Advanced receivers for submillimeter and far infrared astronomy
Kooi, Jacob Willem

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2008

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 18-08-2019
Chapter 1

Introduction

1.1 Submillimeter and far-infrared astronomy

Modern astronomy started when Galileo Galilei in 1610 pointed a small telescope towards the heavens. It nearly cost him his life and changed the world. Since then we have become aware that most of the light in the universe resides in the infrared and submillimeter region of the electromagnetic spectrum (1 \(\mu\)m – 1 mm). One of the reasons for this is that at large, the universe is a cold and forbidding place. Its low temperature causes the prevalent dust and molecular gas in galaxies, including our own Milky Way, to exhibit an emission peak in the far-infrared and submillimeter wavelength regimes.

The other reason has to do with the expansion of the universe. In the late 1920s Edwin Hubble (California Institute of Technology) made the important cosmological discovery that spectral lines of galaxies are shifted toward the longer wavelength end (red part) of the electro-magnetic spectrum in an amount proportional to their distances. The explanation of this apparent redshift lies with the Doppler effect, which suggests that galaxies move away from each other with velocities proportional to their distance. This has since become known as the “The Hubble law” and implies that the universe is expanding on the whole. Then in 1965 another great cosmological discovery was made by Arno Penzias and Robert Wilson (Nobel Prize 1978), at Bell labs at the time. They discovered that there is a universal microwave radiation, with a spectrum that corresponds to a ~ 3 K blackbody. This has since been confirmed in exquisite detail by COBE [1], NASA’s cosmic background explorer mission, designed to measure the diffuse infrared and microwave radiation of the early universe. It was found that the cosmic microwave background (CMB) spectrum is that of a nearly perfect blackbody with a temperature of 2.725 ± 0.002 K (PI: John Mather, Nobel Prize 2006). The success of COBE has been followed up by the Wilkinson Microwave Anisotropy Probe (WMAP) [2], another NASA explorer mission, and soon also with the launch of the European Space Agency (ESA) satellite Planck [3]. Planck has more than fifty times the angular resolution of COBE, and at least ten times the sensitivity. It will, like WMAP, observe the anisotropies of the cosmic microwave background and
help answer fundamental questions such as: Will the universe continue its expansion forever, what is the nature of the so-called “dark matter” that appears to count for more than 90% of the total amount of matter in the universe, and furthermore what is “dark energy” in relation to the accelerated expansion of the universe at the earliest epoch of time?

And so we find the submillimeter and far-infrared an interesting wavelength regime (100 µm to 1 mm), one where continuum and line radiation from relatively cool diffuse material, such as interstellar and circumstellar dust and gas, is most abundant (Fig. 1.1). Black bodies with temperatures between 5- and 50 K peak here, and gasses with temperatures between 10 and $\sim 200$ K emit their brightest molecular and atomic emission lines.

Broadband thermal radiation from dust grains is the most common continuum emission process in this wavelength range. It is also here that luminosities of a wide variety of astronomical sources obscured by interstellar or circumstellar material may be derived. This is because ultraviolet, visible, and near-infrared photons from stars are absorbed by the dust grains in the interstellar medium and re-emitted as a modified graybody spectrum in the far-infrared. And due to a very high optical extinction,
only a small amount of dust suffices to absorb the short-wavelength UV radiation. In this way galaxies, or even quasars, emit a significant fraction of their radiation in the far-infrared. The spectral energy density (SED) of the dust emission in the submillimeter and far-infrared thus provides quantitative information of the temperature distribution of dust particles, their composition, size and spatial distribution. And because submillimeter measurements are little effected by dust extinction, studies of the physical condition and energy balances of nuclear regions in dust enshrouded galaxies are greatly facilitated. As a final note, the submillimeter continuum optical depth provides a direct measure of large dust concentrations, such as are present in protostars, circumstellar disks, or in the large amounts of interstellar gas in galactic nuclei.

Knowledge of the submillimeter and far-infrared spectrum is therefore essential in our understanding of the many complex physical and chemical processes that take place in astrophysical objects. For example, the cooling of giant molecular clouds (GMCs) in star forming regions. It is in these cold and dense molecular clouds where young stars are born.

A striking aspect of submillimeter astronomy is that dust emission, molecular lines, and for example the atomic fine-structure line of C\(^+\), which dominates the gas cooling, are all red shifted to the submillimeter band for distant galaxies. At a redshift of 2–4, galaxy formation in the early universe was most active. Thanks to the expanding universe therefore, lines and dust are seen to shift into the submillimeter band and will be detectable on the ground by large telescopes such as the CSO [4], APEX [5], and in the near future ALMA [6]. Even today Spitzer is surveying large areas of the sky for these early universe objects. Spectroscopic follow-up will be required to determine the physics of galaxy evolution.

It is thus seen that the submillimeter and far-infrared wavelength regimes are critical for astronomy. They contain spectral and spatial information on the early stages of star formation in giant molecular gas clouds, both in our own and external (distant) galaxies, and the cosmic background.

### 1.1.1 Interstellar medium

The interstellar medium (ISM) is a complex environment that harbors giant molecular clouds out of which stars – and planetary systems – form. Since the creation of the universe the neutral HI dominated medium is enriched by newly synthesized elements from dying stars that have run their course. This complex interplay between stars and the ISM drives the evolution and thus the observational characteristics of our local galaxy (Milky Way) and external galaxies all the way back to the earliest protogalaxies (Fig. 1.2).

Schematically the ISM can be represented by Fig. 1.3 (Walker et al., 2007). The HII region is a heavily ionized and physically hot (∼ 10,000 K) region in close proximity to young stellar objects (YSO). Profuse amounts of high energy far UV photons dominate these regions ionizing everything in their path. This includes atomic nitrogen, that with an ionization potential of 14.52 eV can only exist in the UV dominated HII regions in ionized N\(^+\) form. Also present in the HII region is C\(^+\) with an ionization potential of 11.26 eV.
Figure 1.2: Optical image of the Hubble deep field. Shown is a bewildering assortment of galaxies, some of which are close to the beginning of the universe $\sim 13.73 \pm 0.12$ billion years ago. Dying stars in each of these galaxies enrich the local interstellar medium, thereby effecting the evolutionary cycle of galaxy formation including that of our own Milky Way. Photo credit: R. Williams (STScI).

The boundary region between the cooler molecular clouds and HII region is known as the photodissociated region (PDR). Here the overall UV photon energy still dominates the heat budget ($6 \text{ eV} < h\nu < 13.6 \text{ eV}$) and modifies the chemical balance of the molecular cloud. Many of the molecular species (such as CO) are photodissociated and C$^{+}$ and C$^{0}$ become highly abundant at the expense of CO in this case. As depicted in Fig. 1.3, the interstellar medium has a variety of phases, from very diffuse regions of low hydrogen column density, to dense diffuse warm clouds, as well as photodissociated regions. Large scale shocks produced by supernovae or stars injecting

Figure 1.3: Schematic representation of the interstellar medium. The various regions with molecular and atomic content are indicated. The dense molecular cloud cools by means of the emission of photons with energies that correspond to the rotational line transitions of the molecules (C. K. Walker \textit{et al.}, 2007).
matter (outflows) are not uncommon.

Again referring to Fig. 1.3, the neutral HI gas in the interstellar medium is driven via a thermal balance into two phases; the diffuse warm neutral medium (WNM) of \( \sim 8000 \) K, and the diffuse cold neutral medium (CNM) at \( \sim 70 \) K. Turbulence and shocks will naturally complicate the thermal equilibrium between the two regions. \( \text{C}^+ \) and neutral atomic hydrogen (HI can be observed with the forbidden hyperfine 21 cm, 1.42040575 GHz, transition) are both present in this neutral medium. In the presence of UV radiation, from say a nearby star, warm diffuse neutral clouds can be destroyed. The resulting ionized region contains in this case hydrogen, carbon and nitrogen ions. Thus we find that ionized carbon is high distributed among the different ISM regions, with \( \text{N}^+ \) isolated to the more heavily (warm) ionized regions. From this model we find that \( \text{N}^+ \) emission provides a measure of the ionizing radiation from young massive stars in the ISM. \( \text{N}^+ \) also has the advantage that it is unaffected by dust extinction. As such ionized nitrogen appears to be a good probe to identify star forming regions. The ratio of \( \text{C}^+/\text{N}^+ \) measures the rate of cloud destruction/formation.

Because molecular hydrogen (\( \text{H}_2 \)) does not have accessible emission line spectra in the submillimeter and far-infrared (the nearest transition lines are in the UV) tracer elements such as \( \text{C}^0 \), \( \text{CO} \), and \( \text{C}^+ \) are commonly used to study molecular clouds. Atomic carbon (\( \text{C}^0 \)) can be found in all types of neutral clouds, from diffuse (Jenkins & Shaya 1979 [7]) to dense molecular gas (Phillips & Huggins 1981 [8]). \( \text{C}^+ \) has been seen extensively by COBE throughout the Milky Way and makes a significant contribution to the gas cooling, being so widely distributed (Bennett et al. 1994 [9]). Atomic carbon is also found to be a good tracer of molecular gas in molecular clouds. For example, their appears to be a linear correlation between the strength of the \( \text{C}^0 \) \( ^3\text{P}_1 \rightarrow ^3\text{P}_0 \) line at 492 GHz and the \( ^{13}\text{CO}_{J=2\rightarrow1} \) line at 220 GHz (Keene et al. 1997 [10]). Higher \( J \) \( \text{CO} \) transitions such as \( ^{12}\text{CO}_{J=7\rightarrow6} \) (807 GHz) trace the warm energetic gas component, whereas higher transitions of HCN and HCO\(^+\) trace the density of the gas that is being warmed by star-formation in clouds and PDRs. Strong mid- and high \( J \) rotational line emission is also a signature of shocked gas in the molecular outflows of newly formed stars.

Although emission lines of \( \text{C}^0 \) and \( \text{CO} \) arise in principle from different physical regions in an ISM cloud, in actual observations of galactic clouds it is found that CI and CO emissions, on average, seem to come from the same physical region (Phillips et al. 2002). This appears to be due to the clumpy or fractal nature of the clouds which allows UV radiation to irradiate regions even deep into the cloud. For external galaxies where individual clouds are unresolved, the different carbon species, including \( \text{C}^+ \), will be observationally coexistent. In these way line ratios may be compared. It is essential therefore to have a good understanding of nearby galaxies in order to provide templates for more distant unresolved objects.

Aside from ionized carbon, water is also a very important molecule in interstellar chemistry (Tielens 2005) [11]. It can be a dominant reservoir of elemental oxygen in the gas phase, and because of its many levels, water is also an important coolant that can dominate the heat balance of a gas. To study and understand the role of water in the ISM, the high resolution Heterodyne Instrument (HIFI) [12] on Herschel [13], will have as its primary objective (it cannot be done from the ground) the study of a
wide range of H$_2$O transitions. A second, nearly equally important objective of HIFI, is to study the role of C$^+$ as the main cooling line in the interstellar medium.

Ionized carbon has a ground-state $^2P_{3/2} \rightarrow ^2P_{1/2}$ CII atomic fine-structure transition at 158 $\mu$m (1.901 THz). The CI atomic carbon ($C^0$) fine-structure transition falls at 370 $\mu$m ($^3P_2 \rightarrow ^3P_1 = 809$ GHz) and 609 $\mu$m ($^3P_1 \rightarrow ^3P_0 = 492$ GHz). N$^+$ has two NII ground-state atomic fine-structure transition lines: $^3P_1 \rightarrow ^3P_0$ at 205.178 $\mu$m (1.470 THz) and $^3P_2 \rightarrow ^3P_1$ at 121.898 $\mu$m (2.459 THz). Each of these lines is very narrow (low pressure, low velocity) and high resolution spectroscopic observations are needed to spectrally resolve them. Much of this thesis is devoted to developing the technology that enables these observations.

1.1.2 Star formation

So far we have discussed the various phases of the interstellar medium, but have not yet touched in detail upon the subject of star formation. Stars are found to form in the dense and dusty molecular clouds of the interstellar medium (Figs. 1.1, 1.3). One of the physical properties of the ISM is in the way gas cooling processes allow clouds to collapse to form stars. In particular, cooling happens through line emission from rotationally excited H$_2$O, diatomic hydride molecules (H$_2$S, CH, HCl, CH+, NH, ...) and, as we have seen, CI and CII.

1.1.2.1 Molecular clouds

The atomic and isotopic abundances of the ISM provide information on the degree of star-formation activity a given area has suffered. A critical factor in the star-formation process is the cooling mechanism, since a cloud cannot collapse all the way to form a star unless it can purge itself of the heat from its compression under gravity. Such cooling is provided by a variety of molecular transitions in the submillimeter band, but mostly by CO, H$_2$O, light hydride molecular species, CI, and CII. The densities of interest in star formation, where molecules dominate, span the range from $10^2$ cm$^{-3}$ to $> 10^{12}$ cm$^{-3}$. Shocked regions from supernovae or outflows from young stellar objects (YSO) are likely to experience violent temperature and density increases, possible causing star formation.

The primary manifestation of star formation in the ISM is through molecular outflow. Many deeply embedded sources were first detected this way, and are known as young stellar objects (YSOs). They are surrounded by small 100 AU disk like regions. Complex line profiles may therefore be expected in shocked regions, showing evidence of turbulence, high velocity winds, and pressure broadening in the observed emission lines. To understand the complex interactions in the different regions of the ISM, to overcome spectral blending, line confusion, and to resolve velocities of the various clouds along the line of site, high resolution heterodyne spectroscopy is therefore required.

Since the temperatures in the cold dense molecular gas ranges from approximately 10 K in the cooler regions to 100 – 200 K in the hotter regions (Phillips & Keene 1992 [14]), the corresponding emission has a frequency range ($h\nu \sim k_B T$) of 0.2 – 4 THz. This is also the same range many molecular species have their rotation and
1.1. SUBMILLIMETER AND FAR-INFRARED ASTRONOMY

Figure 1.4: Schematic presentation of a 30 K interstellar molecular cloud. The spectrum includes dust continuum, molecular rotation- and atomic fine-structure line emissions (Phillips et al. [14]). The suite of lines forms an excellent tool to probe the molecular and ionized gas conditions after young stars disrupt their environment of birth.

atomic fine-structure transitions. In Fig. 1.4 we show a representation of a cloud with 30 K dust and gas temperatures. A blackbody curve bounds the emission with a Wien peak near the 1.9 THz ionized carbon transition. The dust in the cloud is optically thick and has a continuum emission on the blackbody curve. In the millimeter band (shaded areas) heavy molecules with carbon chemistry prevail. In the submillimeter band light molecules with low angular momentum and high rotation speeds dominate. In general the higher-J energy levels probe physically conditions of higher temperature, e.g. $^{12}\text{CO}_{J=7\rightarrow6}$ (807 GHz) vs. $^{12}\text{CO}_{J=2\rightarrow1}$ (230 GHz). CI and CO are observable in both warm ($T>70$ K) and cold ($T<20$ K) neutral medium clouds, but CII is only observable in warm clouds, and therefore is not a good tracer in a cold cloud.

The many rotational molecular, atomic, and ionic fine-structure lines thus form an excellent probe of the physical and chemical conditions in the molecular cloud, of high velocity stellar winds, protostellar and planet forming regions, galactic nuclei, and galaxy evolution. Analysis of the obtained data provides information on the pressure, elemental and isotopic abundances, molecular inventory of photodissociated...
regions, and a general understanding of the interstellar medium.

1.1.2.2 Spectral line surveys

As can be seen from Fig. 1.5, a line survey of the Orion (OMC-1) molecular cloud shows a forest of lines incredibly rich in chemical detail. Among the heavy molecules

Figure 1.5: Top) The Caltech Submillimeter Observatory (CSO) 325 – 360 GHz survey by Schilke et al. (1997); Middle) CSO 600 – 720 GHz survey by Schilke et al. (2001); Bottom) CSO 794 – 840 GHz line survey by Mehringer et al. (2001). Spectra from E. van Dishoeck [15].
are also some simple light molecules. Surveys such as those shown in Fig. 1.5 thus provide an unbiased view of the molecular inventory of a wide range of objects. Moreover, the large number of lines of individual molecules presented in these spectra allow for a detailed study of the physical conditions in the emitting gas. For example, a spectral survey can tell us when lines are saturated. The line ratios can also provide important information on the temperature and local density, whereas the molecular abundance provides a measure of the evolutionary state of the object. In addition, the cooling rate of the gas can be estimated (CO, CI, H$_2$O). Surveys also address major questions in star-formation such as: How does the chemical composition depend on the mass and luminosity of an object (Dishoeck 2000 [15]). Measuring the molecular inventory in UV radiation fields, PDR shock regions, and molecular clouds will also aid our understanding of how stars and galaxies form, and why we are here.

The whole question of the gas phase chemistry of molecular clouds thus depends on a good understanding of the physical conditions, which may be revealed by detailed studies of atomic species. Ground based survey data has provided a complex, yet comprehensive picture of the physical and chemical evolution of star-forming regions. Heterodyne spectroscopy is particular useful in this regard because it has the spectral resolution necessary to resolve the kinematic components of the gas clouds.

### 1.1.3 Galaxies

Studies of the ISM in our own galaxy, and observations made with large aperture telescopes toward nearby galaxies, can be used as templates for more distant galaxies at z=0.5–3 when galaxy formation was in full swing. The different phases of the ISM are discernable with large aperture telescopes for nearby galaxies in “our” local group, such as M31 (the Andromeda galaxy), M83 (spinwheel galaxy in Hydra), the LMC, SMC and others. Herschel [16] (by virtue of no atmosphere and high spatial resolution, 11.2” beam at the 1.901 GHz CII transition) will be especially well suited to study important carbon and water cooling lines. Thanks to the high spectral resolution provided by heterodyne receivers, deconvolution of the observed line profiles can be applied to disentangle the different sources of emission. This technique provides discrimination between the narrow, cold phase lines and the broader (more turbulent) lines from warm ionized gas.

For a long time, detection of extragalactic molecules was restricted to a few molecular species. Even though ∼ 30 molecules have been discovered outside the Milky Way, no extragalactic molecular line surveys have been completed. Such line surveys can however provide a detailed (and necessary) comparison of the chemical composition, isotopic abundances, and excitation conditions between extragalactic and galactic sources (Güsten et al. [17]). Key extragalactic questions to be addressed may be summarized as: Understanding of the evolution of galaxies, the nature of the interstellar medium, and the process controlling the formation of stars in galaxies.

Being able to study a galaxy with a few beams across will provide a more global aspect of the interstellar medium. Namely, the different antenna beams trace different phases of the ISM (diffuse and molecular clouds, the warm PDRs), and characterize therefore the global properties of the ISM in galaxies. Submillimeter spectroscopy of nearby galaxies is likely to reveal regions of star formation or shock wave activity,
which should provide correlations among the various atomic and molecular species. The correlations obtained in this way tell us about the overall chemistry, and help test physical models for the interstellar medium.

With the Heterodyne Instrument (HIFI) [12] on board Herschel [16], observations of a large number of molecular and atomic transitions detectable in starburst- and the nuclei of nearby galaxies will for the first time enable a detailed analysis of the physical properties, and to some extent the structure of the star forming interstellar medium in these galaxies. In particular, the brightness of C$^+$ cooling lines will serve as an indicator of high mass star forming rate.

A major challenge for future observations of distant galaxies will be to observe strong emission lines such as CII and the so far unmentioned presence of atomic oxygen (OI), as they redshift into the submillimeter atmospheric windows. Atomic oxygen is also expected to be present in molecular clouds, since the ionization potential is quite high (13.62 eV). OI has two atomic fine-structure transitions in the far-infrared. They are: $^3P_0 \rightarrow ^3P_1$ at 147 $\mu$m (2.060 THz), and $^3P_1 \rightarrow ^3P_2$ at 63 $\mu$m (4.745 THz). To detect these lines at high-redshift, large telescopes and excellent SIS receivers, such as the continuous comparison or correlation receiver of Chap. 8, shall be required. It may in fact be the only method available to learn in detail about the chemical and dynamic processes which take place in high red-shifted galaxies.

### 1.1.4 Solar system

#### 1.1.4.1 Planets

The discussion should not end without a brief mention of solar system atmospheric chemistry, or planetary aeronomy, since this is also within the capability of heterodyne
1.1. SUBMILLIMETER AND FAR-INFRARED ASTRONOMY

Mars $^{10}O_{101}$ April 17 - May 3, 1999

Figure 1.7: Detection of the ground level $^{10}O_{101} - 1_{01}$ water line at 556.936 GHz on Mars, as observed by SWAS [18]. Compared are various water distributions (From Gurwell et al. 2000 [19].)

receivers. Because of telluric $H_2O$ and $O_2$ lines, ground based observations of solar system bodies are only possible in the submillimeter atmospheric windows. In the case of airborne platforms such as SOFIA [20] or high altitude ballooning, observations of planetary $H_2O$ remains impossible. Thus observations from space have a distinct advantage over ground based observations, albeit at a much larger expense.

Regardless of these issues, remote sensing in the millimeter and submillimeter is an important tool for studying the chemical composition of atmospheres [21]. Molecular lines are often narrow, though strong $H_2O$ lines can be as wide as 2 GHz, thanks to pressure broadening. Heterodyne spectroscopy provides the capability of measuring individual line shapes, and is therefore especially well suited for vertical profiling of atmospheres. An added advantage of the submillimeter regime is that it is relatively immune to dust extinction, as would be the case with infrared observations.

SWAS [18] and ISO [22] have in fact both detected water in planetary systems. ISO has detected $H_2O$ in the four giant planets, and on Saturn’s moon Titan. CO and $CO_2$ have also been detected by ISO in Jupiter, Saturn, and Neptune, but interestingly not on Uranus. SWAS with its high resolution Schottky based heterodyne receiver (Sec. 3.1) has observed the ground state $H_2O 1_{10} - 1_{01}$ line at 556.936 GHz in Jupiter and Saturn. $H_2O$ and CO are thought to be deposited in the upper atmospheres by comets, such as the well observed 1994 Schoemaker-Levy Jupiter impact. In the Martian atmosphere some of the identified molecular species, aside from of course $CO_2$ are: $CO$, $H_2O$, $O_2$, $O_3$, and $H_2$, the latter species observed with FUSE, the far-infrared spectroscopic explorer from NASA [23].

For the high resolution heterodyne instrument on the Herschel space observatory [12, 13] the primary science goals are to obtain a detailed characterization of $H_2O$ in the Martian atmosphere. This means measuring a variety of backbone water
transitions, HDO, and isotopic ratios of H$_2^{18}$O and H$_2^{17}$O. The latter two will establish the important hydrogen/deuterium ($H/D$) ratio with high precision. Chemically speaking, preferential escape of hydrogen relative to deuterium causes an enrichment of the atmospheric abundance of the latter over geological time. Measurement of atmospheric HDO and H$_2$O provides an estimate of this enrichment process, and may be compared to the 4.5 billion year atmospheric evolution of for instance the Earth ($H/D \sim 1000$), Venus ($H/D \sim 100$), and comets (next Section). Also anticipated with the Herschel space observatory are searches in the Martian atmosphere for OH, HO$_2$, H$_2$O$_2$. These species are however subject to seasonal variations.

The study of planetary atmospheres provides therefore important clues for understanding the origin, evolution, and formation processes of terrestrial and giant planets. It is clear from everyday experiences on Earth that planetary atmospheres undergo continuous processing through their lifetimes, for example: Volcanism on Io. Both transport (diffusion, convection, escape and winds) and physicochemical effects (photochemistry, electrochemistry, thermochemistry) act to change the state of planetary atmospheres. A high resolving power ($R > 10^6$, Chap. 2) enables the retrieval of the vertical distribution of the observed molecule, the thermal profile, and wind velocity [21]

Submillimeter Fourier-transform spectroscopy ($R \sim 10^4$) such as demonstrated from the CSO by Weinstein and Serabyn (1994) [24], the LWS Fabry-Perot instrument on ISO [25], and the Imaging Fourier Transform Spectrometer instrument (SPIRE) [26] on Herschel are also very important in the study of planetary atmospheres. These instruments provide a lower resolving power than heterodyne instruments, but are able to measure over a far greater instantaneous signal (RF) bandwidth. This is good for lines that are too broad to be detected with a high resolution heterodyne instrument such as troposphere absorption lines.

In the Earth atmosphere ozone (O$_3$) is a very important trace molecule [27] as it protects the environment from harmful UV-B radiation (for example, the phytoplankton population of Antarctica which affected the entire food chain – from krill to penguins and seals). It is well known that O$_3$ is a powerful oxidizing agent and that it is readily converted to O$_2$ via a photosynthesis reaction of UV light, Cl, CLO, NO and NO$_2$. The highly oxidative OH hydroxyl is the catalyst in this process, and has been found to play a central role in the stratospheric (10 – 50 km) ozone production process [28]. It is however, due to its low abundance (highly reactive, short life time), difficult to observe [29]. In the near future, it may now be possible to observe OH directly via the 1.835 THz, 2.510 THz, and/or 3.545 THz molecular emission line with an HEB heterodyne mixer (Sec. 4.3). For example, with balloon-borne observation from the Terahertz Limb Sounder (TELIS) [30]. As a final word, O$_3$ also plays an important role in the stratospheric heating/cooling balance. Depletion of ozone in the stratosphere is likely to affect the radiative heat balance in the entire atmosphere, with unforeseen consequences. Some other molecules (green house gasses) that affect the heat balance of the atmosphere, aside from CO$_2$, are methane (CH$_4$), and nitrous oxide (N$_2$O). Both have emission lines in the submillimeter, and are in detectable with the laid out heterodyne techniques of this thesis.
1.1.4.2 Comets

Having retained and preserved pristine material from the solar nebula at the moment of their accretion, comets contain unique clues to the history and evolution of the solar system. The study of comets follows the natural progression from interstellar matter to solar system bodies and their formations [34].

Water is the primary constituent of cometary ice, and evaporation drives the dynamics of cometary outgassing at small (< 4) heliocentric distances. Active comets release both gas (mainly water) and dust, with water playing an important role in the thermal balance of a cometary atmosphere as a cooling agent via the emission of its rotational lines. Recent detections of the \( \text{H}_2\text{O} \ \text{J} = 1_{10} \rightarrow 1_{01} \) ground state rotational transition line at 556.936 GHz by SWAS in Comet C/1999 [31], and in Comet C/2001 by Odin [33] (Fig. 1.8) have shown the possibility of cometary outgassing observations. The spectra of water in comets yields information on the amount of water evaporated from the nucleus, the kinetics of the cometary atmosphere, coma expansion, and provides reference levels for minor volatile abundances. Due to the low escape velocity of cometary gasses, a < 1 km s\(^{-1}\) frequency resolution will be required to resolve cometary emission lines. Thus high resolution heterodyne spectroscopy (\( R \geq 10^7 \)), see also Sec. 2.1) will provide the only means available for these kind of observations. As an interesting side note, the vastly improved sensitivity and frequency range of HIFI [12] over the uncooled Schottky mixers (Sec. 3.1) flown onboard SWAS [18], Odin [32], and the microwave radiometer instrument (MIRO) onboard the ROSETTA [35] probe to comet 67P/Churyumov-Gerasimenko (2014), will allow for the study of a number of rotational \( \text{H}_2\text{O} \) transition lines at a variety of comets. Other cometary lines that should be observable with HIFI are: HDO, NH\(_3\), H\(_2\)S, OH, CH\(_2\), NH, H\(_3^+\), and OH\(^+\).

Since the cometary gas is cold (10–100 K), the rotational transitions at the lowest energy states are most intense. Thus it is important to look primarily for the fundamental ground state rotational water transition line as was so successfully done with SWAS and Odin.
The deuterium abundance is a key parameter for studying the origin and early evolution of the solar system (previous Section). Simultaneous observations of H$_2$O and HDO determine the $H/D$ ratio in cometary water. For example, in comet Hyakutake and Hale-Bopp a ratio of $\sim 3300$ was obtained. This is useful for refinement of theoretical models describing the early solar nebula dynamics. These comets were long period, coming from the Oort cloud (believed to be approximately 1 light-year from the Sun). Measurements of the $H/D$ ratio of comets from the “nearby” Kuiper belt (a bit beyond Neptune at $\sim 55$ AU) will again help constrain solar nebula models.

Needless to say, due to telluric water vapor, most of these observations can only be done from space.

1.2 Coherent (heterodyne) detection

Atomic and molecular line astrophysics is demanding enhanced capabilities in terms of frequency coverage, spectral resolution, bandwidth, sensitivity, stability (quality of the data products), frequency agility, and ease of use. To help fulfill these demanding requirements, this thesis investigates advanced heterodyne receiver techniques in the submillimeter and terahertz frequency regimes.

1.2.1 Challenges

Coherent detection, the research topic of this thesis, facilitates high resolution spectroscopy by down converting the phase and amplitude of the incoming signal to a much lower intermediate frequency. Hence the terminology “coherent detection”. To accomplish this task, a local oscillator signal is required. One of the difficulties of far-infrared heterodyne spectroscopy has been the limited availability of local oscillator (LO) signal power. Recent advances in (waveguide based) local oscillator multiplier technology [36, 37, 38] have opened the field of molecular spectroscopy to $\sim 2$ THz. Above this frequency Quantum Cascade Lasers (QCL) [39] have in recent years shown much promise. In fact, the QCL can be expected to play a significant role as local oscillator at frequencies where conventional waveguide techniques cease to be practical [40].

Another difficulty has been the lack of sensitive heterodyne mixers at terahertz frequencies. Schottky diode mixers have been available for quite some time, but suffer from a large (mW) LO power requirement. Niobium based Superconductor-Insulator-Superconductor (SIS) mixers, well established in the submillimeter, do not operate in the terahertz due to an energy gap limitation of the superconducting material, and non-negligible parasitic junction capacitance of the tunnel barrier (Chaps. 4 & 5). The introduction of the superconducting NbN-film Hot-Electron Bolometer (HEB) mixer [41] has shown much promise as a moderately sensitive, low LO-power demanding, terahertz mixer.

HEB mixers are thermal devices that operate near the superconducting transition temperature ($T_c$). By design the hot electron bolometer is planar with very little parasitic device capacitance. As such HEB mixers have no known upper frequency limit, and have demonstrated sensitivities in the terahertz regime that approach six
1.2. COHERENT (HETERODYNE) DETECTION

times the quantum noise limit [42]. This is well beyond the capability of conventional SIS mixers.

Despite some of the for-mentioned advantages of HEB mixers, they do suffer from a number of issues, none the least of which is a limitation in IF bandwidth. The limited IF bandwidth is in large part due to the thermal time constant(s) of the excited (hot) electrons, phonons, and the surrounding bath temperature. In practice the IF gain bandwidth of NbN HEB mixers lies in the range 2.5 – 5 GHz. Though adequate at submillimeter frequencies (200 – 1000 GHz), at terahertz frequencies a 2 – 3 GHz IF bandwidth constrains the obtainable velocity coverage (Doppler shift). This limits observations to galactic sources and some nearby galaxies. A second disadvantage of a small IF bandwidth is the efficiency at which molecular line surveys may be carried out. One of the big challenges in radio astronomy instrumentation is therefore to extend the IF bandwidth of HEB mixers.

Yet another difficulty is the limited sensitivity of HEB mixers. This is due to a relatively low conversion efficiency (-6 to -10 dB) and a high level of thermal noise at the mixer output. A SIS mixer, on the other hand, can have near unity down conversion efficiency [43], and a mixer output noise very close to the quantum noise limit. As a long term goal several groups are trying to extend SIS technology into the terahertz regime with the use of high bandgap materials (Chap. 5).

In case of extended line sources, such as molecular clouds, an improvement in observing efficiency is obtained by combining multiple pixels into a heterodyne array receiver (Chap. 9). This is not unlike multi-pixel charge coupled devices (CCDs) in the optical and near infrared and direct detection bolometer cameras in the submillimeter. In fact, HEB mixers, unlike SIS mixers which have to have a magnetic field applied to suppress Josephson currents, are well suited for array applications.

For point source observations, raw sensitivity with coherent (heterodyne) detection can be attained by utilizing both polarizations, with the use of a continuous comparison or correlation receiver, and for ground based observations the use of a sideband separating mixer [44, 45]. The above mentioned configuration requires two mixing elements. Still further improvement in sensitivity may be gained with the use of balanced configurations. For example, a dual polarization balanced sideband mixer. These receivers configurations are discussed in some detail in Chaps. 3 & 8. Extending this technology to terahertz wavelength will form a significant challenge in the decade(s) to come.

Amplitude instability of the local oscillator (LO) poses a problem to SIS and HEB mixers alike. However, as noted in Chaps. 10, the problem increases in severity with operating frequency. Minute fluctuations in the LO signal, as a consequence of rapid modulation of the LO-mixer standing wave cavity and/or changes in the output power of local oscillator result in a loss of integration efficiency, platforming in the backend spectrometer, and baseline distortion due to imperfect “signal-reference” subtraction. Balanced receivers with inherent common mode amplitude rejection properties (Chap. 8), form an attractive solution to this problem. Moreover, since balanced receivers have built in robustness to excess amplitude modulated (AM) noise from synthesizer driven local oscillator sources, they also facilitate automated (frequency agile) line surveys (Chap. 2).
Instrument stability is also of utmost importance in the case of high-redshift extragalactic observations, where the observed signals are deeply embedded in the noise. Proper system stability ensures efficient integration of the noise with the highest possible fidelity. Since essentially all HEB observations in the terahertz will be done from high elevations, or space (due to telluric \(\text{H}_2\text{O}\) line absorption) it is questionable whether sideband separation in the terahertz frequency range is the correct approach. The challenge that faces HEB receiver development in the terahertz is therefore to construct (balanced) arrays with adequate IF bandwidth. It is the expectation that the theory and techniques outlined in this thesis contribute to this development.

1.2.2 Why high resolution spectroscopy

By virtue of the coherent receiver down-conversion process to an intermediate frequency, it is possible to achieve very high spectroscopic resolution (\(R=\nu/\Delta\nu \geq 10^6\)), thereby facilitating fine-structure atomic and molecular spectroscopy in the submillimeter and far-infrared. As we have seen, this field encompasses a wide variety of important research topics ranging from protostellar disks and molecular outflows, to galaxy formation.

As was alluded to in the introduction, the detection of dust, or continuum radiation in the submillimeter and terahertz provides a powerful method of studying stars deeply embedded inside dense molecular clouds. It also facilitates the measurement of dust concentrations around protostars, as well as the study of galaxy formation. It is for this reason that a variety highly successful (SCUBA-2, SHARC-2) \([46, 47]\) low resolution direct detection (bolometer) cameras either have been, or are currently in the process of being developed.

To examine the spectral energy density (SED) of objects higher spectral resolution is however needed. For example, the use of an incoherent Fourier transform spectrometer or grating spectrometer provides resolving powers of up to 10,000 (Sec. 2.1). ISO-LWS, the low resolution Fabry-Perot spectrometer (\(R \sim 200\)) on board ESA’s Infrared Space Observatory (ISO) \([22]\) has very successfully obtained spectral information on a wide range of galactic and extragalactic objects. ISO was operational between November 1995 and May 1998. PACS and SPIRE, both low to medium resolution photometer cameras on board the, soon to be launched, Herschel space observatory \([16]\) are expected to study the far-infrared wavelength regime for the first time.

Scientifically speaking, it is currently understood that observations of continuum emission from dust provide insight into the structure of molecular clouds, star forming regions, photodissociated regions (PDRs), and galactic formation (Motte 1998 \([48]\), Dishoeck 2000 \([15]\)), while the spectral energy distribution (SED) provides an indication of the evolutionary state of the objects (Lada 1999 \([49]\)). To study however the physics and chemistry of the gas, only high resolution molecular line observations suffice. For example, the ratios of lines originating at different energy levels of the same molecule provide a convenient way of measuring the temperature structure, local density, and abundance. The ability to disentangle the many gas components responding to these various astrochemical processes requires the high spectral resolution provided by a heterodyne instrument. Given that the research objective of this
thesis is “advanced heterodyne receivers technology”, we have ignored in our discussion topics related to the continuum detection of dust, Sunyaev–Zel’dovich effect, and the cosmic background radiation. These are better treated elsewhere in literature.

1.3 Contribution to the field of far-infrared astronomical instrumentation

A global overview of the contributions to the field of far-infrared astronomy is provided in Table 1.1. The following two subsections discuss the thesis outline and impact on astronomy.

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>SIS and HEB mixer fundamentals &amp; technology development.</td>
</tr>
<tr>
<td>5</td>
<td>NbTiN film, AlN tunnel barrier, 850 GHz twin-slot mixer.</td>
</tr>
<tr>
<td>6</td>
<td>IF bandwidth and mixer gain of NbN HEB mixers</td>
</tr>
<tr>
<td>7</td>
<td>High current density (waveguide) AlN-barrier SIS receiver.</td>
</tr>
<tr>
<td>8</td>
<td>Balanced, Correlation, and Sideband separating receivers.</td>
</tr>
<tr>
<td>9</td>
<td>SIS and HEB mixers on SOI substrates for array applications.</td>
</tr>
<tr>
<td>10</td>
<td>HEB stability and the influence of optical standing waves.</td>
</tr>
</tbody>
</table>

1.3.1 Thesis outline

This thesis begins with a brief introduction to the field of submillimeter and terahertz astronomy. From there it proceeds to outline the submillimeter detection and operational requirements in Chap. 2. This is an important Chapter as boundary conditions are defined for the astronomical instrumentation we concern ourselves with.

Once the operational requirements are established a variety of receiver configurations are introduced in Chap. 3. In particular we touch upon the concepts and implementations of the double sideband (DSB) single-ended receiver, the balanced receiver, the correlation receiver, and the sideband separating receiver.

At its core, frequency down-conversion or heterodyning is accomplished with an highly non-linear mixing element. In the submillimeter, below 1000 GHz, SIS tunnel junctions are now commonly employed thanks to their superior conversion gain and sensitivity. In the terahertz regime, at frequencies beyond 1000 GHz, hot-electron bolometer mixers are at the present time the element of choice. The mixing process in both mixers is fundamentally very different. In Chap. 4 we address the physics of each device, and look in detail into a number of device related peculiarities. With the theory covered, the focus shifts in subsequent Chapters to a variety of realized receiver
designs in the laboratory and on the telescope. We address in Chap. 5 an all NbTiN film quasi-optical mixer that operates in the 800 – 920 GHz atmospheric window. As a result of our work, low loss and high energy gap NbTiN superconducting films have been demonstrated, and since then been incorporated in HIFI mixers bands 3, 4, and 5 [50, 51]. Also in Chap. 5 we explore for the first time the use of an AlN tunnel barrier. The results were so impressive that this technology is currently being baselined at institutions like JPL, TU Delft, the University of Virginia, and the University of Köln.

In Chap. 6, our research leads the way to addressing such important issues as the IF bandwidth and mixer gain of (NbN) hot electron bolometers. Then in Chap. 7 a recently installed (at the Caltech Submillimeter Observatory (CSO)) double sideband high-current density AlN-barrier SIS Technology development receiver (Trex) is introduced. Trex was constructed to prove many of the underlying technologies required in advanced receiver designs (Chap. 8). At the present time Trex serves as a facility heterodyne instrument for high spectral resolution wide bandwidth observations, and extended observations with the SMA [52].

In the first part of Chap. 8 we look in detail at the theory, characteristics, and implementation of balanced receivers, designed to cover the important 180 – 720 GHz atmospheric window(s). The second part of Chap. 8 discusses a correlation, or continuous comparison, receiver to operate between 280 – 420 GHz (covers two atmospheric windows with one SIS mixer). Finally, the last Section of Chap. 8 concerns itself with a 600 – 720 GHz (Alma band 9) sideband separating receiver. Theory, implementation, and measurement results are presented. This is the first time a sideband separating receiver has been implemented at such a high frequency. In our treatment, instantaneous signal and IF bandwidth, sensitivity, and sideband ratio are important considerations.

An important, so far un-discussed, receiver concept is the heterodyne focal plane array. Future astronomical developments will increasingly demand the mapping speed enhancement afforded by multi-pixel focal plane arrays. This thesis will devote Chap. 9 towards integrated arrays receivers with actual realized devices and measurement results.

In all of the above outlined receiver designs, instrument stability is of primary importance. It affects the observed baseline quality and integration efficiency. Indeed, there are many ways in which the stability of the receiver may be compromised. The theory and actual measurement results will be presented (Chap. 10).

### 1.3.2 Impact on astronomy

In the last decade, the field of submillimeter and terahertz astronomy has seen a rapid development. The reasons are many, but several relevant to this thesis stand out: High speed computer technology combined with sophisticated (commercial) electromagnetic simulation software has enabled mixer and local oscillator designs to be much better understood and optimized. A significant amount of money, and thereby expertise, has been injected into the field by large programs such as HIFI [13, 12] and ALMA [6]. In addition there has been a revolution in the automation industry which has enabled commercially viable micro-milling machines. This has allowed waveguide
Table 1.2: Some of the instruments and observatories that have benefited from this work.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Contribution</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIFI</td>
<td>Mixer bands 1,3,4,6,7, Systems stability.</td>
<td>4,5,6,8,10</td>
</tr>
<tr>
<td>HEAT</td>
<td>HEB mixer, System stability, IF bandwidth.</td>
<td>4,6,10</td>
</tr>
<tr>
<td>TELIS</td>
<td>IF bandwidth integrated receivers.</td>
<td>8</td>
</tr>
<tr>
<td>STO</td>
<td>HEB mixer, System stability, IF bandwidth.</td>
<td>4,6,10</td>
</tr>
<tr>
<td>SuperCam</td>
<td>SIS on SOI, Multi pixel development.</td>
<td>9</td>
</tr>
<tr>
<td>ALMA</td>
<td>Band 10: AlN/NbTiN films, Band 5,9: 2SB design.</td>
<td>5,8</td>
</tr>
<tr>
<td>Casimir</td>
<td>AlN barrier/NbTiN ground plane.</td>
<td>5</td>
</tr>
<tr>
<td>CSO</td>
<td>SIS mixer development, Stability, IF bandwidth.</td>
<td>4,5,7,8,A–C</td>
</tr>
<tr>
<td>APEX</td>
<td>Waveguide transition, System stability.</td>
<td>7,A</td>
</tr>
</tbody>
</table>

Techniques to be extended to nearly 2 THz [36, 37]. And finally, receiver backend technology has seen a tremendous advancement, to the point where it is now possible to process IF bandwidths of multi-pixel focal plane arrays. Much of the research in this thesis has benefited from the above mentioned revolution.

Many of the ideas and concepts developed in this thesis are either already implemented, or are in the process of being implemented in submillimeter and terahertz instrumentation and telescopes around the world. And, with the high resolution Heterodyne Instrument (HIFI) on ESA’s Herschel satellite [16], soon also in space. Table 1.2 serves as a summary.

Examples of instruments and projects that have utilized various aspects of the presented research are, aside from HIFI [12]; the High Elevation Antarctic Telescope (HEAT) [53], the TERahertz and submillimeter LImb Sounder (TELIS) [30], the Stratospheric Terahertz Observatory (STO), SuperCam [54], Atacama Large Millimeter Array (ALMA) [6], Casimir [55], and the Caltech Submillimeter Observatory on top of Mauna-Kea, Hawaii [4].

HEB mixer development for HIFI mixer bands 6 & 7 has encouraged fundamental research into the physics of (NbN) hot electron bolometers at several laboratories, and in particular by Prof. Teun Klapwijk’s group at the Kavli Institute of Nanoscience, Delft University of Technology [56, 57, 58, 59]. The research into HEB devices is driven by the urgent need to improve the sensitivity, IF bandwidth and system stability of hot electron bolometers mixers. Much progress has been made in recent years, with issues such as IF bandwidth and mixer gain addressed in Chap. 6 and system stability in Chap. 10.

For HEAT, the hot electron bolometer mixers are anticipated to be fabricated at the TU Delft (NbN films from Gol’tsman group, Moscow State Pedagogical University, Russia [60]). In this case, implications from Chaps. 6 & 10 are directly applicable. Similarly, HEB instrumentation on STO will also benefit directly from the presented research.

IF bandwidth has also been an issue with the integrated flux-flow (SIS) mixers on TELIS [30]. Here integrated techniques developed in Chaps. 7 & 8 have been applied.
very successfully [61]. In case of the SuperCam [54], the 64 pixel heterodyne array camera from the University of Arizona (Prof. Walker’s group), silicon-on-insulator (SOI) technology is employed to produce very thin (3 µm) silicon substrates with gold beam leads. These are needed to move away from conventional single pixel quartz substrates to large format arrays, and to facilitate scalability to terahertz frequencies (Chap. 9).

ALMA mixer band 10 [6] and the Casimir instrument on SOFIA [20] use the NbTiN films and SIS AlN barrier technology of Chap. 5, as developed for HIFI mixer bands 4 & 5 [50, 51]. Furthermore, the ALMA band 5 (163 – 211 GHz) sideband separating mixer work at GARD (Group for Advanced Receiver Development at Chalmers TU, Sweden) [62] employs the quadrature hybrid outlined in Chap 8, and the ALMA band 9 (602 – 720 GHz) sideband separating mixer [63] the research and development of Chaps. 7 & 8.

It should also be observed that instrument stability is a common thread throughout this thesis (Appendix A). It is fair to say that all the above mentioned instruments have benefited from this work. And last but not least, to assist the science at the Caltech submillimeter Observatory, much effort has gone into the development of wide RF and IF bandwidth balanced– and correlation receivers, as outlined in Chap. 8.
Bibliography


[3] [Online]. Available: http://www.rssd.esa.int/Planck/


[16] [Online]. Available: http://sci.esa.int/


[22] [Online]. Available: http://www.sci.esa.int/iso/


[26] [Online]. Available: http://www.spire.rl.ac.uk/


[35] [Online]. Available: http://www.esa.int/esaMI/Rosetta/


