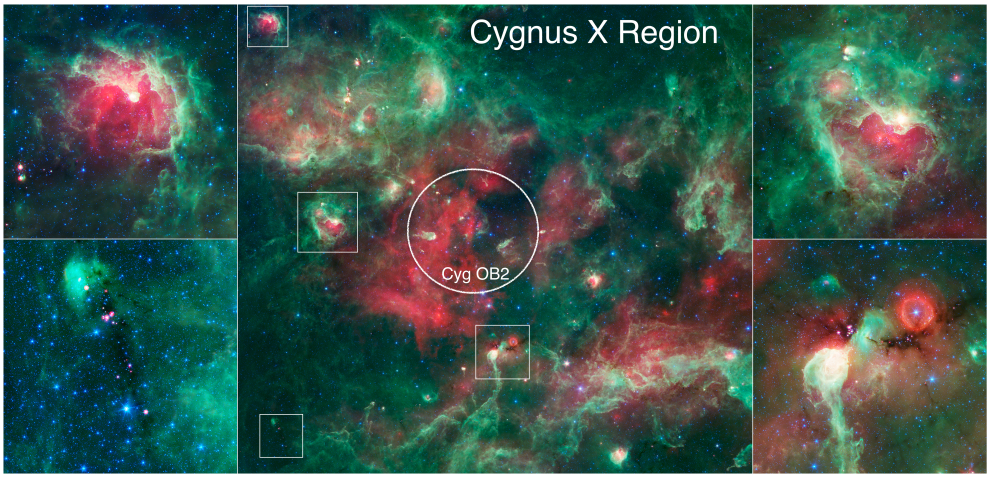

Figure 1.3: The Cygnus X star-forming region, a bubbling cauldron of star birth surrounding the Cyg OB2 association — observed with the *Spitzer* Space Telescope: 3.6 μm in blue, 4.5 μm in blue-green, 8.0 μm in green, and 24 μm in red. PDRs traced by UV-heated PAHs appear bright in green-white. Massive stars that have blown bubbles and cavities in the dust and gas are evident where green-white PDR emission surrounds a bubble of glowing red emission from warm dust (ionized gas). The brightest, yellow-white regions are warm centers of star formation. Tendrils of dust appear green, and stars generally appear as blue point sources. The pillar-like and elongated features directed away from Cyg OB2 have been shaped by the stellar radiation and winds from this massive ($M_\star \approx 2 \times 10^4$ M$_\odot$) association. The boxed zoom-in regions show massive star(s) (formation) in AFGL 2636 (upper left) and DR22 (upper right). The lower left and right (DR15) images show clouds that are so thick to be absorbed at the mid-infrared wavelengths of Spitzer. Young stars, visible as red points, are buried in the dark clouds. Image credit: NASA / JPL-Caltech / Harvard-Smithsonian.

by CO (and isotopes and transitions up the ladder) and several other molecules such as HCN, CS, HCO^+, SO, etc. As in the Milky Way, directly detecting the emission of cold diffuse clouds is difficult. Observations of the [C I] 370 and 609 μm lines help, and cold neutral gas can be inferred using [C II] or FIR dust continuum observations together with CO and H I (e.g., Bolatto et al. 2013a; Herrera-Camus et al. 2017). Ionized gas can be studied to deep emission measures through hydrogen recombination lines at $\lambda \lesssim 2$ μm (i.e., Lyα, Hα, Hβ, Paα) and other optical emission lines such as [N II], [O II] and [S II], and through FIR fine structure lines of [C II] 158 μm, [N II] 122 and 205 μm and [O III] 52 and 88 μm. However, the WIM (and/or diffuse ionized gas) is difficult to differentiate between the discrete (or unresolved) sources of stars in face-on galaxies (e.g., Oey et al. 2007). Studies of edge-on systems have been particularly illuminating for the WIM (e.g., Rossa & Dettmar 2003a,b; Levy et al. 2019).

When we look outside of the local universe, the neutral gas content is relatively unknown. Neutral H I has been detected in individual systems to $z \lesssim 0.2$. The neutral gas content of the universe is probed in large numbers through e.g., Lyα and Mg II absorption systems in quasar spectra, though studying individual systems responsible for absorption has been difficult and the origin of the absorption systems remains under investigation (Wolfe et al. 2005). H I 21 cm absorption from intervening and associated systems has so far been detected in just over 200 systems (e.g., Kanekar & Briggs 2004; Morganti & Oosterloo 2018). [C II] seems to predominantly trace the ionizing photon luminosity through PDRs emission in galaxies (though not well for galaxies with FIR luminosities of $> 10^{11}$ $L_\odot$); with ALMA, galaxies of $z > 4$ can be investigated in this manner (but it is difficult in intermittent redshifts) (e.g., Carilli & Walter 2013). Galaxies have traditionally been probed across all redshifts through some of the brightest emitting lines of ionized gas, though the contribution from diffuse gas and ionized gas is not often possible to distinguish. Before ALMA, CO molecular gas was only detected in ∼200 systems, but this has steadily been increasing revealing the molecular gas history of the universe (e.g., Decarli et al. 2016). In conclusion, it is challenging to trace the ISM of galaxies outside of the local universe.

1.2 (Clustered) Star Forming Regions

The vast majority of massive stars form in clusters (we refer the reader to reviews by Lada & Lada 2003; Portegies Zwart et al. 2010; Krumholz 2014; Krumholz et al. 2019). 70% of O stars reside in young clusters or associations (Gies 1987). Furthermore, most of the field population can be identified as runaways (de Wit et al. 2005). Massive young clusters tend to form with compact (∼1 pc) sizes or in somewhat extended regions (∼7 pc) which may be unbound (associations). Giant molecular clouds, the birth sites of stars and star clusters, have typical sizes of 50 pc and masses of $\sim 10^5$ $M_\odot$. Filaments and clumps form within them through the combined action of turbulence and self-gravity, and massive clusters tend to form at the intersection of filaments. Cloud collisions may also be important for their formation.

In the Galaxy, the nearest young cluster that has formed O stars is 414 ± 7 pc away (Menten et al. 2007) in Orion. The Orion Nebula Cluster (ONC), with a total stellar mass of ∼4000 $M_\odot$ and age of $1 - 2$ Myr (e.g., Zari et al. 2017, and references

therein), is one of the most comprehensively studied regions of massive star formation. Its massive stars, an O7 star and two early B types, have created a bubble filled with hot X-ray gas and warm ionized gas that is a few pc in diameter and sits at the edge of its parent molecular cloud (e.g., Pabst et al. 2019). In the association Cyg OB2 ($d \approx 1.5$ kpc), massive star formation ($M_\star \sim 2 \times 10^4$ M$_\odot$) has occurred in a region of relatively low mass surface density and the system is not gravitationally bound (see Figure 1.3). Cyg OB2 has an age spanning 3-5 Myr and thus is no longer actively forming stars. While Cyg OB2 has clearly been influential in dispersing its molecular cloud — a ∼100 pc region has been cleared and ionizing radiation continues to eat away at the nearby clouds — evidence for (subdominant) triggered star-formation is also present in small globules and pillars (e.g., Schneider et al. 2016). Moreover, the Cygnus X region is accompanied by new regions of active star formation spread across the entire complex (see Figure 1.3).

The most massive ($> 10^5$ M$_\odot$), young clusters — super star clusters (SSCs) — in the local universe are often found in starbursting regions and merging galaxies (e.g., Portegies Zwart et al. 2010; Whitmore et al. 2010; Linden et al. 2017). The efficiency with which matter is converted into stars is higher in SSCs — $\epsilon \sim 0.5$ compared with the typical value of ∼0.05. They provide key insights into galaxy formation as cluster formation is expected to be a main mode of star formation in high intensity, starburst environments (e.g., Kruijssen 2012) and thus possibly dominant during the peak epoch of star formation in the universe ($z \sim 1 - 3$). The presence of globular clusters also indicates that star formation in massive clusters was a hallmark of the era of galaxy formation.

In nearby starbursting galaxies ($d < 4$ Mpc), such as M 82 (e.g., Tsai et al. 2009), NGC 253 (Leroy et al. 2018) and NGC 4945 (Emig et al. 2020b), SSCs populate the nuclear starbusting regions. In M 82, 15 SSCs or more are found within the central starburst (see Figure 1.4; McDonald et al. 2002). The concerted efforts are capable of driving galaxy scale outflows, reflecting the profound impact they have on lifecycle of galaxies (see Figure 1.4).

1.3 Radio Recombination Lines

Radio recombination lines provide a powerful tool to study the characteristics of neutral and ionized phases of the ISM. As ions and electrons recombine to form atoms, electrons left in high principal quantum numbers and which cascade to lower energy levels (or may be excited to higher energy levels) are observable through the spectral signature of radio recombination lines (RRLs). RRL transitions in these Rydberg states have (some of the) properties of the Bohr model of the atom. RRLs are typically observed at the radio frequencies of 350 GHz ($\lambda \sim 0.86$ mm) down to 10 MHz ($\lambda \sim 30$ m), corresponding to low energy α-transitions ($\Delta\mathsf{n} = 1$) of principal quantum numbers $\mathsf{n} = 26$ – 869 [1]. At such high principal quantum numbers, for example at the largest detected (δ; $\Delta\mathsf{n} = 4$) transition in space of $\mathsf{n} = 1009$ (Stepkin

[1] In this notation, n represents the principal quantum number of the final state

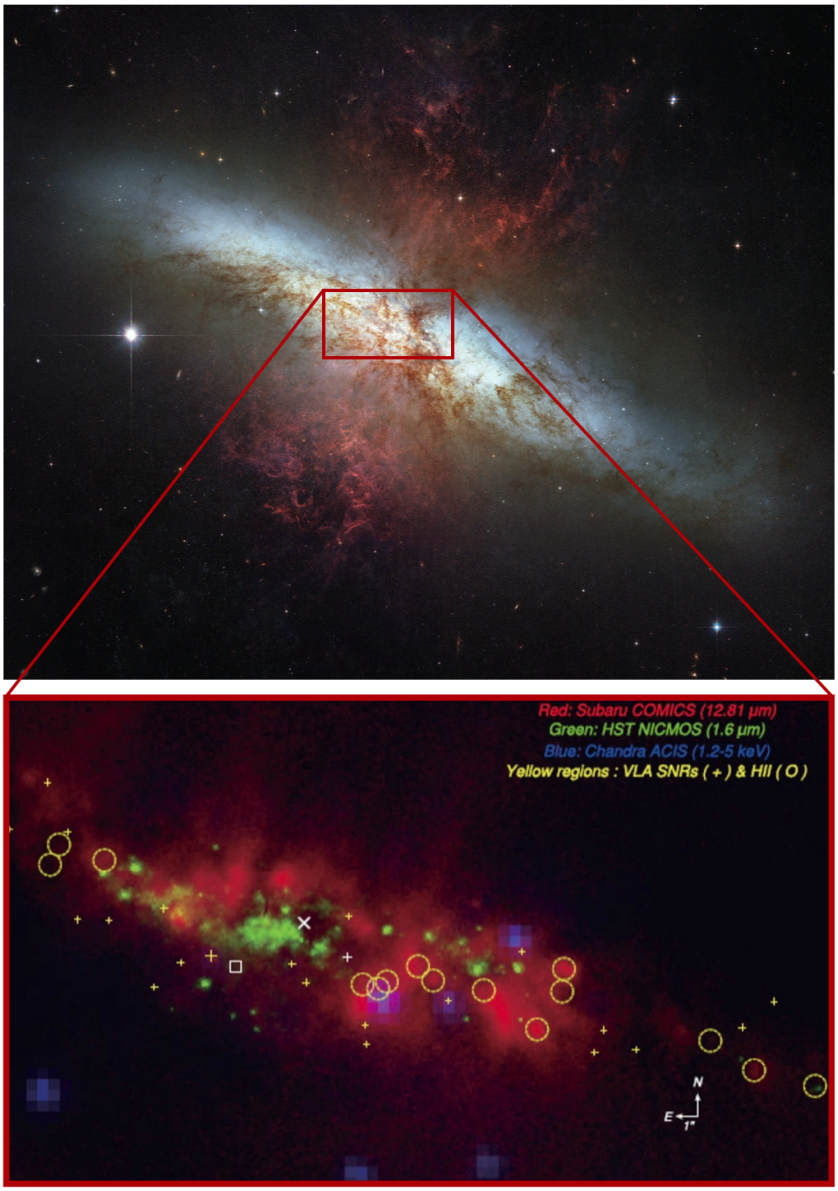

Figure 1.4: In the top panel, the disk of the nearby (3.5 Mpc) edge-on galaxy, M82, glows in white-blue and the prominent galactic winds from the nucleus are traced by the PAH emission in red. Super star clusters (yellow circles) in the nucleus (in the bottom panel) drive the outflow. Image credits: top panel NASA / ESA / The Hubble Heritage Team and bottom panel from Gandhi et al. (2011).

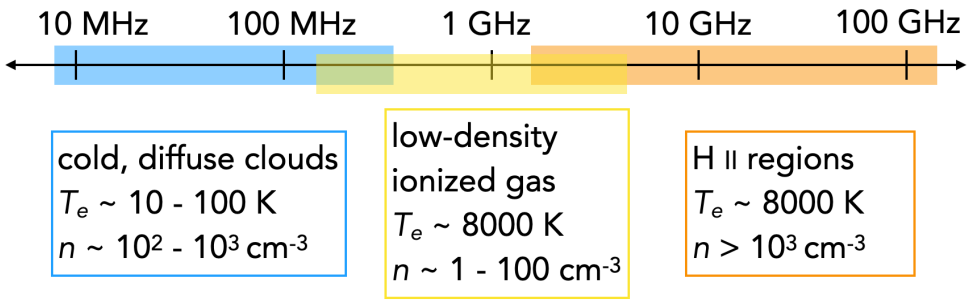

Figure 1.5: The gas components and their frequencies traced by radio recombination lines. Emission from cold diffuse clouds is probed by $\nu \lesssim 500$ MHz, low-density ionized gas by 300 MHz $\lesssim \nu \lesssim 5$ GHz, and classic H II regions by $\nu \gtrsim 1$ GHz.

et al. 2007), the Bohr diameter of an atom[2] reaches $(2n^2\hbar)/(Zk_e e^2 m_e) \approx 110$ μm — for scale, the average width of a human hair is roughly 75 μm!

Observations have now sampled the full frequency coverage of the RRL spectrum, primarily through hydrogen, carbon, and helium RRLs. Broadly speaking, three gas components have been found to dominate line emission, as shown in Figure 1.5. At high frequencies ($\gtrsim 10$ GHz), recombination lines arise from spontaneous transitions in dense ($n_e \sim 10^3 - 10^4$ cm^{-3}) ionized gas — conditions associated with compact H II regions and PDRs. At the intermediate frequencies of 0.3 – 1 GHz, emission by a warm and low-density ($n_e \sim 1 - 100$ cm^{-3}) ionized component dominates. Whereas at low frequencies ($\lesssim 250$ MHz), recombination lines arise in the cold-neutral medium stimulated by external radiation fields; carbon is typically the most prominent species, and spectral lines can be observed in either absorption or emission.

In this section we introduce and discuss the physics, primary use cases, and at times some historical information of the three different frequency regimes of RRL emission. Lastly, we give special emphasis to extragalactic RRLs with additional background. We refer the reader to "Radio Recombination Lines: Their Physics and Astronomical Applications" (Gordon & Sorochenko 2002) as well as Roelfsema & Goss (1992) for additional information.

1.3.1 High frequency recombination lines

The most common observations of RRLs are Galactic in origin and arise from spontaneous transitions found in optically-thin free-free emitting gas. When the electron temperature T_e of the thermal free-free emission from ionized gas accurately characterizes the relative populations of electrons in bound atomic levels, the system is well approximated by local thermodynamic equilibrium (LTE). Ionizing radiation is balanced by recombinations and line emission is described by the Saha-Boltzmann distribution. This description characterizes RRLs at frequencies of around 5 – 20 GHz arising from typical H II regions with densities of $10^3 - 10^4$ cm^{-3} and electron temperatures of ∼7000 K.

[2] $\hbar$ is the reduced Planck constant, Z is the atomic number, k_e is the Coulomb constant, e is the charge of an electron, and m_e is the mass of an electron.

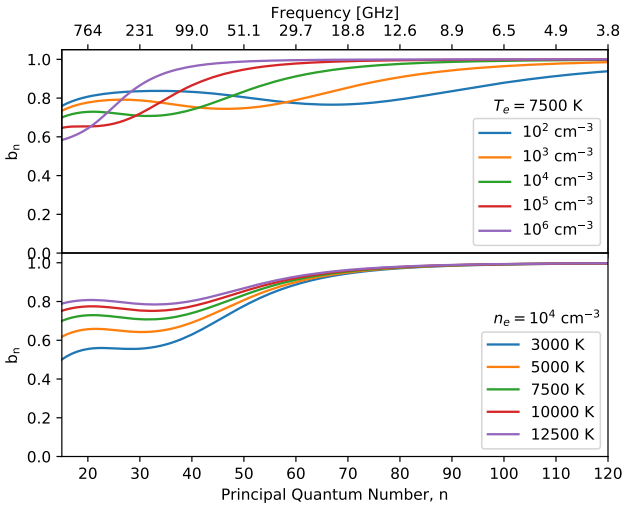

Figure 1.6: The departure (from LTE) coefficient of spontaneous emission, b_n, shown as a function of principal quantum number (frequency). Top shows the variation with density for a single temperature ($T_e = 7500$ K. Bottom shows the variation with temperature for a single density ($n_e = 10^4$ cm^{-3}).

As the recombination cross section of a Bohr atom rapidly decreases ($\propto \mathsf{n}^4$) towards smaller principal quantum numbers and thus higher frequencies, of order $\mathcal{O}(100$ GHz), collisions between protons and electrons become less important, and the observed line intensity is not exactly set by the kinetic motion of the electrons. Radiative processes, which dominate the smallest principal quantum numbers ($\mathsf{n} \lesssim 10$), influence the level populations; smaller n are underpopulated as compared with a Boltzmann distribution. The population rates into an energy level are not exactly balanced by rates out of the level, and a correction coefficient (b_n) for a departure from LTE is necessary. Multiplicative b_n factors to the line intensity, as shown in Figure 1.6, at a given n depend (mildly) on gas conditions found in typical H II regions; however for low-density gas, they become more important and variable across frequency.

The physics that is extracted through high-frequency recombination lines has been a major workhorse in studies of photo-ionized gas and massive star formation in the local universe. The intensity of emission from a RRL is dependent upon the electron temperature and the emission measure (EM)[3] of the ionized atoms. As the ionization balance is in steady state, recombinations are balanced by ionizations. (We refer the reader to Chapter 3 for more detailed equations.) The ionizing photon rate which maintains the ionized state of the gas is also directly proportional to the EM (and temperature through the recombination coefficient). Connecting these properties, the RRL intensity effectively measures the ionizing photon rate. In obscured regions, which may be in our Galactic plane or in extinguished regions in other (especially local, edge-on) galaxies, RRLs provide access to total ionizing luminosities (and thus star formation rates) free from dust extinction. High-frequency RRLs can also be used

[3] $EM = \int n_e n_i \, \mathrm{d}\ell$ where n_e and n_i are the volume number density of electrons and ions, respectively, and ℓ is the pathlength integral.

to extract the electron temperature when combined with continuum observations (at the same frequency); since the volume of gas emitting line and continuum emission is the same, dependencies on the EM, beam filling factor and distance to the source cancel out when taking the line to continuum ratio and only the temperature remains to be directly solved. With the kinematic information provided by the lines and if the region is spatially resolved, components can be distinguished and thus pathlengths determined in order to also infer electron densities. Deeper observations may also detect helium RRLs (offset by just 122 km s^{-1} from hydrogen lines) which follow these same physical principals for emission; by comparing the integrated line strengths, the helium abundance is determined.

1.3.2 Low frequency recombinations lines from cold partially ionized gas

Recombination lines at low frequencies ($\lesssim 250$ MHz) inherently arise from relatively cold gas, $T_e = 10 - 100$ K, with low electron densities, $n_e = 0.01 - 0.1$ cm^{-3} (Shaver 1975b) — characteristic conditions of the diffuse ISM. Since background radiation ($I_{\rm BC}$) typically dominates over the continuum emission of the cloud at low frequencies, stimulated effects dictate the RRL line intensity (Goldberg 1966), and the optical depth of the line is approximated by $-\tau_{\rm line} \approx I_{\rm line}/I_{\rm BC}$. Under these conditions, the observed emission depends on temperature as $T_e^{-2.5}$ (Shaver 1975b; Salgado et al. 2017b) and thus preferentially traces cold clouds. Low-frequency RRLs are quite faint — emitting/absorbing 0.1 – 0.01% of the continuum — which attributes to the difficulty in detecting them; however, as the lines appear more frequently at low frequencies (see Figure 1.9), line stacking (e.g., Emig et al. 2020a) can aid in their detection.

Carbon is typically the brightest RRL emitter at the lowest frequencies for a number of reasons. Carbon has a lower ionization potential ($E = 11.3$ eV) than hydrogen (13.6 eV) and remains ionized when the bulk of the gas (traced by hydrogen) is neutral. In the cold, diffuse ISM where ionizing radiation can not excite hydrogen, the [C II] $^2P_{1/2}$ – $^2P_{3/2}$ fine structure line at 158 μm serves as the dominant coolant (Tielens 2005) and temperatures stabilize near 70 K (Heiles & Troland 2003). Carbon is abundant (compared with other "metals") in the ISM (Cardelli et al. 1996). Ionized carbon at these temperatures also has the unique property of dielectronic capture (Watson et al. 1980), in which an electron recombines with a carbon ion by simultaneously exciting the [C II] $^2P_{1/2}$ – $^2P_{3/2}$ fine structure line, leaving additional (compared with LTE) carbon atoms with electrons in high Rydberg states. Dielectronic capture greatly enhances RRL emission at high quantum levels ($\mathsf{n} \gtrsim 250$; Salgado et al. 2017a).

Carbon RRLs are an unprecedented physical diagnostic of the diffuse ISM. Being maser-like, the distribution of the quantum level population, characterized by the correction factor for stimulated emission β_{n} and the departure coefficient b_{n}, is strongly dependent upon the temperature and density conditions of the gas (see Figure 1.7; Salgado et al. 2017a,b). This product, $b_{\mathsf{n}}\beta_{\mathsf{n}}$, which is directly proportional to the integrated line strength, is set by the relevant interactions which influence atomic physical processes; only recently have computing resources allowed for (i) the full

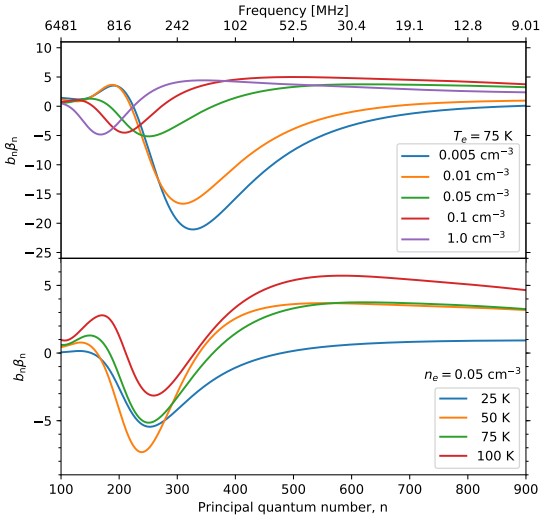

Figure 1.7: The strong dependence of carbon RRL emission (as a function of principal quantum number) on the physical conditions of cold diffuse clouds, for which the product of the departure and stimulation coefficients ($b_n\beta_n$) is directly proportional to the line intensity. Carbon RRLs show an enhancement in emission — where negative values of $b_n\beta_n$ indicate emission and positive values indicate absorption — at small principal quantum numbers (high frequencies), and they are often observed in absorption at the lowest frequencies.

angular momentum treatment for carbon and (ii) with transitions to large enough principal quantum number to be included (Salgado et al. 2017a). Because the density and temperature are constrained simultaneously by both the relation for the line intensity and the $b_n\beta_n$ coefficients, the pathlength can also be uniquely determined. (We refer the reader to Chapter 4 for the detailed equations.) The capability to derive the density, temperature and pathlength in the diffuse ISM is revolutionary because traditional probes of the ISM have not, apart from temperature, provided this physical information (e.g., Heiles & Troland 2003; Roshi & Kantharia 2011).

Detailed studies of gas along the line of sight to Cassiopeaia A, a supernova remnant in our Galaxy and one of the brightest radio sources in the sky at low frequencies, have demonstrated and shaped our understanding of carbon RRLs (Kantharia et al. 1998, and references therein) (Asgekar et al. 2013; Salas et al. 2017; Oonk et al. 2017; Salas et al. 2018; Chowdhury & Chengalur 2019). Large-area surveys in the Galaxy have also demonstrated the ubiquity of low-frequency carbon RRL emission, especially using the Ooty Radio Telescope (Anantharamaiah 1985a; Erickson et al. 1995; Kantharia & Anantharamaiah 2001; Roshi et al. 2002). However, because of the spatial resolutions (with the necessary sensitivity) and limited frequency coverage to constrain gas properties, a clear characterization of the origin of the emission in the ISM has not been made. Moreover, the important implication of stimulated emission possibly detectable in galaxies outside of the local universe (Shaver 1978) — since the line emission is proportionally to the continuum intensity of background radiation —

has not (yet) been fully utilized in extragalactic objects.

Ionized hydrogen may also be stimulated by an external radiation field at low-frequencies (e.g., Shaver 1975b,a; Anantharamaiah 1985a; Roshi & Anantharamaiah 2000, 2001). Indeed hydrogen RRLs have been observed from a cold medium (Gordon & Sorochenko 2002, and references therein) (Lockman 1989; Roshi & Anantharamaiah 1997, 2001; Oonk et al. 2017, 2019; Emig et al. 2020a). Cold ionized hydrogen layers are thought to originate in photoevaporating flows (typically when observed in conjunction with a broad hydrogen RRL of a typical H II region) or maintain ionization through cosmic rays thus probing cosmic ray ionization (for a review see Gordon & Sorochenko 2002).

1.3.3 Hydrogen RRLs from warm low-density ionized gas

At the intermediate frequencies of 0.3 – 1 GHz, a warm (T$\sim$7000 K) and low-density ($n_e \lesssim 10^3$ cm^{-3}) ionized component dominates hydrogen RRL emission in the plane of the Galaxy (Shaver 1976; Lockman 1976; Anantharamaiah 1985b, 1986; Heiles et al. 1996b; Roshi & Anantharamaiah 2000; Alves et al. 2015). Firstly, it is possible this gas may be dominated by spontaneous emission from ionized gas that has a lower (electron) density than a typical H II region and would be characterized by emission measures of around $EM \lesssim 10^6$ pc cm^{-6}; in this case, non-LTE effects need to be considered. Secondly, the RRL emission may be dominated by stimulated effects — thus largely characterized by the properties of low-frequency RRLs as described in Section 1.3.2 — and would arise from gas of lower density than in the spontaneous case (i.e., $n_e \lesssim 10$ cm^{-3}). Thirdly, spontaneous and stimulated effects may both influence the line intensity. Without sampling the line properties across frequency and comparing it with continuum emission at a high enough resolution, it is difficult to distinguish between these cases (e.g. Shaver 1976). While RRLs have been detected with stimulated effects (Pedlar et al. 1978; Anantharamaiah 1985a; Heiles et al. 1996a; Roshi & Anantharamaiah 2000), in regions separated from ionizing sources (Lockman 1976; Heiles et al. 1996b), and at low frequencies (Shaver 1976; Anantharamaiah 1985a; Roshi & Anantharamaiah 2001), pinning down the origin and properties of gas probed in this frequency regime remains open.

1.3.4 Extragalactic observations

RRLs detected in external galaxies (beyond the Magellanic Clouds) have primarily been observed at frequencies above 1 GHz from classic H II-emitting regions in star-bursting regions, accounting for 21 galaxies detected to date[4]. We refer the reader to

[4]To the best of our knowledge, detections of RRLs above 1 GHz have been reported in the following extragalactic objects: M 82 (Shaver et al. 1977; Chaisson & Rodriguez 1977; Bell & Seaquist 1977, 1978; Shaver et al. 1978; Bell et al. 1984; Seaquist et al. 1985; Puxley et al. 1989; Roelfsema & Goss 1992; Seaquist et al. 1994, 1996; Rodriguez-Rico et al. 2004), NGC 253 (Seaquist & Bell 1977; Mebold et al. 1980; Anantharamaiah & Goss 1990, 1996; Puxley et al. 1997; Mohan et al. 2002, 2005; Rodriguez-Rico et al. 2006; Kepley et al. 2011; Bendo et al. 2015; Eisner et al. 2019), MRK 668 (OQ 208; Bell & Seaquist 1980; Bell et al. 1984), NGC 2146 (Puxley et al. 1991; Zhao et al. 1996), NGC 3628 (Anantharamaiah et al. 1993; Zhao et al. 1997), IC 694 (merging galaxy in ARP 299; Anantharamaiah et al. 1993; Zhao et al. 1997), NGC 1365 (Anantharamaiah et al. 1993), NGC 3690 (merging galaxy in ARP 299; Anantharamaiah et al. 1995; Zhao et al. 1997), ARP 220

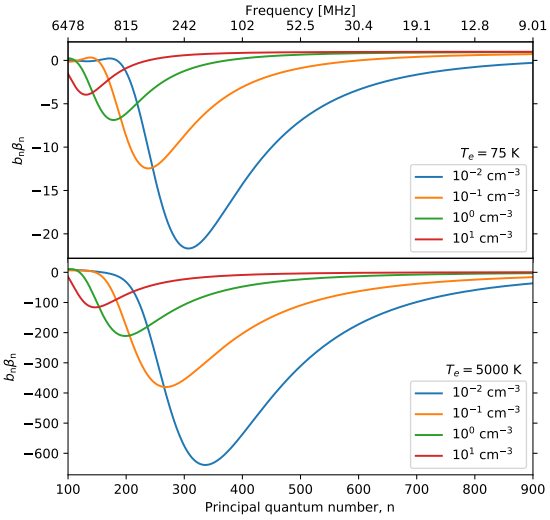

Figure 1.8: The strong dependence of the departure and stimulation coefficients ($b_n\beta_n$) from hydrogen RRL emission on the electron density of cold, partially ionized hydrogen (top, with $T_e = 75$ K) and warm, low-density ionized gas (bottom, with $T_e = 5000$ K).

Gordon & Sorochenko (2002) for a thorough discussion laying out the early advancements of probing and establishing the first extragalactic detections. Briefly, the first detections were established in M 82 by Shaver et al. (1977) and NGC 253 by Seaquist & Bell (1977), and only in the early 1990s were other galaxies detected, in large part due to renewed interest with the Very Large Array (e.g., Anantharamaiah et al. 1993). Through these and subsequent observations, star formation rates, ionized gas temperatures and densities, properties of super star clusters, and stellar kinematics have been established in the central starburst regions of nearby galaxies. Because the lines are relatively faint and trace starbursting regions (at galaxy centers), searches to detect high-frequency RRLs have not always been fruitful i.e., at low resolution, in the presence of a (Seyfert) AGN, and at frequencies below which (edge-on) starburst regions become optically thick to thermal continuum emission (e.g., Bell & Seaquist 1978; Bell et al. 1984; Anantharamaiah et al. 1995; Zhao et al. 1996; Phookun et al. 1998; Roy et al. 2008; Izumi et al. 2016; Luisi et al. 2018; Eisner et al. 2019).

Some of the extragalactic RRL observations at frequencies of about 1 – 10 GHz give support for line emission which is enhanced compared to LTE values as inferred from a comparison with thermal continuum components. In some sources (and at

(Anantharamaiah et al. 1995; Zhao et al. 1996; Anantharamaiah et al. 2000; Rodriguez-Rico et al. 2005), M 83 (Anantharamaiah et al. 1995; Zhao et al. 1996), NGC 4945 (Anantharamaiah et al. 1995; Roy et al. 2010; Bendo et al. 2016; Emig et al. 2020b), Circinus (ESO 97-G13; Anantharamaiah et al. 1995; Roy et al. 2008), NGC 660 (Phookun et al. 1998), NGC 5253 (Mohan et al. 2001; Rodriguez-Rico et al. 2007; Bendo et al. 2017), Henize 2-10 (Mohan et al. 2001), M 33 (toward NGC 604 Araya et al. 2004), NGC 3256 (Roy et al. 2005; Michiyama et al. 2020), NGC 1808 (Roy et al. 2008), IC 342 (Balser et al. 2017), M 51 and NGC 628 (Luisi et al. 2018).

particular frequencies) excess LTE values result once the free-free continuum starts to become optically thick and stimulates the line emission in a narrow frequency range (e.g., Anantharamaiah et al. 1993; Seaquist et al. 1996; Zhao et al. 1996, 1997; Phookun et al. 1998; Mohan et al. 2002, 2005; Rodriguez-Rico et al. 2006; Balser et al. 2017). Also, in a number of sources, a(n additional) low-density component which is stimulated by non-thermal background radiation is indicated by modeling the line and continuum emission. For example, the very first extragalactic detections in M 82 indicated stimulated emission was a major factor in the line intensities and originated from gas in front of a background non-thermal source (Shaver et al. 1977; Bell & Seaquist 1977; Shaver et al. 1978; Bell & Seaquist 1978; Roelfsema & Goss 1992; Rodriguez-Rico et al. 2004). Modeling the RRL emission from ∼100 GHz down to 1 GHz in NGC 253 and ARP 220, Mohan et al. (2005) and Anantharamaiah et al. (2000), respectively, showed that stimulation of a lower-density ionized component likely dominates emission below about 4 GHz (see also Mebold et al. 1980; Anantharamaiah & Goss 1990; Zhao et al. 1996; Mohan et al. 2002; Eisner et al. 2019). NGC 4945 is the brightest RRL emitter at and below 8 GHz, detected at 3″ with the Australia Telescope Compact Array (Roy et al. 2010); stimulated effects are possibly present. Detections of RRLs in the Circinus galaxy and NGC 1808 with the same observing set up as NGC 4945, find even stronger contributions of stimulated emission (≥90%; Roy et al. 2008). Early RRL observations of MRK 668 (OQ 208; Bell & Seaquist 1980; Bell et al. 1984) at 2.6′ (4.3′) resolution at 10.5 (6.1) GHz showed that stimulated RRL emission is detectable — with a peak line to continuum ratio of 0.3 (0.5) % — in a GHz-peaked radio source at a redshift of $z = 0.0763 \pm 0.0002$.

The RRL results coming from extragalactic observations in the first few years indicated that stimulated emission is important (Shaver et al. 1977; Bell & Seaquist 1977; Shaver et al. 1978; Bell & Seaquist 1978; Mebold et al. 1980; Bell & Seaquist 1980) and theory was formalized (Shaver 1976, 1978). The implication that RRLs could be observed in radio-bright high-z sources was realized. However, two additional (single dish) searches to deep sensitivities but with the narrow bandwidths at the time, totaling 33 sources in all, were not promising (Churchwell & Shaver 1979; Bell et al. 1984). Using the 305 m Arecibo Observatory, (Churchwell & Shaver 1979) targeted radio sources with redshifts in the range $z = (0.002 - 3.530)$ and which have at least one of the following qualities: (a) a low-frequency turnover in the SED, (b) low polarization, (c) low-frequency variability, (d) a fairly strong continuum, (e) a large optical emission-line region, (f) low electron density of ionized gas indicated by optical forbidden lines, (g) narrow optical lines, since the bandpass fidelity limits line detections to widths of ≤400 km s^{-1}, and (h) infrared emission. Observations of 21 sources at 1.4 GHz with a 10 MHz bandwidth receiver resulted in no detections[5] with a typical (median) 5σ limit to the peak line flux of 10 mJy and a peak line to continuum ratio of ∼ 0.1%. The observations had a 4.2 km s^{-1} channel width, but the upper limits consider a 400 km s^{-1} line width. In the same study, three sources were observed — 4C 28.25, 3C 270.1, and 3C 298 — at 430 MHz with a 2.5 MHz bandwidth receiver targeting carbon RRLs in intervening absorption systems. Upper limits for a 10 km s^{-1} line width are a peak line flux density of 7 – 18 mJy and a

[5]A tentative detection was found in the local source NGC 4631 but has not been followed up or confirmed since.

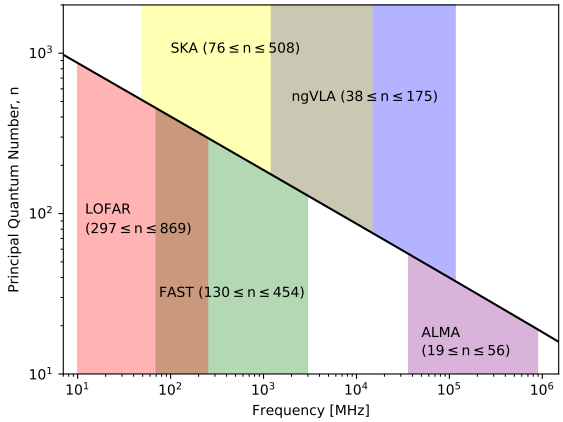

Figure 1.9: The principal quantum number, n, of the α-transitions ($\Delta\mathsf{n} = 1$) of radio recombination lines plotted as a function of frequency. We highlight the (planned) frequency ranges of telescopes that are currently operable or will be in the foreseeable future and that provide among the best sensitivity and resolution to characterize cold and warm diffuse gas and dense ionized gas in extragalactic sources through RRL observations.

peak line to continuum ratio of $(0.09 - 0.4)\%$. The other high-z search for RRLs was reported by Bell et al. (1984) using the Effelsberg 100 m dish and the 4.8 GHz receiver (thus $2.4'$ resolution) with two 40 MHz bandwidths. No ambiguous detections were reported in 9 sources which have redshifts ranging from $z = (0.0048 - 2.365)$. The 3σ upper limit to the peak line flux densities are $5 - 20$ mJy with a channel width of 20 km s^{-1}. Searches did not continue, likely because of the limited hardware at the time — inadequate resolution at low-frequencies (<1 GHz and especially below 500 MHz) and narrow bandwidth receivers.

1.4 Telescopes

The LOw Frequency ARray (LOFAR) and the Atacama Large Millimeter/submillimeter Array (ALMA) are two aperture synthesis telescopes which recently began operation, accessing portions of the electromagnetic spectrum with the sensitivity and resolution not previously capable. In this section, we briefly introduce these two telescopes. We discuss the new capabilities to observe radio recombination lines from these telescopes, thanks to more than 800 lines (per each element) populating frequencies below 950 GHz (see Figure 1.9). We also briefly introduce the advantages to studying thermal radio emission that these telescopes bring.

1.4.1 LOFAR

The LOw Frequency ARray (LOFAR; van Haarlem et al. 2013) is a new generation, digitally-intensive radio interferometer operating in the frequency range 10 – 250

Figure 1.10: *Top:* The yellow-filled circles mark the locations of LOFAR stations throughout Europe. LOFAR stations are centralized in the Netherlands and spread throughout nine European countries – the Netherlands, Germany, Poland, United Kingdom, France, Sweden, Ireland, Latvia, and Italy. *Bottom:* A zoom-in on LOFAR stations. The LOFAR "superterp" is the central circular patch of land containing multiple LOFAR stations that affords excellent short-spacing u,v coverage. Portions of three additional core stations are also seen in the peripheries of the image. High Band Antenna (HBA) stations are identified by consecutive segments of dark tiles. In the remote and international stations of LOFAR, a station consists of one consecutive grouping of HBA tiles; however, for the core stations (and as shown in this image), an HBA station is made up of two "ears" or two groupings of HBA tiles. Low Band Antenna (LBA) stations are shown in the image as dense groups of individual antenna elements interspersed within the HBA stations.

MHz. The array consists of a large number of simple, inexpensive dipole antennas that are not steerable, but fixed in place. Two types of antennas provide coverage in two bands: the low band antennas (LBA) at 10 – 90 MHz and high band antennas (HBA) at 110 – 250 MHz. LOFAR is the first telescope of its kind in the Northern Hemisphere and will uniquely remain so for the foreseeable future.

The heart of LOFAR is located in the Netherlands — as the telescope was designed and built by the Netherlands Institute for Radio Astronomy (ASTRON) — with stations currently spread across nine European countries (see Figure 1.10). The array is configured with 24 core stations (CS) within a radius of 2 km, and 6 of those stations reside on raised land 320 m diameter referred to as the "superterp" (see Figure 1.10). The 14 remaining stations of the full Dutch array, referred to as the remote stations (RS), extend out to a radius of 90 km. The international stations of LOFAR have baselines of more than 1800 km.

LOFAR is an extremely flexible telescope, offering multiple observing modes (beam-formed, interferometric) and vast ranges of spectral, timing and spatial resolutions. Through observing the low-frequency sky with unprecedented resolution, depth, and fidelity, the key science areas where LOFAR is making an impact include: H I 21 cm observations of the epoch of reionization; northern sky continuum surveys of AGN, star-forming galaxies, and galaxy clusters; pulsars; ultra high energy cosmic rays; (Galactic) ISM studies of supernovae, low-density ionized gas, and RRLs; transient (stellar) sources; magnetism; and, solar physics and space weather.

In particular, LOFAR is allowing for the exploration of RRLs in the Galaxy and in extragalactic sources by meeting the following technical requirements: 1) sensitivity to frequencies (principal quantum numbers) where stimulated (enhanced) emission is prevalent, 2) large fractional bandwidths that observe 50 to 100s of RRLs simultaneously, enabling changes in line properties to be identified and/or more sensitive searches through line stacking, 3) adequate and flexible spectral resolutions for Galactic ($\sim$1 km s^{-1}) and extragalactic ($\sim$10 km s^{-1}) observations, and 4) spatial coverage and resolution to large clouds while also resolving the $\lesssim$1–100 pc emitting regions. These requirements have inhibited wide-spread, in-depth studies of low-frequency RRLs in the past, largely due to the low spatial resolutions and the narrow bandwidths of traditional low-frequency instruments — owing to the difficulty of calibrating low frequency observations affected by the ionosphere. Using LOFAR, the unprecedented science of RRLs is being demonstrated (Asgekar et al. 2013; Oonk et al. 2014; Morabito et al. 2014; Oonk et al. 2017; Salas et al. 2017, 2018, 2019; Emig et al. 2019, 2020a).

Continuum studies of thermal radio emission have uncovered and characterized phases of warm, low-density ionized gas in the past (e.g., Westerhout 1958; Hoyle & Ellis 1963). Low-frequency observations provide a crucial probe of the turnover frequency of low-(electron) density ionized gas where the free-free emission becomes optically thick. LOFAR provides the unique resolution to resolve differences in cloud properties in order to investigate influences of energy and momentum injected by stellar feedback in warm ionized gas, the transport and leakage of ionization radiation, and turbulence of ionized gas.

Figure 1.11: Antennas of the Atacama Large Millimeter/submillimeter Array (ALMA). Fifty of these antennas have 12 m diameters and make up the "12 m array", used for sensitive, high-resolution imaging. Green circles in this image mark four additional 12 m antennas that are used for total power observation (the "TP array"), and interior to those are twelve closely spaced 7 m antennas (the "7 m array"). The 7 m and TP arrays enhance wide-field imaging of extended structures and are collectively referred to as the Atacama Compact Array.

1.4.2 ALMA

The Atacama Large Millimeter/submillimeter Array (ALMA)[6,7] currently operates in the frequency range 84 – 950 GHz (3.6 – 0.32 mm). The array is located on the Chajnantor plain of the Chilean Andes (lat.= $-23.02917°$, long.= $-67.754649°$), a region with exceptionally dry and clear sky conditions. The array consists of 66 antennas that can be positioned in different configurations over 192 antenna foundations with baselines as large as 16 km. A portion of the antennas are shown in Figure 1.11. Since first light in 2010 and becoming fully operational in 2013, ALMA has had a revolutionary impact on many fields of astrophysical research, notably star formation – from resolving individual stellar systems to galaxy scale properties and at the highest redshifts. The high-sensitivity and high-resolution capabilities have enabled the plethora of atomic and molecular species in this wavelength range to be explored together with dust and free-free continua.

In particular, ALMA affords a breakthrough in studying the most massive (and densest) forming stellar clusters. Direct optical and even near-infrared observations of forming clusters are complicated by large amounts of extinction present in the recently formed, largely embedded regions. Analyses of long wavelength free-free emission and hydrogen recombination lines of star clusters offer an alternative, extinction-

[6] `https://www.almaobservatory.org`

[7] A comprehensive description of the ALMA observatory and its components can be found in the ALMA Technical Handbook: `https://almascience.org/documents-and-tools/cycle8/alma-technical-handbook`

free probe of the ionizing gas (and therefore ionizing radiation and stellar mass) surrounding young star clusters (Condon 1992; Roelfsema & Goss 1992; Murphy et al. 2018). However, achieving a spatial resolution matched to the size of young clusters $\mathcal{O}(1$ pc) (Ryon et al. 2017) in galaxies at the necessary frequencies and sensitivities has only recently become possible thanks to the ALMA. Together with molecular species and dust continuum observations, a comprehensive picture can be assessed towards understanding their formation, identifying the dominant feedback processes at each stage of cluster evolution, determining which clusters survive as gravitationally bound objects, and linking all of these processes to the galactic environment (Turner et al. 2017; Oey et al. 2017; Leroy et al. 2018; Rico-Villas et al. 2020; Krieger et al. 2020; Emig et al. 2020b; Levy et al. 2021; Villas et al. 2020).

On larger (unresolved) scales, observations of RRLs with ALMA are providing tracers of total star formation rates and characterizing physical properties of ionized gas. At low ALMA frequencies affects due to stimulated emission which may be present at traditionally employed radio frequencies, 1–10 GHz, do not complicate line observations and the continuum is less contaminated by synchrotron emission. Sources including NGC 253 (Bendo et al. 2015, Mills et al., in preparation), NGC 3256 (Michiyama et al. 2020), NGC 4945 (Bendo et al. 2016; Emig et al. 2020b), and NGC 5253 (Bendo et al. 2017) have been investigated through RRLs.

1.5 In this thesis

This thesis is focused on investigating the interstellar medium and the influence of massive star formation on global (galactic) scales to sub-cloud pc scales. We focus on the influence of massive stars through the warm gas which they ionize in two different environments: a massive excavated region of relatively low surface mass density star formation (Chapter 2), and a massive clustered mode of efficient star formation (Chapter 3). We explore new methods of investigating the interstellar medium through RRL observations, resulting in the first detection at cosmological distances (Chapter 4 and 5).

In **Chapter 2**, we investigate the influence of massive stars on their environment by tracing low-density photoionized in the Cygnus X region. The Cygnus X region ($d \sim 1.5$ kpc) is one of the most massive star forming complexes in our Galaxy, in which the Cyg OB2 association (age of 3-5 Myr and stellar mass 2×10^4 $M_\odot$) has a dominant influence. We observe the Cygnus X region at 142 MHz using the LOFAR and correct for missing short-spacing information during image deconvolution. Together with data from the Canadian Galactic Plane Survey, we investigate the morphology, distribution, and physical conditions of low-density ionized gas in a $4^\circ \times 4^\circ$ ($\sim$100 pc $\times$ 100 pc) region at a resolution of $2'$ (0.9 pc). The radio emission in the region analyzed is almost entirely thermal (free-free) at 142 MHz, with emission measures (EM) of $10^3 < EM$ [pc cm^{-6}] $< 10^6$. As filamentary structure is a prominent feature of the emission, we use `DisPerSE` and `FilChaP` to characterize their radial profile distributions. The distribution of profiles has a characteristic filament width of 3.6 pc and a power-law distribution ($\beta = -1.8 \pm 0.1$) in peak EM down to our completeness limit of 4200 pc cm^{-6}. The electron densities in the filaments range

between $10 \lesssim n_e$ [cm^{-3}] $\lesssim 400$ with a median value of 38 cm^{-3}, remarkably similar to [N II] surveys of ionized gas. Cyg OB2 may ionize up to two-thirds of the total ionized gas and the ionized gas in filaments. Indeed we find the majority of filaments are likely photoevaporating surfaces flowing into a surrounding diffuse ($\sim$5 cm^{-3}) medium. However, stellar winds of Cyg OB2 may create a minority of the ionized filaments as dissipated turbulence or shock heated ionized gas. We discuss Cygnus X as a source of extended low-density (ELD) ionized gas, replenished over the lifetime of Cyg OB2 by ten of the typical filaments we characterize.

In **Chapter 3**, we turn to NGC 4945, a nearby (3.8 Mpc) galaxy hosting a nuclear starburst and Seyfert Type 2 AGN. We use the Atacama Large Millimeter/submillimeter Array (ALMA) to image the 93 GHz (3.2 mm) free-free continuum and hydrogen recombination line emission (H40α and H42α) at 2.2 pc (0.12″) resolution. Our observations reveal 27 bright, compact sources with FWHM sizes of 1.4–4.0 pc, which we identify as candidate super star clusters. Recombination line emission, tracing the ionizing photon rate of the candidate clusters, is detected in 15 sources, 6 of which have a significant synchrotron component to the 93 GHz continuum. Adopting an age of $\sim$5 Myr, the stellar masses implied by the ionizing photon luminosities are $\log_{10}(M_\star/\mathrm{M}_\odot) \approx 4.7$–6.1. We fit a slope to the cluster mass distribution and find $\beta = -1.8 \pm 0.4$. The gas masses associated with these clusters, derived from the dust continuum at 350 GHz, are typically an order of magnitude lower than the stellar mass. These candidate clusters appear to have already converted a large fraction of their dense natal material into stars and, given their small free-fall times of $\sim$0.05 Myr, are surviving an early volatile phase. We identify a point-like source in 93 GHz continuum emission which is presumed to be the AGN. We do not detect recombination line emission from the AGN and place an upper limit on the ionizing photons which leak into the starburst region of $Q_0 < 10^{52}$ s^{-1}.

Chapter 4 reports the first detection of RRLs at cosmological distances with observations of the radio quasar 3C 190 ($z = 1.1946$. 3C 190 was observed with the HBA of the LOFAR and processed using newly developed techniques for spectral analysis. We report the detection of RRLs in the frequency range 112 MHz – 163 MHz in the spectrum of 3C 190. Stacking 13 α-transitions with principal quantum numbers $\mathsf{n} = 266 - 301$, a peak 6σ feature of optical depth $\tau_{\mathrm{peak}} = (1.0 \pm 0.2) \times 10^{-3}$ and FWHM $= 31.2 \pm 8.3$ km s^{-1} was found at $z = 1.124$. This corresponds to a velocity offset of -9965 km s^{-1} with respect to the systemic redshift of 3C 190. We consider three interpretations of the origin of the RRL emission: an intervening dwarf-like galaxy, an active galactic nucleus-driven (AGN) outflow, and the inter-galactic medium. We argue that the recombination lines most likely originate in a dwarf-like galaxy ($M \sim 10^9$ M$_\odot$) along the line of sight, although we cannot rule out an AGN-driven outflow. We do find the RRLs to be inconsistent with an inter-galactic medium origin. With this detection, we have opened up a new way to study the physical properties of cool, diffuse gas out to cosmological distances.

In **Chapter 5**, we lay out the technical methods of searching for faint RRLs, both in processing and preparing new spectral line observations with LOFAR and how a spectrum can be searched in redshift space for RRLs. RRLs have been largely unexplored outside of our Galaxy. Next-generation low-frequency interferometers such as LOFAR, MWA, and the future SKA will, with unprecedented sensitivity,

resolution, and large fractional bandwidths, enable the exploration of the extragalactic RRL universe. We describe methods used to (1) process LOFAR high band antenna (HBA) observations for RRL analysis, and (2) search spectra for RRLs blindly in redshift space. We observed the radio quasar 3C 190 ($z \approx 1.2$) with the LOFAR HBA. In reducing these data for spectroscopic analysis, we placed special emphasis on bandpass calibration. We devised cross-correlation techniques that utilize the unique frequency spacing between RRLs to significantly identify RRLs in a low-frequency spectrum. We demonstrate the utility of this method by applying it to existing low-frequency spectra of Cassiopeia A and M 82, and to the new observations of 3C 190. Radio recombination lines have been detected in the foreground of 3C 190 at $z = 1.12355$ (assuming a carbon origin) owing to the first detection of RRLs outside of the local universe (first reported in Emig et al. 2019). Toward the Galactic supernova remnant Cassiopeia A, we uncover three new detections: (1) stimulated Cϵ transitions ($\Delta \mathsf{n} = 5$) for the first time at low radio frequencies, (2) Hα transitions at 64 MHz with a full width at half-maximum of 3.1 km s^{-1}, the most narrow and one of the lowest frequency detections of hydrogen to date, and (3) Cα at $v_{\rm LSR} \approx 0$ km s^{-1} in the frequency range 55–78 MHz for the first time. Additionally, we recover Cα, Cβ, Cγ, and Cδ from the -47 km s^{-1} and -38 km s^{-1} components. In the nearby starburst galaxy M 82, we do not find a significant feature. With previously used techniques, we reproduce the previously reported line properties. RRLs have been blindly searched and successfully identified in Galactic (to high-order transitions) and extragalactic (to high redshift) observations with our spectral searching method. Our current searches for RRLs in LOFAR observations are limited to narrow (< 100 km s^{-1}) features, owing to the relatively small number of channels available for continuum estimation. Future strategies making use of a wider band (covering multiple LOFAR subbands) or designs with larger contiguous frequency chunks would aid calibration to deeper sensitivities and broader features.

1.6 Future Outlook

A major take away from this thesis is that the physical conditions in the diffuse ISM can be explored through observations of RRLs in and/or against radio bright galaxies out to high redshift. This opens the door to investigations in large populations of external galaxies of cold, diffuse gas and/or of diffuse ionized gas using existing facilities like the LOFAR, VLA, GMRT, WSRT, MeerKAT, and ASKAP. In Figure 1.12, we plot the (brightest) galaxies at 178 MHz observed with the 3C(R) catalog that will enable the deepest searches at low frequencies. In the immediate future, it is possible to follow up the 3C 190 analysis with RRL observations at lower frequencies (i.e., with LOFAR's LBA) and higher frequencies (P-band, 300–500 MHz) to confirm the detection and constrain physical properties of the gas. It is also clear that M 82 needs to be re-observed to significantly deeper sensitivities (50 hours) and at higher spectral resolution using LOFAR.

In the future, the Square Kilometer Array (SKA)[8] can revolutionize (extragalactic) radio recombination lines (Peters et al. 2011; Oonk et al. 2015; Manti et al. 2016).

[8] `https://www.skatelescope.org`

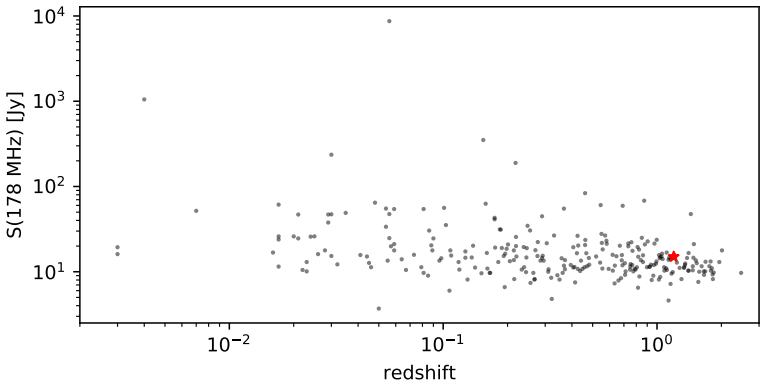

Figure 1.12: 278 (of the 298) sources of the revised third Cambridge catalog (3CR) — which includes radio sources in the Northern Hemisphere with flux densities of $S(178\ \mathrm{MHz}) \geq 9$ Jy — with known redshifts as reported in Spinrad et al. (1985). The 3CR catalog consists of the brightest sources in the low-frequency sky that will enable the deepest searches of RRLs from the diffuse ISM with current telescopes. In the future, SKA will enable deep searches in more than 10 000 galaxies. 3C 190, the source from which an RRL detection was reported by Emig et al. (2019), is shown with a red star at $z = 1.2$.

It will be a radio interferometer with the world's largest collecting area and fastest survey speed (expected operational for the community by 2028). With two different arrays, the SKA1-low ($50 - 350$ MHz) located in Australia will have a maximum baseline of 65 km ($6''$ resolution at 200 MHz) and deliver a sensitivity approximately eight times deeper than that of LOFAR, and the SKA1-mid ($0.35 - 15$ GHz) will be operated from South Africa (current MeerKAT site) with a maximum baseline of 150 km ($0.1''$ resolution at 5 GHz) with a sensitivity about six times deeper than the VLA. Additional updates (planned as SKA2) would add another order of magnitude in sensitivity. With LOFAR, clouds with column densities of 3×10^{20} cm^{-2} on scales of $>10'$ can be probed. With SKA, we can detect clouds with $N(\mathrm{H}) > 5 \times 10^{19}$ cm^{-2} on scales $>3'$. Or to put this in a different perspective, with LOFAR we can study absorption towards the sources in the 3C catalog while with SKA, all sources in the 7C catalog come into play (Oonk et al. 2015).

NGC 4945 and NGC 253 are the first galaxies in which super star clusters are being uncovered and resolved with ALMA (Leroy et al. 2018; Emig et al. 2020b) and being directly probed with RRLs using ALMA (Emig et al. 2020b, Mills et al., in preparation). ALMA has great potential to uncover many more SSCs and constrain (i) the fraction of star formation that occurs in this clustered mode, (ii) the timescales for the onset and completion of cluster formation, (iii) the net efficiency (fraction of gas converted to stars) and efficiency per free-fall time, (iv) the dominant feedback and support mechanisms for protoclusters as a function of time (Levy et al. 2021), and (v) the mechanism that ends cluster formation. A significant fraction of galaxies with existing RRL detections at lower resolution or lower frequency seem to harbor SSCs; populations of young SSCs likely remain to be uncover from the currently remaining extragalactic sources with RRL detections.