Mémoire

Auteur : Blond Hanten, Elsa
Promoteur(s) : Jehin, Emmanuel
Faculté : Faculté des Sciences
Diplôme : Master en sciences spatiales, à finalité approfondie
Année académique : 2023-2024
URI/URL : http://hdl.handle.net/2268.2/21661

Faculty of Sciences
Master in Space Sciences
2023-2024

Investigation of unidentified lines in optical cometary spectra

Elsa Blond Hanten

Supervisor:
Prof. Emmanuël Jehin

Members of the evaluation committee:
Prof. Emmanuël Jehin
Prof. Damien Hutsemékers
Prof. Denis Grodent
Dr. Johan de Keyser

Comets are like cats: they have tails, and they do precisely what they want.

David H. Levy
David Levy's Guide to Observing and Discovering Comets, p.1, Cambridge University Press

Painting of the comet C/1858 L1 (Donati) taken from the Kometarium website.

Abstract

The first objective of this thesis focuses on enhancing the Atlas of cometary lines developed by Hardy [1] and the ULiège COMETA team in 2022. This Atlas is based on the spectra of comets C/2002 T7 (LINEAR), C/2016 R2 (PANSTARRS), C/2002 X5 (Kudo-Fujikawa), and 8P/Tuttle obtained by the ULiège team using the high-resolution Ultraviolet and Visual Echelle Spectrograph (UVES). The use of spectra that are centred and offset in the tail direction allows the identification of new lines, especially ionic ones. A total of 579 new lines have been identified in the spectra, among which 57 were previously unidentified lines in the Atlas. The remaining 522 lines are new lines that had not previously been detected in the version of the Atlas developed by Hardy. The second objective of the thesis focuses on developing a fluorescence model of imidogen (NH). We describe how the model was set up, and then present the results obtained from the model. The thesis ends with numerous perspectives on continuing the identification of new lines using different comparison techniques.

Keywords: comet, cometary atlas, spectrum, fluorescence, UVES

Acknowledgement

First and foremost, I would like to thank my supervisor, Professor Emmanuël Jehin, for giving me the opportunity to embark on this master's thesis. I feel very fortunate to have progressed under his inspiring supervision and to have benefited from his in-depth knowledge and sound advice. His passion for cometary research is contagious. The trip to Morocco he organised, even though he could not be physically with us, was a turning point in my very young career, because I came back with stars in my eyes, leaving volcanoes to the stage of a child's dream.

I am profoundly grateful to Professor Philippe Rousselot for introducing me to the fluorescence model; his precious guidance allowed me to deepen my theoretical knowledge. Professor Rousselot and his team warmly welcomed me in March 2024 at UTINAM Institute of the University of Franche-Comté, where I had the great opportunity to experience a different research environment, gain valuable scientific knowledge, and visit the impressive Besançon Observatory.

This thesis would never have been possible, in the truest sense of the word, without the work of Pierre Hardy. My dissertation is truly an extension of the ULiège team's Atlas. Thank you, Pierre, for generously sharing your knowledge and the work you accomplished during your master's thesis.

I realise what a privilege it is to go to university and acquire in-depth knowledge of the things I am passionate about. ULiège is one of those places that formed me in every way. I would like to thank my fellow students with whom I have shared this journey. Thanks to the Parsercle and its members, for the collective efforts to restore the telescope on the roof, for side-by-side learning of administrative and legal procedures, and for the yummy truffles. Special thanks go to Justin. Over the past 5 years, we have conversed in a colourful language, understandable only to insiders; he has kept my mind on its toes with his daily riddles.

Thanks to my parents for their unconditional support in all my undertakings, and especially during the final sprint of my master. And last but not least, I thank my cat Pat, for passing on to me his unfailing zen attitude. Comets are like cats ...

Contents

Introduction

Comets have always fascinated me. Their appearance in the sky is one of the most beautiful natural spectacles. Comets are believed to have brought water, the source of life, to Earth. To understand how comets contributed to the emergence of life on Earth, it is first necessary to understand their composition. This subject is under investigation and is the focus of the research in this master's thesis.

Comets have long captivated humans, but our perception of them has evolved over time. In ancient times, comets were considered signs of misfortune and were attributed to atmospheric phenomena. This perception shifted during the Renaissance, which marked the starting point for a more scientific understanding of comets as distant celestial objects. Scientific interest in comets is multifaceted. Alongside knowledge of comet orbits and their formation processes, scientific research also focuses on their composition. To better understand their chemistry, researchers analyse data collected from a variety of sources, such as optical, infrared, sub-millimetre (sub-mm), and radio spectroscopy, as well as mass spectroscopy. Mass spectroscopy, as used in the Rosetta mission, stands out as the most effective method for the precise chemical analysis of comets, enabling researchers to identify their composition. However, a limitation of mass spectroscopy is its requirement for in situ measurements. Using optical spectroscopy which was the first instrumental technique developed to analyse the composition of comets and also the method used by the ULiège comet team, researchers have identified molecular lines in cometary spectra. With the identification of molecular lines in optical spectra, scientists realised the need to catalogue them. Accordingly, cometary molecular lines are mapped in atlases.

The atlas of Comet 122P/de Vico [2], developed by Cochran and Cochran in 2002, is often considered the reference cometary atlas. Its spectrum ranges across wavelengths from 3830 Å to 10192 Å with gaps starting from 5777 Å. The authors identified 12,219 emission lines and detected 4,055 unidentified lines. Several older and more recent atlases also exist, such as those of Comet 153P/Ikeya-Zhang (Cremonese et al., 2006 [3]) and Comet C/2020 F3 (NEOWISE) (Cambianica et al., 2021 [4]).

In 2004, the ULiège comet team used the Ultraviolet and Visual Echelle Spectrograph (UVES) to obtain a spectrum of Comet C/2002 T7 (LINEAR), ranging from 3100Å to 10000Å. This spectrum has one of the highest resolutions and was used by Hardy [1] and the Liège COMETA[1] (COmets, METeors & Asteroids) team to create a new Atlas[2] in 2022. Owing to its higher resolution, 25,345 lines were detected in the spectrum, 20,417 of which were identified, whereas 4,928 lines remain unidentified. Despite the existence of numerous atlases, some lines have remained unidentified for over 20 years.

My master's thesis builds on the atlases of Cochran and Cochran, and Hardy. The former serves as a reference for both identified and unidentified lines, whereas the latter embodies the indexing and cataloguing that I aim to enhance through my work. Specifically, my thesis fits into the existing body of work as follows (see Figure 1.3 in Section 1.2):

Cochran and Cochran based their atlas on line lists from the existing literature. Their atlas includes both identified and unidentified lines. In 2022, Hardy started to design an Atlas using the data obtained from UVES spectra of Comet C/2002 T7 (LINEAR). He developed a code that automatically detected and identified the emission lines of Comet C/2202 T7. Most of the lines were identified using the ExoMol database, but some of the identified lines stem from the atlas of Cochran and Cochran. Due to the use of a high-resolution spectrum and the availability of updated line lists, Hardy's Atlas lists more identified and unidentified lines than that of Cochran

[1] www.cometa.uliege.be

[2] The ULiège cometary Atlas is always written with a captial A.

and Cochran. Regarding the contribution developed in the framework of this thesis, I include the unidentified lines from Hardy's Atlas as well as other unidentified lines in my analysis, relying on data from the same spectrograph as that used by Hardy. Therefore, at first sight, it appears that my approach in this master's thesis is similar to that of Hardy. However, it differs from Hardy's in that I use other spectra, thereby multiplying the possibilities of better understanding the observed lines and finding new ones. By using the spectra of several comets and considering that different comets have different compositions, valuable information can be gained on other molecules. Furthermore, by using spectra from the same comet taken at different positions in the comet atmosphere (towards the Sun or in the tail direction), the lines of ions and those of neutral molecules can be clearly distinguished.

The positions of the spectral lines of iminogen (NH) are known with high precision. The fluorescence model developed in the framework of this thesis provides new information about this molecule. Indeed, the purpose of a fluorescence model is to predict the intensities of the NH lines. When these intensities are known, the abundance of the molecule in a certain environment can be determined. The implementation of the model will also offer a more theoretical approach to the topic of cometary lines.

The thesis is organised into six sections. Section 1 introduces cometary atlases, particularly the Atlas developed by the ULiège team on which the thesis focuses. The aim of the dissertation is to add to the existing cometary Atlas produced by the ULiège team and to enrich it by identifying new molecular lines. The objectives and purpose of the thesis are also presented in this section. Section 2 discusses the evolution of the perception of comets over time and includes general comments about comets. To highlight the scientific interest in comets and the state of scientific knowledge about them, to which my work aims to contribute, I have provided a historical context of human curiosity about comets. In addition to the classification of comet orbits, anatomy, and composition, this section introduces the comets that I have been studying. In Section 3, which is dedicated to spectroscopy, I explain the theoretical aspects of how spectroscopy reveals the origin of both atomic and molecular emission lines in cometary spectra. Subsequently, I introduce the concept of fluorescence, which is crucial for the fluorescence model of NH that I developed in my research. Section 4 describes the methods used in this study to measure and process the cometary spectra and their characteristics. In addition to outlining the methods employed to discover new lines in the spectra, I present a step-by-step approach for constructing the fluorescence model. The results are released in Section 5 and effectively contribute to the development of the ULiège cometary Atlas by increasing the number of identified and decreasing the number of unidentified lines. Regarding the novel fluorescence model, the associated results can be defined as intermediate and can serve as a reference for perfecting the model through further research. Section 6 concludes the dissertation, providing various scientific perspectives on, for example, potential further improvements to the Atlas. Finally, Appendix A (and a web link) includes one of the wavelength tables of Lew (1976, [5]), which I digitised and recalculated; they may be useful to other researchers. Appendix B contains the Python code devolped for the NH fluorescence model.

1 Cometary Atlas

The aim of this section is to introduce the cometary Atlas and the purpose of the thesis. The Atlas is the platform that archives and organises the identified lines in the cometary spectra. Beyond its storage function, the Atlas is also the instrument we use during the line identification process. Indeed, it allows to compare different molecular lines, different spectral regions, and different cometary spectra (Section 1.1). Identifying unidentified lines, as well as new lines, is one of our key concerns in this thesis (Section 1.2).

1.1 Atlas, a tool to map cometary lines

Before discussing the specifics of the cometary Atlas, which served as the primary tool for this master's thesis, we will briefly explain the origin of the word "atlas", what an atlas is, and how it can benefit cometary research.

1.1.1 From "Maps" to "Atlas" to "Cometary Atlas"

Atlas was considered both a legendary king of Mauritania and a Titan of Greek mythology, condemned by Zeus to hold up the sky [6]. Although people have used maps for millennia, it was not until the 16th century that civilisations began producing comprehensive atlases. During this period, Mercator's term "atlas" replaced the previous term "cosmography":

> Atlas referred to a portrait of King Atlas, a mythical African monarch. King Atlas invented the first celestial globe. A celestial globe is a ball-shaped map of the stars and constellations. Celestial globes were very important in navigation, when sailors used stars to determine their position at sea. Mercator showed King Atlas to demonstrate his importance to navigation. This was the first time the term was applied to a collection of maps. Eventually, "atlas" came to be used for any book of maps. [https://education.nationalgeographic.org/resource/atlas/]

The evolution from maps to atlases unfolded in several stages. It began with the geographical knowledge of the ancient world (Ptolemy, 2nd century) and progressed through medieval maps created for strategic and commercial purposes, such as the 1246 compilation by a canon of Mainz and the Ebstorf world map. The 13th century saw the rise of portolan charts, which emphasised precise nautical surveys and introduced the term "tabula". During the Renaissance, maps became more specialised, depicting the globe and various aspects like politics and linguistics, leading to the modern atlas. Modern atlases emerged with works like Mercator's *Tabulae geographicae* (1585) and Ortelius's *Teatrum orbis terrarium* (1570), transitioning from universal to national atlases and distinguishing between general and thematic atlases.

Nowadays, the term atlas can be used in a broader sense to refer to a collection of data and charts, either in the form of a book or an online platform. The ULiège Cometary Lines online Atlas is a collection of cometary lines that enables comparisons and analyses between different spectral regions from different comets, which fulfils the task highlighted by Hruby (2015) cited by the ICA Commission on Atlases. The first atlases established in the field of astrophysics mapped stars and other celestial bodies.

The necessity of consolidating all this information into a single location has increased as the number of molecules detected in comets has expanded. This process started in 1864, when the first cometary molecule, C_2, was detected by Donati in the C/1864 N1 Tempel. The detection of the second and third molecules, CN and C_3, followed 17 years later with a photographic spectrum of Comet C/1881 K1. Since then, many cometary molecules have been identified in great comets like Halley, Kohoutek, Morehouse, Hyakutake, and Hale-Bopp. Since 2015, ESA's Rosetta mission to Comet 67P/Churyumov-Geraismenko has led to the identification of more than 40 new molecules [7].

1.1.2 The Atlas of the ULiège cometary team

As the number of known cometary molecules has increased, the need for a support that combines and classifies their emission lines has grown. It is in this context that the first cometary atlases were created. In 1958, Swings and Haser published the *Atlas of Representative Cometary Spectra* [8], and Arpigny *et al.* published their *Atlas of Cometary Spectra* in 1987 [9]. The goal of these and future atlases is to assist scientists in identifying molecules in their own spectra (Rahe, 1981 [10]).

Currently, the most commonly used atlas is the one published by Cochran and Cochran [2] in 2002 of Comet 122P/de Vico. In 2022, Hardy [1] presented a new Atlas to make use of the very high-resolution spectrum of Comet C/2002 T7 (LINEAR) obtained by the ULiège team. From the strong basis established by Hardy, in this work, we further investigate the unidentified lines that remain in this Atlas. The website hosting the Atlas was developed by Sandrine Sohy of the COMETA team, ULiège. The Atlas as an online tool is a valuable asset compared to conventional atlases. It provides an interactive platform for cometary scientists as it caters to their needs. Ongoing contributions to the Atlas give researchers instant access to updated versions. In this sense, the online Atlas can be qualified as a *living atlas*. A new version of the online Atlas was developed in the framework of this thesis[3]. This site allows users to visualise any region within the spectral range of T7 (3075–9980 Å), selecting between the two spectra that are already in the system, T7 and 16R2, or uploading their own spectra. It is possible to choose the molecules one wants to plot from the list of molecules that are already identified in the Atlas. Additionally, it is possible to retrieve a table that presents the wavelengths, the species, and the transition details for specific molecular lines. Figure 1.1 shows the homepage of the online Atlas and Figure 1.2 an example of the plots it presents.

[3]Our version of the Atlas can be accessed via the following URL: http://www2.orca.ulg.ac.be/Cometary_atlas_Elsa/Cometary_atlas.html

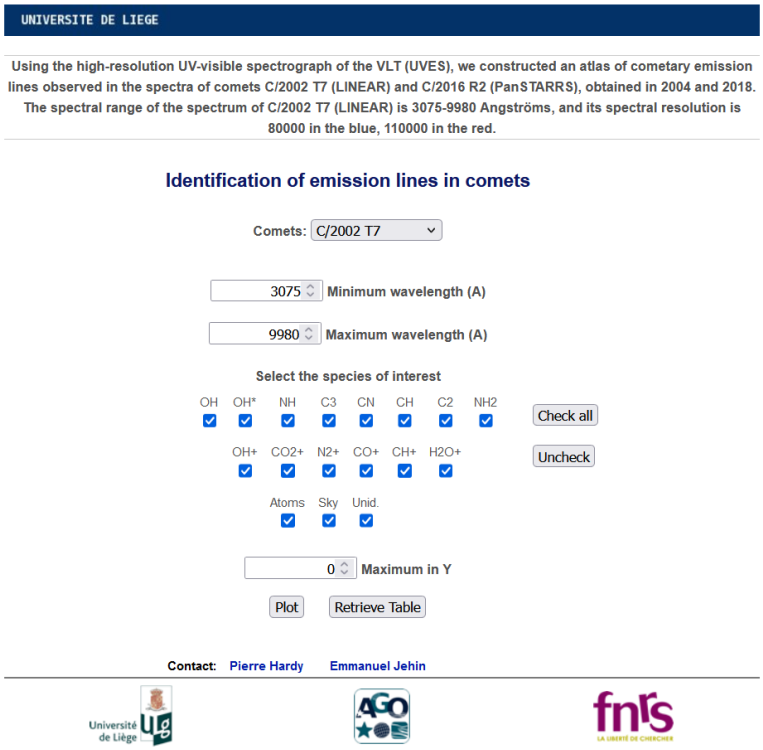


Figure 1.1: Homepage of the ULiège Cometary Atlas.

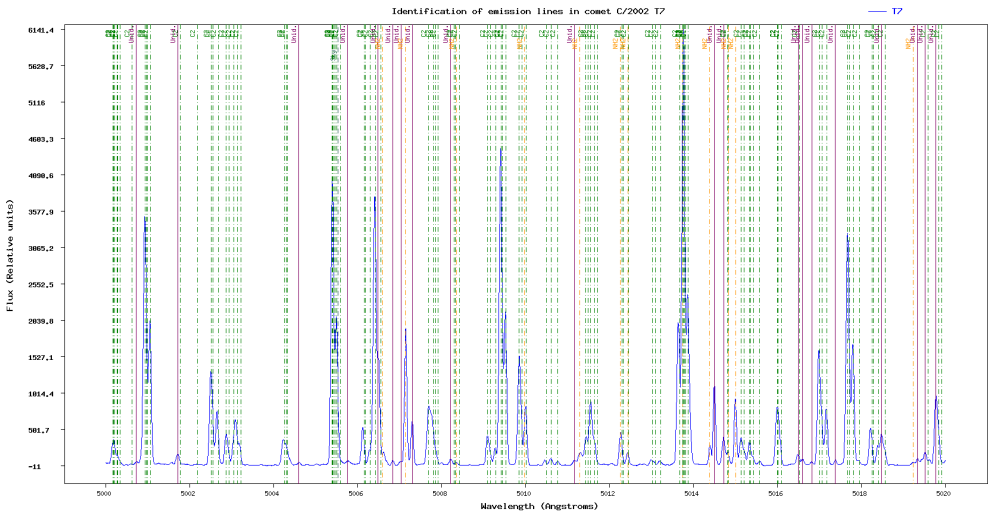


Figure 1.2: An example of a plot produced by the ULiège Cometary Atlas.

The process of identifying new spectral lines used this Atlas, several available spectra, and existing literature on comet emissions, as well as line lists that Hardy has not yet explored. From there, a trial-and-error approach was employed. We analysed the unidentified bands in the spectrum to identify patterns that might indicate the molecular origins of the lines. When we came across a promising band in an article or line list, we checked the Atlas to see if the band was visible in any

of the spectra. This part of the work required a detailed analysis of the entire spectral range of the various spectra.

Line lists in articles often indicate a theoretical wavenumber in cm^{-1} obtained in a vacuum. However, the spectra of the comet were taken on the ground, so the light had to pass through the atmosphere. Therefore, the theoretical wavenumbers in the vacuum need to be transformed into wavelengths measured in angstroms in the air. The formula proposed by Morton (1991, [11]) was used for this purpose:

$$\frac{\lambda_{vac}}{\lambda_{air}} = 1 - 8.34254 \times 10^{-5} + \frac{2.406147 \times 10^{-2}}{130 - s^2} + \frac{1.5998 \times 10^{-4}}{38.9 - s^2} \quad (1)$$

where $s = \frac{10^4}{\lambda_{vac}}$ with λ in Å.

1.2 The aim of the thesis

The goals of this research are directly tied to the current status of the ULiège Atlas, as detailed in the previous section. Presently, the Atlas includes 4,928 lines that have yet to be identified. Thus, the main objective of this work is to identify a subset of these lines. This will be achieved through an extensive review of the literature to discover new and updated line lists, as well as by keeping track of new identifications of bands in cometary spectra. The identification process will involve comparing spectral regions within individual spectra and across different cometary spectra. In cases where the molecular origin of an emission band cannot be definitively determined, efforts will focus on narrowing down potential candidates by eliminating unlikely options. By incorporating additional spectra beyond the central spectrum of C/2002 T7 (LINEAR), the research aims to reveal new emission bands. These newly detected bands will undergo the same identification procedures as those applied to previously detected bands. All verified discoveries will be added to the Atlas, thereby increasing its utility. As a result, the Atlas will be viewed as a dynamic resource, continually updated as new advancements emerge in the field.

The final part of this thesis involves developing a fluorescence model of the diatomic molecule imidogen (NH). This model will enable the determination of molecular line intensities and facilitate the calculation of the molecule's abundance. The model will be constructed using equations from Zucconi and Festou [12] to calculate the population of energy levels, with molecular constants sourced from the literature.

In conclusion, this thesis aims to expand the ULiège Atlas, thereby enhancing its value as a resource for cometary research within the scientific community.

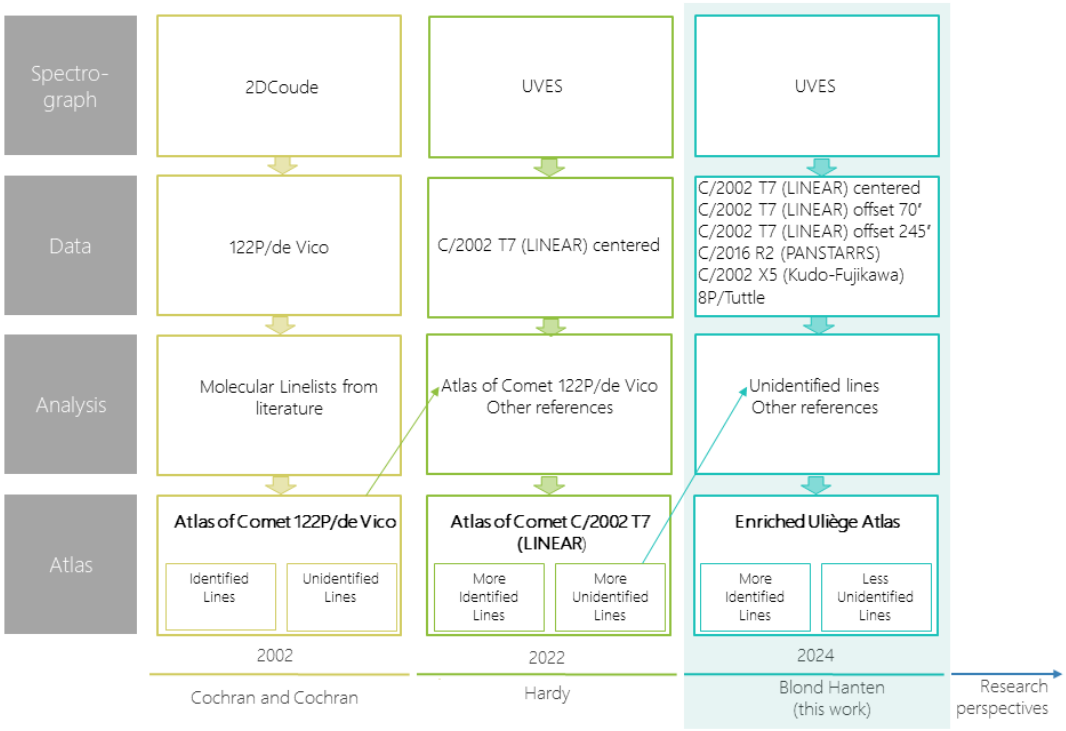


Figure 1.3: How my master's thesis builds on existing atlases and enriches the ULiège Atlas.

2 Comets

2.1 Relics from the past: from superstition to scientific quest

2.1.1 Antiquity and Middle-Ages: Burning Beams and Windy Exhalations

Since the dawn of humanity, comets have been a great source of fear, apprehension, and mystery, as shown by historical reports and ancient artwork. For much of human history, comets have had a reputation as the root of disastrous events, such as the death of Attila the Hun or the fall of Constantinople. Frequent superstitions surrounding around comets include both warnings about future tragedies and the bearing of happy news, depending on the believer [13]. Pliny the Elder held unique convictions about comets: that these small icy bodies brought terror everywhere but in Rome(for instance during the reign of Nero, known for his cruelty). In this city, he viewed comets as a blessing; for example in 44 BC, during the funeral games of Cesar, a comet skimmed the sky (Herbert-Brown, 2007 [14]). Pliny's curiosity for comets pushed him to study their shape and appearance, as presented in the drawing shown in Figure 2.1.

Beliefs and imagination surrounding comets have taken on different forms. Comets have been represented as "a burning torch of an extraordinary size, to which one gave the name of burning beam" (Diodorus, 90–30 BC, Historical Library). They were also seen as "windy exhalations" and belonging to the sublunary world (Aristotle). This view was contradicted by Seneca (4 BC–65 AD), who considered that comets may be more distant than the Moon. He wrote a text announcing a scientific approach to comets., stating the following: "Why be surprised that comets, which show themselves so rarely to the world, are not yet for us constrained by fixed laws, and that one does not know from where they come and where they will stop. [...] An age will come when what is mystery to us will be revealed by time and by the studies accumulated for centuries. [...]" (*Des questions naturelles*, Seneca [15])

In ancient Christianity, comets were seen as favourable omens: the guiding star, often identified as a comet, announcing the birth of Christ.

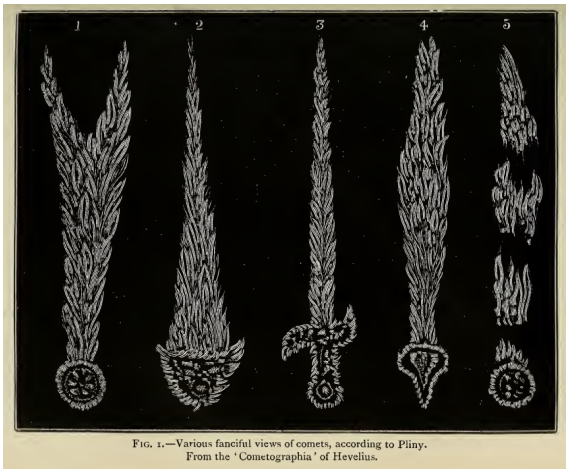


Figure 2.1: Drawings of comets by Pliny the Elder. Source: *The Mysteries of Time and Space* (Proctor, 1883).

2.1.2 Renaissance: Beginning of the scientific approach

The scientific approach began to develop during the Renaissance. Representations of comets became more realistic and the first European scholars tried to answer the questions raised by Seneca. Their main contributions are as follows:

- Describing the apparent motion of six comets (Paolo Toscanelli, 1397–1482);
- Calculating the distance of a comet, estimated –erroneously– as nine Earth radii (Johannes Müller von Königsberg, 1436–1476, *De Cometae*, 1531);
- Fixing the distance to comets at a greater distance – supra-lunar (Gerolamo Cardano, 1501–1576);
- Observing that the direction of the comet's tail points away from the Sun (Girolamo Fracastoro, ca. 1478–1553, and Petrus Bienewitz or Bennewitz, Apianus or Apian, 1495–1552);
- Obtaining a reliable distance for the comet in 1577 (Michael Mästlin, 1550–1631, and Tycho Brahe, 1546–1601);
- Demonstrating that comets are not atmospheric phenomena by comparing the position of the comet relative to nearby stars using two simultaneous measurements (Tycho Brahe, 1546–1601, and Thaddaeus Hagecius, 1525–1600); and
- Exploring for the first time the nature of comets, thereby opposing Galileo and Tycho Brahe and ending with Johannes Kepler's (1571–1630) appropriate description: "The head is like an accumulated nebula, somewhat transparent; the tail or beard is an emanation of the head, expelled by the Sun's rays in the opposite direction, and in this continual outpouring the head is finally exhausted and consumed, so that the tail mean the death of the head" (Johannes Kepler, *Pamphlet Tychonis Brahei Dani Hyperaspistes*, 1625).

2.1.3 17th century: "the causes and effects of comets"

During the 17th century, scientists continued their efforts to understand comets. One of the first known scientific symposia, the symposium on "the causes and effects of comets" was held in Paris in 1665. Scientific contributions to this symposium were diverse, even contradictory, as summarised below:

- the movement of a comet (Adrien Auzout, 1622-–1691);
- the paths of comets;
- their trajectories;
  - straight or nearly straight (erroneously; Kepler);
  - circular (Mästlin and Tycho Brahe) or oval (Tycho) orbits around the Sun; and
  - parabolic trajectories (Johannes Hevelius, 1611–1687, *Descriptio cometae anno 1665*, *Cometographia*, 1668).
- their periodicity (Pierre Petit, 1598–1667, *Essay on the nature of comets of 1665*; Jean-Dominique Cassini, 1625–1712), although the examples he observed in the newly created Paris Observatory were in fact two different comets.

2.1.4 Newton: "About the motion of bodies that rotate" [4]

This historical overview of knowledge concerning comets would be incomplete without highlighting Newton's contribution. Isaac Newton (1643–1727) applied his theory of universal gravitation to the movement of planets around the Sun and demonstrated the three empirical laws established by Kepler. Newton discovered that Kepler's theories aligned with comet observations. He also developed a technique to determine the parabolic orbit of a newly observed comet by measuring

[4] We borrowed the title of this subsection from the book that Newton offered to his friend Halley "De motu corporum in gyrum".

its successive positions relative to stars. He even convinced his friend, Edmond Halley, with his demonstration.

Newton's ideas were revolutionary and received limited acceptance, even among his peers. The posthumous publication of the commented translation of the *Philosophiæ naturalis principia mathematica* into French by Marquise Gabrielle Émilie du Châtelet (1709—1749) helped Newton's ideas to gain recognition, all the more so as it coincided with the return of Halley's comet in 1759.

2.1.5 Halley: the return

Comet Halley has earned its fame due to its regular appearances in the sky. However, it was not until the English astronomer Edmond Halley suggested that the comets observed in 1531, 1607, and 1682 were actually the same comet that it became known that the comet lighting up the sky every 76 years throughout the centuries was the same one. Halley made this assumption by recognising the similarities in the orbits of the comet. This breakthrough led to an understanding of periodic comets, allowing Halley to accurately predict the comet's return in 1758. His scientific accomplishment was rewarded by naming the comet "1P/Halley". Looking back in history, knowing that Comet Halley traces and probably had traced its path in the sky every 76 years, astronomers were able to identify Comet Halley in more than 2000-year-old observations. Figure 2.2 shows one of the most famous recordings of comet Halley, depicted in the Bayeux tapestry which portrays the battle of Hastings in 1066, and in which Comet Halley makes an appearance [16].

Comet Halley's repeated return has triggered a craze for comets and a wave of comet enthusiasts. Due to their rarity, many amateur and professional astronomers aspire to discover a new comet. Knowledge could not have reached its current level, without the efforts of centuries of comet hunters who spotted new icy bodies in the sky. Since, an array of new comets have been discovered thanks to surveys and robotic telescopes. Two examples of such telescopes are LINEAR (Lincoln Near-Earth Asteroid Research) and Pan-STARRS (Panoramic Survey Telescope And Rapid Response System), having discovered comets T7 and 16R2 that we introduce Section 2.6.

Figure 2.2: Bayeux Tapestry, showing the battle of Hastings and Comet Halley. Source: Anglo-Saxon Archaeology, taken from www.nasa.gov.

2.1.6 HERstory: hidden figures

Science, particularly astronomy, has long been dominated by men; in contrasts, women were nearly invisible in the field until the 20th century. Despite this, some women managed to leave a mark on history, overcoming significant obstacles to have their work on comets recognised (Observatoire de Paris [17]):

- assisting her husband in his observations of comets: Sophia Brahe (1556–1643);
- translating and commenting Newton's *Philosophiæ naturalis principia mathematica* into French: Gabrielle Emilie, Marquise du Châtelet (1709–1749)
- predicting new comets: Maria Winkelmann (1670–1720), however her observations of a non-periodic comet were attributed to her husband;
- predicting the return of Halley's comet in 1759: Nicole-Reine Lepaute (1723–1788);
- discovering eight comets: Caroline Herschel (1750–1848);
- discovering a comet: Maria Mitchell (1818–1889); and
- observing comets: Dorothea Klumpke (1861–1942).

2.2 Comets: Remnants of the early Solar System and their role in the origins of life on Earth

Comets are part of the family of small bodies and are relics of our early Solar System. Their formation dates back to a time where the protoplanetary disk surrounded the proto-Sun. The strong temperature gradient in this disk allowed the median outer region to be lower than the freeze out temperature of molecules. Consequently, the molecules in this region are not chemically active and settled within ices that comets are made of. Since cometary material has undergone limited transformation, it is amongst the most pristine material in the Solar System. Therefore, studying comets provides valuable insights into the composition of the early Solar System.

Additionally, comets might have brought water on Earth. Indeed, the planetesimal that later evolved into the Earth was formed during the early stages of the Solar System, in the inner part of the planetary disk. At that time, this region was too hot to allow the planetesimal to contain ice. As such, the water currently present on Earth must have been brought at a later time. The hypothesis that water was acquired from inward-moving ices once the inner Solar System cooled below the condensation point has been rejected due to the action of proto-Jupiter. Therefore, it is plausible that water was brought to Earth by comets after the dissipation of the solar nebula (Alexander et al., 2018 [18]). In addition to potentially bringing water, comets may have played a crucial role in the development of life on our planet since it is thought that comets also brought organic material and molecules essential for prebiotic evolution (McKay, 2021 [19]). Several scenarios can lead to life on Earth; broadly, these can be divided into two classes: the molecules for life have an endogenous or exogenous origin. An endogenous origin suggests that the required molecules formed on Earth through various processes, such as those involving hydrothermal vents at the bottom of the ocean, which influence ocean acidity. In contrast, an exogenous origin posits that prebiotic molecules formed in space and were delivered to Earth by comets or asteroids. For this scenario to be plausible, three conditions need to be fulfilled:

1. a production mechanism for the material outside of Earth;
2. an interplanetary reservoir where material can be stored, such as asteroids or comets; and
3. the survival of the material upon the impact.

The detection of glycine and phosphorus by the Rosetta-Philae mission in the coma of Comet 67P/Churyumov-Gerasimenko in 2016 supports the exogenous hypothesis. Indeed, glycine is the

simplest amino acid and is often found in proteins while phosphorus is a building block for DNA. This evidence suggests that comets could indeed carry the necessary ingredients for life. However, it is increasingly believed that both endogenous and exogenous formation processes contributed to creating the prebiotic world, which later evolved into a more complex biotic system. The exact processes that led to the origins of life are still unknown, but studying comets may help to fill some of these knowledge gaps [20].

The name "comet" originates from the Greek word "*komēˉ*" meaning hair, which later transformed into "*komētēˉs*" meaning long-haired (star) (Whipple, 1974 [21]). Indeed, comets appear as fuzzy or hairy stars in the sky and have been observed over the last few millennia. Their peculiar appearance coupled with their rarity and rapid movement relative to the background stars have induced strong reactions, including fear among ancient peoples.

Despite being the most numerous celestial bodies in the Solar System, comets are only rarely seen. This is due to the location of their reservoirs. The Kuiper Belt is the closest reservoir to Earth. It is located just outside the orbit of Neptune, at 35 to 48 au from the Sun, and has a flat annulus shape. This reservoir is thought to contain 200 million icy bodies larger than 2 km, the total mass of which adds up to a tenth of the mass of the Earth. The second reservoir is the Oort Cloud, which is located at a much greater distance, between 20,000 and 100,000 au from the Sun. This larger reservoir is believed to contain a trillion objects larger than 2 km.

Cometary objects remain in their respective reservoir, until a gravitational perturbation provokes a change of trajectory. They then have the possibility to become active comets if their new trajectory brings them closer to the Sun (Ory, 2021 [22]).

The active life of a comet outside of its reservoir does not last forever. Indeed, comets will inevitably disappear after some time. Their fate varies among the four following scenarios (Oort, 1986 [23], Jones *et al.*, 2018 [24], [25]):

- collision with either the Sun or another planet, for example Comet C/1979 Q1 (SOLWIND) which vaporised in the Sun's atmosphere;
- disintegration due to tidal forces from the Sun or through the rapid sublimation of the ices inside the comet. Indeed, comets are only loosely bound due to their rubble pile structure;
- extinction when the comet has exhausted all of its volatile material; and
- ejection from the Solar System due to the gravitational perturbation of a giant planet.

2.3 Orbits and classification of comets

Comets move according to one of three types of orbits: elliptical, parabolic, and hyperbolic. In astronomy, it is common to organise objects into different classes. This allows for a better overview and the possibly to detect certain patterns in these classes. It is therefore only natural that such classes also exist for comets. One of the first-developed classification schemes was based on the comet period. Figure 2.3 illustrates the traditional way of classifying comets based on period. However, simulations by Levison & Duncan (1994) [26] have demonstrated that the periods of comets may change over their lifetime, causing comets to move from one family to another; however, an ideal taxonomy should assure that the classification of an object does not change over time. For this reason, Levison proposed a new taxonomy based on the Tisserand parameter T [27]. The Tisserand parameter is defined as follows:

$$T = \frac{a_J}{a} + 2\sqrt{\frac{(1-e^2)a}{a_J}}\cos i \tag{2}$$

where a is the semi-major axis, i is the inclination, and a_J is the semi-major axis of Jupiter. The family tree for this new classification is shown in Figure 2.4 and has the advantage that comets

should not change family during their lifetime.

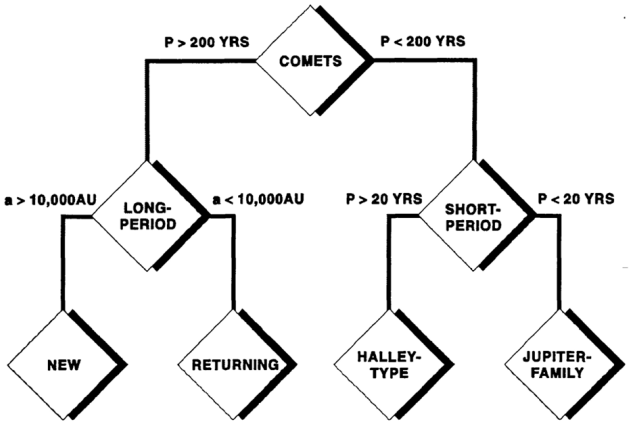


Figure 2.3: Traditional family tree classifying comets based on their period. Source: Levison (1996) [27].

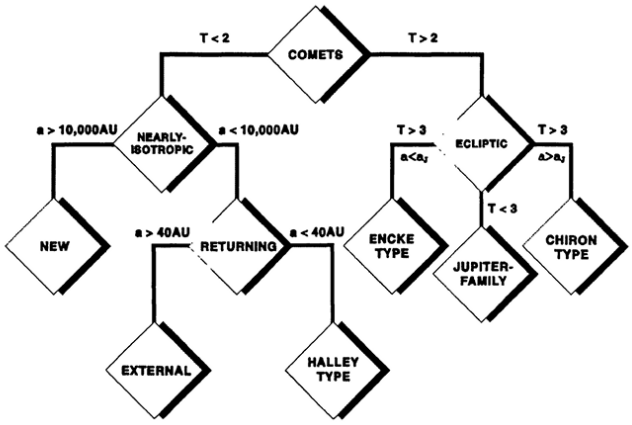


Figure 2.4: New family tree classifying comets using the Tisserand parameter. Source: Levison (1996) [27].

2.4 Anatomy of comets

The unique structure of comets has contributed both to their name and to their ability to instill curiosity throughout human history. Indeed, the appellation of "hairy star" stems from the fact that one observes the coma, which is the comet's atmosphere. Figure 2.5 illustrates the anatomy of a comet. The nucleus constitutes the heart of the comet and has a size only of the order of 10 km (Feldman, 1982 [28]). The nucleus is hidden by the surrounding coma which extends to much greater distances, although the extent of the coma depends on the chosen definition (Crifo & Rodionov, 1999 [29]). Tails are another important component of comets. Figure 2.5 presents two tails, the plasma tail and the dust tail. The former is driven by the solar wind; it is narrow and points in the direction opposite to the Sun. The molecules constituting the plasma tail have been ionised by the Sun's UV radiation and move at high velocities, causing the plasma tail to be almost rectilinear. The blueish colour of this tail arises from CO^+ emission at 4270 Å. The white dust

tail is formed due to the ejection of dust grains from the nucleus when volatiles sublimate. The competing forces acting upon the ejected dust include the gravitational force from the Sun versus the radiation pressure from the Sun. Radiation pressure develops as a result of the sunlight hitting dust particles and transferring momentum to them. Dust particles come in different sizes and masses and therefore have different trajectories, causing the dust tail to be broad and curved. For heavy dust particles, the gravitational force dominates over the radiation pressure. Consequently, these dust particles remain on the orbit of the comet, dispersing progressively. When Earth passes through the trail of heavy grains left behind by orbiting comets, shooting stars appear in the sky. The hydrogen envelope is a cloud surrounding the coma that is invisible to the naked eye. It extends up to 100 Mkm from the nucleus in the case of Comet Hale-Bopp. Hydrogen is released when ultraviolet light separates the water molecules of comets [25].

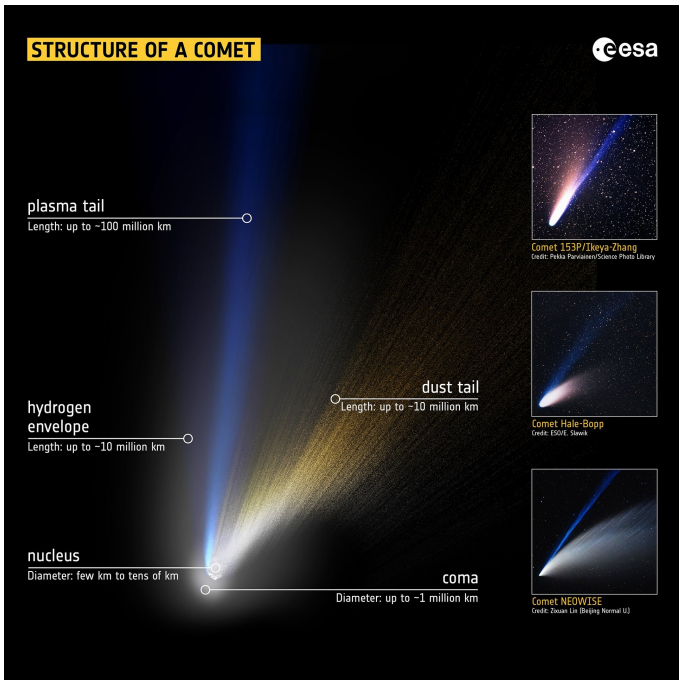


Figure 2.5: Structure of a comet. Source: ESA [30].

2.5 Composition of comets

In situ measurements performed by space missions, for example Giotto's visit to Comet 1P/Halley ([31]), have allowed a deeper understanding of the composition of comets. Fred Whipple, an American astronomer who discovered six comets, introduced the theory of the "dirty snowball" to describe the nucleus of Comet 2P/Encke [32]. Indeed, in Whipple's theory, the nucleus is made of ice composed of volatile materials that traps dust within. Each time the comet passes near its perihelion, the surface ice strips away as it sublimates, leaving surface dust behind. This explains the low albedo of comets. Missions visiting Comet 1P/Halley have reinforced Whipple's theory, measuring an albedo of only 4 %, comparable to that of charcoal. Whipple's theory was widely accepted until the Rosetta mission to 67P/Churyumoy-Gerasimenko demonstrated that this comet contains more dust than ice. It therefore seems that the ice/dust ratio may be unique for each comet [33].

The primordial molecular composition of the cometary nucleus can only be studied through fly-bys or rendez-vous missions [34]. These analyses yield information about the *parent molecules*, i.e. the molecules that have not yet been altered. Remote sensing measurements, as used in

this work, only allow to probe the coma and not the nucleus itself. The coma is composed of *daughter molecules* resulting from photo-processing of the volatile parent molecules. Nevertheless, the effectiveness of detecting daughter molecules through remote sensing and inferring the parent molecules has been proven [35].

Methods to probe the molecular complexity of comets include spectroscopy in the visible, ultraviolet, infrared, and radio domains, as well as mass spectroscopy, chromatography, sample return, and micrometeorite collection [35]. This master's thesis is based on spectroscopy measurements, a more detailed description of which is given in Section 3. To outline the common molecules detected within comets, Table 1 provides a non-exhaustive list of molecules that have been detected in comets with ground-based measurements, where water is by far the most abundant molecule.

Table 1: Molecules, radicals, ions, and atoms observed in comae detected using radio (rad), infrared (IR), visible (vis), ultraviolet (UV) spectroscopy or mass spectroscopy (MS). Table adapted from Crovisier (2005) [35].

	spectroscopy type		spectroscopy type
H_2O	rad, IR, MS	N	UV
H_2O+	vis	N_2+	vis
H_3O	rad	NH_3	rad, IR, MS
OH	rad, IR, UV	NH	vis
H	vis, UV	NH_2	IR, vis
H_2	UV	HCN	rad, IR, MS
O	vis, UV	HNC	rad
O+	UV	CN	rad, IR, vis
CO	rad, IR, UV, MS	CH_3CN	rad, MS
CO_2	IR, MS	HC_3N	rad
CO+	rad, vis, UV	HNCO	rad
CO_2+	vis	NH_2CHO	rad
C	UV	H_2S	rad, MS
C_2	IR, vis	CS	rad, UV
C_3	vis	SO	rad
CH_4	IR	SO_2	rad
C_2H_2	IR, MS	OCS	rad, IR
C_2H_4	MS	H_2CS	rad
C_2H_6	IR, MS	NS	rad
C_{42}	IR	S_2	UV
CH	IR, vis	S	UV
CH+	vis	Na	vis
CH_2	MS	K	vis
H_2CO	rad, vis	Ar	UV?
HCO+	rad		
CH_3OH	rad, IR, MS		
HCOOH	rad		
CH_3CHO	rad		
$HCOOCH_3$	rad		

More recently, a great number of new molecules have been identified thanks to mass spectroscopy performed during in situ measurements. Such rendez-vous missions include the Rosetta mission that has discovered the molecules presented in Figure 2.6.

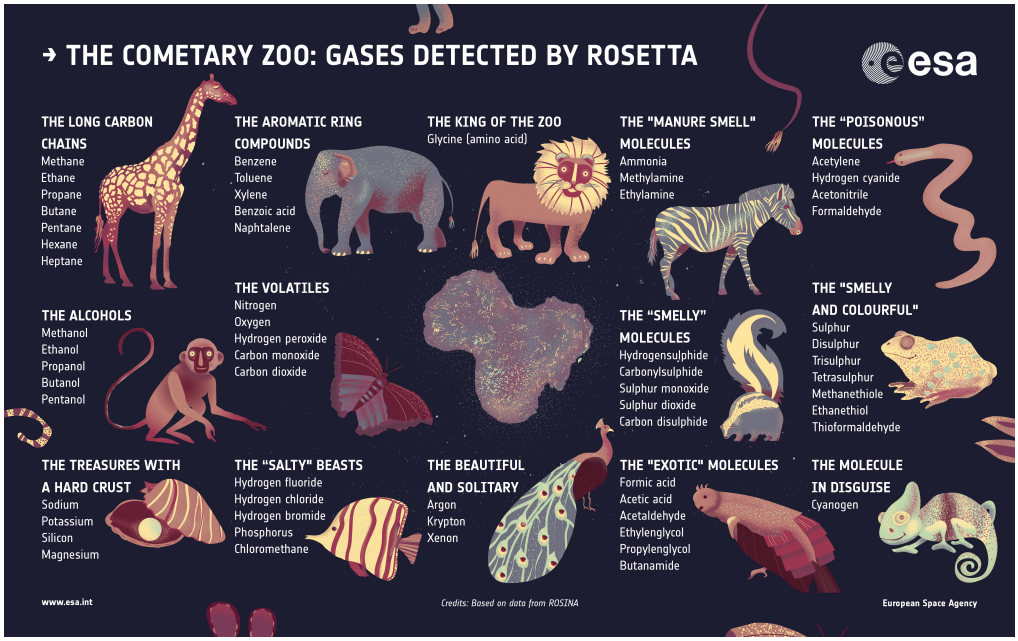


Figure 2.6: Gases detected in 67P/Churyumov–Gerasimenko by Rosetta. Source: [36].

2.6 Introducing the comets used in this work

The initial version of the Atlas developed by Hardy [1] is based on the spectrum of Comet C/2002 T7 (LINEAR). Because Hardy's work on the Atlas provides the foundation for this master's thesis, C2/2002 T7 remains the central comet of interest herein. It can, however, be fruitful to analyse spectra stemming from different comets in order to compare their emission lines. Indeed, different comets can present different compositions, exerting a major impact on their spectra. For this reason, in addition to the original centred spectrum of C2/2002 T7, spectra of comets C/2016 R2 (PanSTARRS), C/2002 X5 (Kudo-Fujikawa), 8P/Tuttle and C/2002 T7 (LINEAR), where the slit has been offset by 70″ and 245″ in the tail, are used.

Before describing the aforementioned comets, we present the observational details summarised in Table 2.

Table 2: Summary of the observations used. r is the heliocentric distance, $\dot{r}$ is the heliocentric velocity, Δ is the geocentric distance, $\dot{\Delta}$ is the geocentric velocity, and $Offset$ is the position of the slit relative to the centre of the comet.

Comet	*CalendarDate*	r(au)	$\dot{r}$(km/s)	Δ(au)	$\dot{\Delta}$(km/s)	$Offset('')$
C/2002 T7 (LINEAR)	2004-05-06	0.680	15.800	0.609	-65.600	5
	2004-05-06	0.680	15.800	0.609	-65.600	5
	2004-05-27	0.955	25.740	0.447	57.400	70
	2004-05-28	0.970	25.900	0.479	59.000	70
	2004-05-28	0.970	25.900	0.481	59.300	70
	2004-06-10	1.170	26.840	0.953	64.450	70
	2004-06-11	1.184	26.850	0.990	64.260	245
	2004-06-12	1.184	26.850	0.992	64.380	245
	2004-06-12	1.200	26.870	1.028	64.210	245
	2004-06-13	1.215	26.870	1.065	64.030	245
C/2016 R2 (PanSTARRS)	2018-02-14	2.75	-5.9	2.43	20.1	0
	2018-02-16	2.75	-5.8	2.46	20.2	0
C/2002 X5 (Kudo-Fujikawa)	2003-02-19	0.70	43.0	0.86	-5.0	0
	2003-03-07	1.06	37.0	0.99	29.4	3
8P/Tuttle	2008-01-16:2008-02-04	1.03	0.0	0.50	23.0	0

C/2002 T7 (LINEAR):
On 14 October 2002, the LINEAR (LINcoln Near-Earth Asteroid Reserch) (Stokes *et al.*, 1998 [37]) project discovered an asteroidal object which turned out to be Comet C/2002 T7 (LINEAR) (T7 hereafter), as announced by the IAU (International Astronomical Union) on 29 October 2002. Comet T7, shown in Figure 2.7, has a hyperbolic orbit and reached its perihelion in April 2004 [38]. It came close to the Sun with a perihelion distance of 0.615 au which allowed the ULiège team to get spectra of this very active comet with a very high signal to noise ratio, much better than the deVico spectrum (Cochran and Cochran 2002).

Figure 2.7: C/2002 T7 (LINEAR). Image by Víctor Ángel Buso, Gustavo Mazalán and Mariani Ascheri taken on 27 April 2004. Source: [38].

C/2016 R2 (PANSTARRS):
C/2016 R2 (16R2 hereafter), also called the blue comet, was discovered by the PANSTARRS (Panoramic Survey Telescope And Rapid Response System) (Kaiser *et al.*, 2002 [39]) telescope on 7 September 2016. Its deep blue tail stemming from the emission of ionised carbon monoxide (CO^+) is shown in Figure 2.8. In addition to being a very CO-rich comet, 16R2 has been found to be N_2-rich, suggesting that it belongs to a group of comets that we rarely get an opportunity to observe: those that formed beyond the N_2 ice line (Biver *et al.*, 2018 [40]). Moreover, Opitom *et al.* (2019, [41]) have shown that its composition is peculiar with usual molecules much less abundant than in other comets allowing to observe other lines rarely seen in typical comets. Its uncommon composition makes 16R2 a valuable asset in the identification of new lines.

Figure 2.8: C/2016 R2 (PANSTARRS), Observations taken on 18 January 2018. Source: ESO/SPECULOOS Team/E. Jehin, M. Kornmesse.

C/2002 X5 (Kudo-Fujikawa):
On 13 December 2002, Kudo discovered a comet that was independently discovered by Fujikawa the next day. Following the IAU nomenclature, this comet was named after their discoverers: C/2002 X5 (Kudo-Fujikawa) (X5 hereafter). It had a parabolic orbit and reached its perihelion, only 0.19 au from the Sun, on 29 January 2003. Figure 2.9 shows X5 two days before reaching its perihelion.

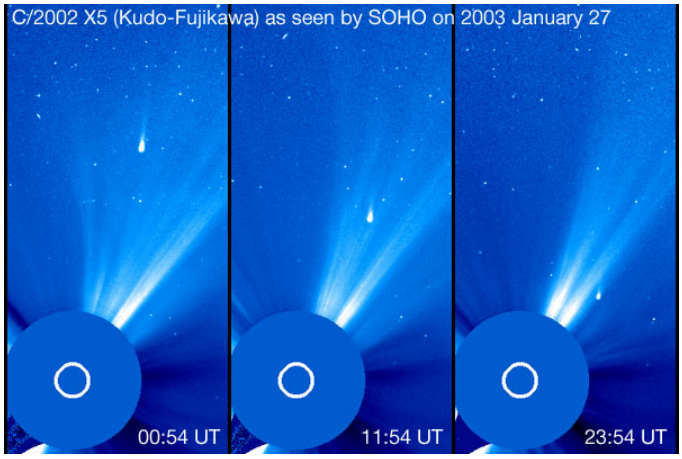


Figure 2.9: C/2002 X5 (Kudo-Fujikawa) near its perihelion taken on 27 January 2003. Source: SOHO (SOlar and Heliospheric Observatory).

8P/Tuttle:
P.F.A. Méchain was the first to observe 8P/Tuttle, shown in Figure 2.10, in January 1790 and several other astronomers were also able to observe it in the following month. H.P. Tuttle rediscovered the comet in 1858 and his name was later attributed to the comet. 8P/Tuttle has a short period of 13.6 years which explains the "P" in its name, standing for periodic. In 2010, using radar imagery, Harmon *et al.* [42] discovered that 8P/Tuttle is a contact binary, meaning that it is bilobate. 8P/Tuttle is responsible of the Ursid meteor shower happening in December [38].

Figure 2.10: 8P/Tuttle. Image by Hunter Wilson taken on 26 Decembre 2007 . Source: Sky&Telescope [43].

Table 3 summarises additional characteristics of the four comets by presenting the date of discovery and some orbital information.

Table 3: Date of discovery, type of comet, date of perihelion and perihelion distance for the four comets introduced: T7, 16R2, X5, and 8P/Tuttle. Source: Cometography [38], Small-Body Database (JPL, NASA) [44], The Catalogue of Near-Parabolic Comets (Centrum Badań Kosmicznych PAN) [45].

Comet	Date of discovery	Type of comet	Date of perihelion	Perihelion distance (au)
C/2002 T7 (LINEAR)	2002-10-14	NEW	2004-04-23	0.615
C/2016 R2 (PANSTARRS)	2016-09-07	EXT	2018-05-09	2.595
C/2002 X5 (Kudo-Fujikawa)	2002-12-18	EXT	2003-01-29	0.190
8P/Tuttle	1790-01-09	HFC	2021-08-28	1.027

Initially named after the year they appeared, comets also often bore the name of their discoverer starting in the mid-18th century. In the 20th century, they began to receive a provisional designation comprising the year of their discovery and a lower-case letter indicating the order of discovery within that year. Approximately two years later, they were assigned a definitive designation, which included the year of their perihelion passage and a serial number in Roman numerals. In 1995, this system was replaced by new rules for comet designation, introduced by the International Astronomical Union and is still in force.

Having introduced comets, particularly those significant to this study, it is now essential to explain how comets can be examined. Spectroscopy is a key method for unravelling the physics and chemistry of comets. The following section presents spectroscopy, laying the foundation for understanding the spectra used in this Atlas.

3 Spectroscopy

3.1 General Concepts

Consider the definition from Sinha *et al.* (2023, [46]), that "Spectroscopy is the study of the interaction between matter and electromagnetic radiation". Spectroscopy allows many physical phenomena to be studied and is frequently used in astrophysics. Indeed, light interacts with matter through absorption, emission or scattering, allowing scientific information to be captured. Because each chemical compound has its own spectral signature, scientists are then able to retrace the origin of the light received. Light has a dual nature: it is both particulate and wavelike. However, to describe specific phenomena, it is often easier to view light as either a wave or a particle. Because spectroscopy considers the interaction of light with matter, it is most convenient to consider it as a particle, the photon. The energy of a photon is calculated as follows:

$$E = h\nu \tag{3}$$

where h is Planck's constant and ν is the spatial frequency of the associated wave. This equation demonstrates that light composed of a single frequency, i.e. monochromatic light, is composed of photons with the same energy. This fundamental principle is important for the rest of this discussion (Penner, 2017 [47]).

Analysing light can be achieved by examining its electromagnetic spectrum, wherein light is decomposed into its constituent wavelengths or frequencies. This spectrum can contain either dark or bright lines, known as absorption or emission lines. Joseph von Fraunhofer was the first to detect absorption lines in the Sun's spectrum, caused by elements in the Sun's atmosphere absorbing specific wavelengths of light. Subsequently, Robert Bunsen and Gustav Kirchhoff produced emission lines by heating elements, discovering that each element generates a unique set of emission lines in the spectrum.

Both emission and absorption processes are rooted in quantum mechanics. This work focuses only on the bright lines created by emission processes; the subsequent sections will describe these processes in detail, first for atoms and then for molecules.

3.2 Atomic Transitions

The fact that emission lines provide the signature of certain chemical compounds is due to energy quantification, leading to discrete energy levels for the electron. The energy levels of an atom are calculated using the Schrödinger's equation. For the simplest atom, Hydrogen, the energy levels calculated as follows:

$$E_n = -\frac{\mu e^4}{8h^2\epsilon_0^2 n} = -\frac{R}{n^2} \tag{4}$$

where $\mu = \frac{m_e m_p}{m_e + m_p}$ is the reduced mass, e is the electric charge of an electron, h is the Planck constant, ϵ_0 is the vacuum permittivity, and R is the Rydberg unit of energy (-13.6eV). n is a positive integer and is termed the principal quantum number; it labels the energy level of the electron, where $n = 1$ is the ground state. To reach an energy level of $n > 1$, the atom must gain energy. With increasing n, the difference in energy between two subsequent levels decreases as shown in Figure 3.1, meaning that the atom has to gain less and less energy to jump up one level.

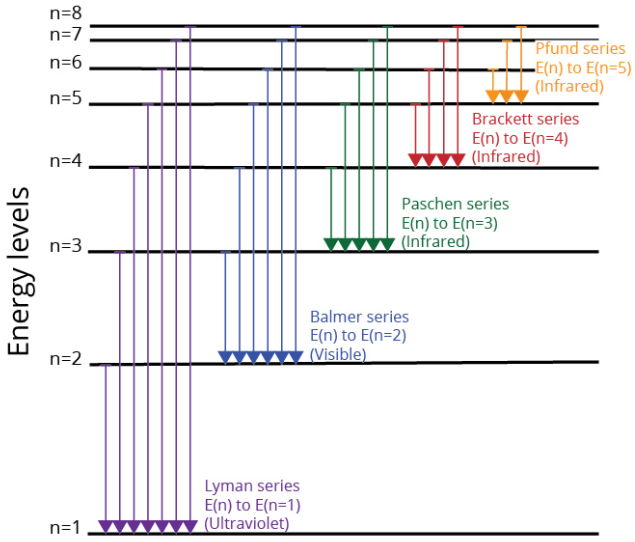


Figure 3.1: Energy level transition diagram for Hydrogen. Source: www.visionlearning.com.

An electron can jump to either a higher or lower level by receiving or releasing energy, respectively. When an electron transitions to a lower level, it releases energy in the form of a photon. If this photon is captured by a spectrometer, it will produce a bright emission line. Considering the example of the Hydrogen atom, if the electron jumps from level $n = 2$ to the ground state $n = 1$ (illustrated by the leftmost purple arrow in Figure 3.1), the photon that will be released through this process will have an energy of E_2-E_1 as follows:

$$E_2 - E_1 = -\frac{R}{4} + \frac{R}{1} = \frac{3R}{4} \approx 10.2eV \tag{5}$$

The energy of a photon from Equation 3 can then be re-expressed as follows:

$$E = h\nu = h\frac{c}{\lambda} \approx 10.2eV \tag{6}$$

where λ is the wavelength of the light, and c is the speed of the light in the vacuum. In spectroscopy , ν is also termed the *spectroscopic wavenumber* and one will either use the wavenumber ν or the wavelength λ to express the energy of a photon.

The energy released when an atom transitions from a higher to a lower level is unique to that atom. This results in the emission of a photon with a specific wavelength characteristic of that atom. These unique wavelengths create emission lines that serve as a fingerprint for each atom or molecule, allowing their identification in a spectrum. The discussion above was based on the Hydrogen atom because it is the simplest case, where the energy levels are well known. The general discussion still stands for atoms having more than one electron, however, the Schrödinger equation becomes substantially more complex, as the potential felt by the electron not only depends on the nucleus but also on the neighbouring electrons.

To precisely characterise the location of the electron in the atom, one uses the four quantum numbers:

- n, the principle quantum number introduced above to describe the main energy level, which is a positive integer starting at 1 ($n = 1, 2, 3, ...$);
- l, the angular momentum quantum number describing the orbital shape, which takes integer values between 0 and $n - 1$ (e.g. if $n = 3$, l can take the values $2, 1$ or 0);

- m_l, the magnetic quantum number describing the orbital orientation, which takes integer values from $-l$ to $+l$ (if $l = 1$, then $m_l = -1, 0, +1$); and
- m_s, the spin magnetic quantum number describing the projection of the spin of the electron along the z-axis, which can either be $+\frac{1}{2}$ or $+\frac{1}{2}$.

The above quantum numbers describe the position of a single electron. To determine the electronic structure of an atom with more than one electron, the total quantum number is used. Consider an atom with x electrons. The total quantum numbers, relevant for spectroscopy, associated with this atom are as follows:

- L, the total orbital angular momentum quantum number, which can take values from $|l_1 - l_2 - ... - l_x|$ to $|l_1 + l_2 + ... + l_x|$;
- S, the total intrinsic spin quantum number, which can vary from $|s_1 - s_2 - ... - s_x|$ to $|s_1 + s_2 + ... + s_x|$ where s is the spin of one electron; and
- J, the total angular momentum quantum number combining the total orbital angular momentum quantum number and the total spin angular momentum , which ranges from $|L-S|$ to $|L + S|$.

Other total quantum numbers, including the total magnetic quantum number and the total spin magnetic quantum number, also exist but that will not be necessary for the following of the explanation (Herzberg, 1950 [48]).

The spectroscopic term is a symbol that describes the electronic arrangement and can be separated into three parts:

1. the principal symbol L corresponding to the total orbital angular momentum, which is assigned to a letter depending on its value: S, P, D, F, G, H, I, ... for L=0, 1, 2, 3, 4, 5, 6, ...;
2. the multiplicity $2S+1$ corresponding to the spin combinations, the value of which determines whether the state is a singlet, doublet, triplet, ...; and
3. the total angular momentum J.

Combining these three terms in a specific way, one can obtain the spectroscopic term ${}^{2S+1}L_J$ [49].

Not all transitions from one energy level to another are allowed. So-called *selection rules* that depend on the transition moment and are defined through the eigenfunctions of the energy states which we will not discuss here, also apply. If this transition moment differs from 0, the transition is allowed; however, if the transition moment is equal to 0, the transition is forbidden. Fortunately, more practical rules have been established to determine whether or not a transition is allowed, without having to compute the transition moment. The selection rules are as follows:

- $\Delta J = J'' - J' = 0, \pm 1$, but $J'' = 0 \not\to J' = 0$
- $\Delta L = L'' - L' = 0, \pm 1$
- $\Delta S = S'' - S' = 0, \pm 1$

where $\not\to$ means "does not combine with". The notation $'$ refers to the upper level whereas $''$ refers to the lower level. Note that these rules apply to the electric dipole transitions, knowing that these transitions are not the only possible ones. Magnetic dipole and quadrupole transitions are 10^5 and 10^8 times less probable than electric dipole transitions; therefore, transitions violating the selection rules are still possible, although much less probable (Herzberg, 1950 [48]).

3.3 Molecular Transitions

Atoms are the fundamental building blocks of molecules; it is therefore essential to first understand the transitions and emission lines in atoms before examining such phenomena in molecules. Complexity increases progressively from a single-electron atom to a multi-electron atom, a diatomic molecule, and then a polyatomic molecule. The electronic transitions observed in atoms also occur in molecules; however, the interaction with other atoms and electrons alters the electronic energy levels. The principle remains the same: electrons can move up or down energy levels by absorbing or releasing energy. In addition to electronic transitions, molecules exhibit rotational and vibrational transitions, which involve the movement of nuclei.

A common simplification used in the calculations of molecular wavefunctions is the *Born-Oppenheimer approximation*, which treats the motion of electrons and nuclei separately. This is justified because electrons are much lighter and move much faster than nuclei. Consequently, heavy nuclei move so slowly that they do not sense the rapid movement of electrons but instead perceive the average field that they create. Another useful approximation is to consider the vibrations and rotations of the nuclei as independent of each other.

The energy associated with each type of transition can be classified as $E_{elec} >> E_{vib} >> E_{rot}$, and the total energy of the molecule is the sum of these three contributions. This hierarchy helps in understanding the structure of the transition bands: electronic transitions define the general structure, vibrational transitions the coarse structure, and rotational transitions the fine structure. This concept is outlines in Figure 3.2 and the different components will be explained below.

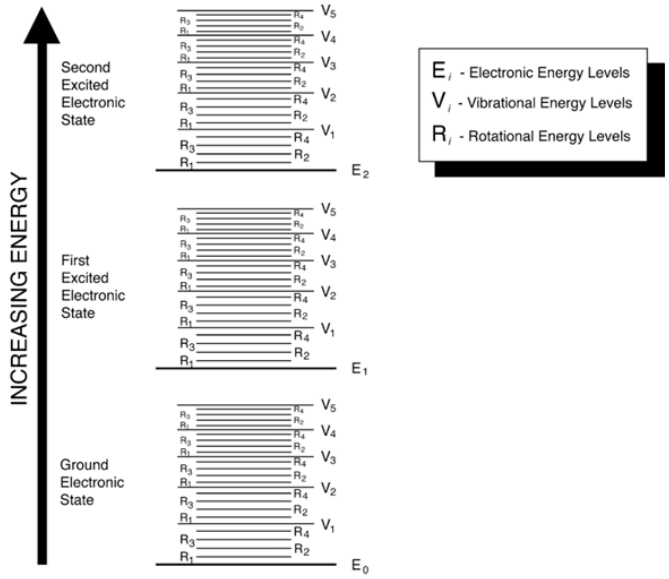


Figure 3.2: Molecular energy level diagram. E_i describes the general structure, V_i the coarse structure, and R_i the fine structure [5]. Source: Penner, 2017 [47].

Rotation spectrum:
The rotation of a molecule can induce emissions in the radio and far-infrared (IR) domains.

[5] The notation in Figure 3.2 does not correspond to the notation in the text. $E_i = E_{elec}$, $V_i = E_{vib}$, and $R_i = E_{rot}$.

The following discussion holds for heteronuclear diatomic molecules. In the case of homonuclear diatomic molecules, no radiation is produced through rotation because the centre of mass of the molecule is perfectly centred between the two atoms, resulting in no net dipole moment change during rotation.

To describe the rotational energy of a diatomic molecule, the rigid-rotor model is often used. According to this model, the rotational energy can be expressed as follows:

$$E_{rot} = J(J+1)\frac{h^2}{8\pi^2 I} \tag{7}$$

where I is the moment of inertia and J is the rotational quantum number, which is a non-negative integer (0, 1, 2, ...). This indicates that the rotational energy is quantised, indicating that only certain discrete energy levels exist.

The rotational energy is frequently expressed in terms of wavenumbers (cm^{-1}) as:

$$F(J) = \frac{E_{rot}}{hc} = BJ(J+1) \tag{8}$$

where B is the rotational constant. Analysing the transition probability reveals that the selection rule for rotational transitions is $\Delta J = J' - J'' = \pm 1$. This implies that transitions occur between consecutive levels, and the energy difference between two consecutive levels is determined as follows:

$$F(J+1) - F(J) = B(J+1)(J+2) - BJ(J-1) = 2B(J+1) \tag{9}$$

The difference between two consecutive rotational energy levels increases by $2B$ with each level, as illustrated in Figure 3.3. This figure also shows the spectrum (in wavenumbers) resulting from rotational transitions. This spectrum is characterised by an equal spacing of $2B$ between each line, such that a purely rotational spectrum is easily identified. Such spectra can be useful to characterise the emitting molecule: by measuring the line spacing, one can determine the rotational constant B, which depends on the inertial moment $I = \mu r_e^2$, where μ is the reduced mass and r_e is the internuclear distance, which is of interest to characterise the molecule.

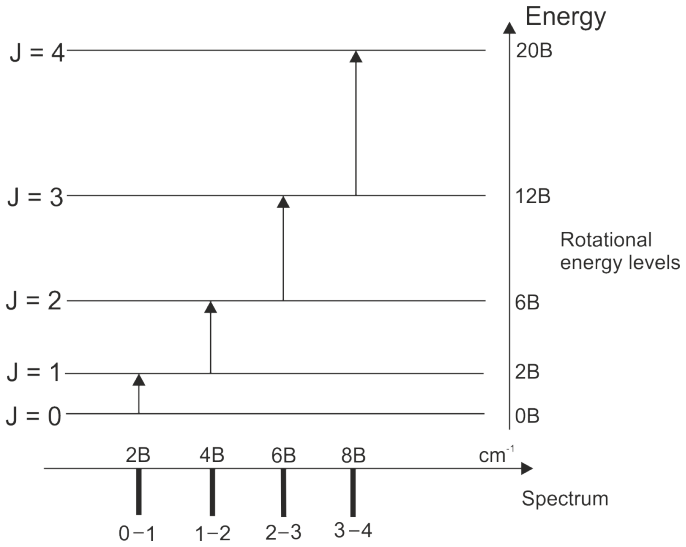


Figure 3.3: Rotational energy levels and transitions, and the resulting spectrum. Source: www.chemtube3d.com.

Although the rigid-rotor model provides a useful approximation, it does not perfectly represent reality. A more rigorous description involves considering non-rigid rotations, which introduce

effects such as vibrational stretching and centrifugal distortion. These effects tend to reduce the rotational constant as the internuclear distance increases because $B \propto \frac{1}{r_e^2}$. Consequently, the spacing between rotational energy levels is slightly reduced (Hanson, 2018 [50]).

Vibrational spectrum:
Vibrations in diatomic molecules are often simplified using the model of a simple harmonic oscillator. In this model, two atoms are linked by a spring and can oscillate around their equilibrium position. Using classical mechanics, the potential energy of such a system is determined as follows:

$$E_{vib} = \frac{1}{2}k(r - r_e)^2 \tag{10}$$

where k is the spring constant, r is the internuclear distance, and r_e is the internuclear distance at equilibrium. When introducing quantum mechanics, the vibrational energy G (in cm^{-1}) for a vibrational quantum number v (where $v = 0, 1, 2, ...$) is expressed as follows:

$$G = \frac{E_{vib}}{hc}\omega_e(v + \frac{1}{2})) \tag{11}$$

where $\omega_e = \frac{\sqrt{\frac{k}{\mu}}}{2\pi c}$ is the harmonic frequency around equilibrium in cm^{-1}. Consequently, the vibrational energy levels are discrete and the selection rule $\Delta v = v' - v'' = +1$ only allows the transition to adjacent levels. The lowest energy level for $v = 0$ is $G_0 = \frac{1}{2}\omega_e$, and the spacing between two adjacent levels is as follows:

$$G_{v+1} - G_v = \omega_e(v + 1 + \frac{1}{2}) - \omega_e(v + \frac{1}{2}) = \omega_e \tag{12}$$

In contrast to the rotational energy levels, the spacing between the vibrational energy levels remains constant with increasing vibrational quantum numbers, as shown by the dashed horizontal lines in Figure 3.4.

However, similarly to the rigid rotor model, the simple harmonic oscillator model does not accurately represent reality. The anharmonic oscillator provides a more realistic depiction and is characterised by additional terms in G known as anharmonic corrections. These anharmonic correction terms break the selection rule $\Delta v = v'' - v' = +1$ established for the harmonic oscillator such that transitions of $\Delta v = +2, 3, 4, ...$ are allowed, but their probability decreases with increasing Δv. Indeed, the correction terms become smaller as the order of the correction increases, leading to smaller probabilities. Similar to the rotational transitions, the addition of correction terms reduces the difference between vibrational energy levels, as indicated by the red horizontal lines in Figure 3.4. The Morse potential shown in Figure 3.4 more accurately depicts reality, because it accounts for the dissociation energy D_e above which the molecule breaks apart.

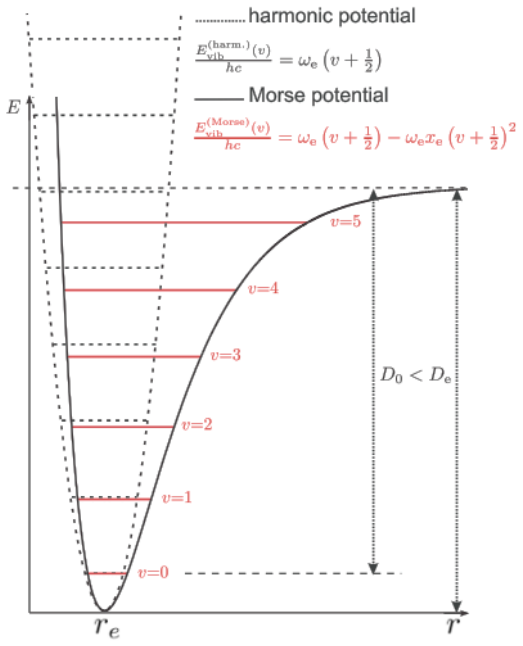


Figure 3.4: Potentials for the harmonic oscillator (dashed line) and the anharmonic oscillator (Morse potential) (solid line). Vibrational energy levels are indicated by the horizontal lines (dashed or red). Figure adapted from Sanli, 2017 [51].

Roto-vibrational spectrum:
When combining rotational and vibrational motions while applying the Born-Oppenheimer approximation, these motions are considered independent. The outcome is a vibrating rigid rotor. The selection rule of the vibrating rigid rotor is a combination of the selection rules for the rotational and vibrational transitions, specifically $\Delta J = \pm 1$ and $\Delta v = +1$. Note that although $\Delta J = 0$ is possible, it is most often observed in electronic transitions and not in pure roto-vibrational transitions. It is possible to have a $\Delta J = 0$ transition in a pure roto-vibrational transition if the vibrational modes carry angular momentum; however, this is not the case for most diatomic molecules. Transitions with $\Delta J = +1$ are termed *R branch*, transitions with $\Delta J = -1$ are termed *P branch*, and transitions with $\Delta J = 0$ are termed *Q branch*. The energy of the roto-vibrational level is given by the sum of rotational energy level F (Equation 8) and the vibrational energy level G (Equation 11), as follows:

$$T(v, J) = F(J) + G(v) = BJ(J+1) + \omega_e(v + \frac{1}{2}) \tag{13}$$

The line positions of the roto-vibrational transitions, obeying the selection rules are thus:

$$\begin{aligned} \mu &= T(v', J') - T(v'', J'') \\ &= \left[BJ'(J'+1) + \omega_e\left(v' + \frac{1}{2}\right)\right] - \left[BJ''(J''+1) + \omega_e\left(v'' + \frac{1}{2}\right)\right] \end{aligned} \tag{14}$$

For the R branch ($\Delta J = +1$) and $\Delta v = +1$:

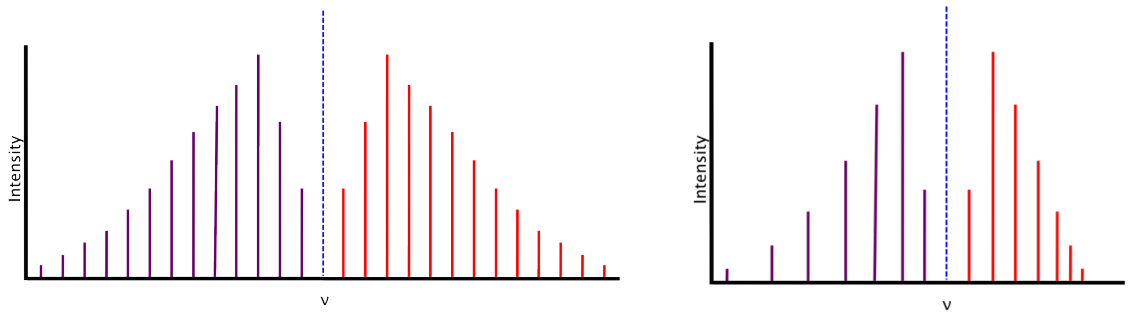
$$\begin{aligned} \mu &= \left[B(J''+1)(J''+2) + \omega_e\left(v''+1+\frac{1}{2}\right)\right] - \left[BJ''(J''+1) + \omega_e\left(v''+\frac{1}{2}\right)\right] \\ &= 2B(J''+1) + \omega_e \end{aligned} \tag{15}$$

For the P branch ($\Delta J = -1$) and $\Delta v = 1$:

$$\begin{aligned}\mu &= \left[B(J''-1)J'' + \omega_e\left(v''+1+\frac{1}{2}\right)\right] - \left[BJ''(J''+1) + \omega_e\left(v''+\frac{1}{2}\right)\right] \\ &= -2BJ'' + \omega_e\end{aligned} \tag{16}$$

This leads to a line separation of $2B$ in the roto-vibrational spectrum, as it was already the case for the rigid rotor (Fig. 3.3). The line intensity corresponds to the transition probabilities, which depend on the probabilities of the molecules being in the considered level. The line intensities are distributed as shown in Figure 3.5a. The R branch is at higher frequencies because the energy jump from the upper level to the lower level is higher for the R branch than the P branch. The sum of all the lines in Figure 3.5a is termed *band* and corresponds to one particular change in vibrational number e.g. $\Delta v = +1$.

As indicated earlier, the vibrating rigid rotor is only a simplified model. In reality, there is also non-rigid rotation, anharmonic vibration, and roto-vibrational interaction, leading to additional terms in the expression of $T(v, J)$. These additional terms cause an increase in the line spacing in the P branch and a decrease in the line spacing on the R branch. In the spectrum of Figure 3.5b, this manifests as a spreading of the P branch and a "squishing" of the R branch [50].



(a) Roto-vibrational spectrum of a band using the vibrating rigid-rotor approximation.

(b) Real case of a roto-vibrational spectrum of a band.

Figure 3.5: The P branch is shown in purple, the R branch in red, and the Q branch in blue. Source: https://chem.libretexts.org/.

Spectroscopic notation for molecules:
The spectroscopic term symbol for molecules is similar to that used one for atoms; however, the naming conventions differ. Furthermore, new quantum numbers are introduced because more numbers are needed to uniquely define the more complex orbitals of molecules (parity and reflection).

- Λ is the projection along the internuclear axis of the total orbital angular momentum number (projection of L along the internuclear axis):

$$\Lambda = \sum_i \lambda_i \tag{17}$$

  where λ_i =0 for σ electrons and ± 1 fro π electrons. It is written in Greek capital letters, depending on its value:

  - $|\Lambda| = 0 \rightarrow \Sigma$
  - $|\Lambda| = 1 \rightarrow \Pi$
  - $|\Lambda| = 2 \rightarrow \Delta$

  - $|\Lambda| = 3 \rightarrow \Phi$

- Ω is the projection of the total angular momentum number along the internuclear distance;
- S is the total spin angular momentum quantum number. $2S + 1$ is called the multiplicity, and according to its value, the state is called singlet, doublet, or triplet, ... state;
- u/g denotes the parity of the orbital of homonuclear diatomic molecules; meaning whether it is symmetric or anti-symmetric for an inversion operator. g (for gerade) stands for a symmetric orbital and u (for ungerade) stands for an anti-symmetric orbital; and
- $+/-$ indicates whether the orbital is symmetric or anti-symmetric for a reflection along a plane that passes through both atoms. Reflection only applies to Σ states ($\Lambda = 0$).

The molecular spectroscopic term is finally given by

$$^{2S+1}|\Lambda|^{(+/-)}_{(g/u)} \tag{18}$$

A letter is usually placed before the spectroscopic term and refers to the energy level of the electronic state. X is used for the electronic ground state and A, B, C, ... for the first, second, third, ... excited level if the excited states have the same multiplicity as the ground state. If they have a different multiplicity, lowercase letters are used, i.e. a, b, c, ... (Ochkin, 2009 [52], and [49]).

Molecular Selection Rules:
The selection rules for molecules are a combination of the roto-vibrational rules and additional rules related to electronic transitions [53]:

- no rule for Δv (outside the harmonic oscillator model) but decreasing probabilities with increasing Δv;
- $\Delta J = \pm 1$ with the possibility of ΔJ=0 if two different electronic or vibrational states are involved;
- ΔS=0, only allowing transitions with the same spin multiplicity $2S + 1$;
- $g \rightarrow u$ or $u \rightarrow g$ for homonuclear diatomic molecules, known as Laporte's rule, allowing only transitions that change parity; and
- $+ \longleftrightarrow +$ or $- \longleftrightarrow -$ for transitions in the Σ state.

When dealing with molecular transitions, it is possible to encounter different notation conventions. One common way to write transition is as follows:

$$B - X \qquad 2 - 1 \qquad rR2(25.5) \tag{19}$$

The three columns correspond to an electronic term, a vibrational term, and a rotational term. In each column, it is usual to first write the initial state and then the final state.

- In the given example, B is the initial electronic state and corresponds to the second excited state, and X is the final electronic state, corresponding to the ground state.
- Similarly, the second column indicates a transition from the vibrational level $v' = 2$ to $v'' = 1$.
- The last column should be written as $^{r}R_{2(25.5)}$ but the transition data are often written in simple .txt files. To avoid any incompatibility, it is more common to use the simple notation $rR2(25.5)$. While the notation for the first two columns is standard, the notation for this third term can vary depending on the preferences of the author. Therefore, it is prudent to check the specific convention used in each source. This is particularly important when dealing with older sources, which may use outdated notation conventions.

3.4 Fluorescence

This introduction to spectroscopy explains the emergence of the emission lines captured in the spectra of comets. Moreover, it serves as the theoretical foundation to my model of NH presented later. In addition, we would like to dedicate this last section specifically to fluorescence as part of the theoretical introductory work of the fluorescence model.

Fluorescence is part of the family of processes called *luminescence*. Luminescence is the phenomenon through which light is emitted due to electronically excited states. Phosphorescence and fluorescence fall into this category. The difference between the two lies in the multiplicity of excited states: is it a singlet or a triplet? The transition from a singlet excited state to the ground state occurs rapidly, on the order of 10ns. This rapidity is due to the pairing between the electron in the excited orbital and the electron in the ground orbital through opposite spins, allowing this transition. In the case of phosphorescence, the excited state is a triplet, in which case the electron in the excited orbital and the electron in the ground orbital have the same spin. The transition is therefore forbidden and takes from milliseconds to seconds to occur. Phosphorescence transitions may take more time, for example in glow-in-the-dark objects. In some rare cases, it is possible to observe mixed singlet-triplet states, in which case the transition time is somewhere between hundreds of nanoseconds to a few microseconds. The name phosphorescence stems from the Greek word for light $\varphi\alpha o\zeta$ (pháos); however, the name fluorescence might be a bit more surprising as fluor is not fluorescent. George Stokes introduced this term while studying the light emitted from the mineral *fluorspar*; however, fluorspar does not get its name from the element fluor, but rather from the Latin words *fluo*, meaning flow, and *spar*, meaning non-metallic mineral (Valeur, 2001 [54]).

Jablonski diagrams, such as that shown in Figure 3.6, provide useful representations for understanding luminescence. The purple and blue lines in Figure 3.6 indicate the absorption of light to jump from the ground to an excited state. After absorption, there can either be an internal conversion, whereby the molecules relax to S_1 in 10^{-12} s. From this point, the molecules can either de-excite to the ground state via fluorescence (green line) or switch to a triplet T_1 state by spin conversion. More time is required to de-excite from T_1; this is called phosphorescence (red line). Phosphorescence is facilitated when the molecules contain heavy atoms because the transition from S_1 to T_1 is more probable. The Jablonski diagram shows that fluorescence gives rise to higher frequency photons compared to phosphorescence, hence the colour green for the more energetic transition and red for the less energetic one. Indeed, the gap between S_1 and S_0 is greater than that between T_1 and S_0 and, as described above, a larger energy gap indicated a more energetic electron (Lakowicz, 2006 [55]).

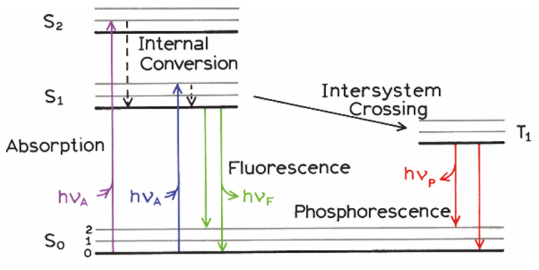


Figure 3.6: Jablonski diagram. S_0, S_1, and S_2 denote the singlet electronic states. The vibrational energy levels are indicated by the finer lines and the number 0, 1, 2. T_1 is the first triplet state. Source: Lakowicz, 2006 [55].

This concludes the theoretical part about how emission lines are created in atoms and molecules.

The next section will establish how to capture the emission lines using spectrographs and what information can be extracted from the resulting cometary spectra.

4 Methods

The methods to tackle the objectives of this thesis outlined in Section 1.2 are detailed below. First, we will discuss cometary spectra, which serve as the foundation for cometary atlases since the process of line identification relies on the peaks within these spectra. We took advantage of the diverse range of spectra taken by the ULiège team to detect and identify new cometary lines. The search and identification of new lines is highly eased by our Atlas presented in Section 1.1, which regroups the findings in one place. The second component of the methodology involves developing a fluorescence model. Indeed, the Atlas only provides information on the location of the molecular lines. However, a model also determines the intensity of the lines, creating the opportunity to compute the abundance of the molecule.

4.1 Cometary spectra

The spectra produced by comets can provide valuable information about their coma composition. The higher the resolution of a spectrum, the more detail it offers. For this reason, the high-resolution spectrum of the bright Comet T7 obtained with UVES was used as the foundation of the Atlas. We will give an overview of the history of spectroscopy and illustrate that the nature of unidentified lines has been an open question for over a century. We will then discuss the UVES instrument and explain how its data can be processed to obtain the clean spectra that we are using to make the Atlas. Finally, we will highlight some of the spectra's characteristics.

4.1.1 The emergence of cometary spectra

As explained in the section dedicated to spectroscopy, electromagnetic radiation is decomposed into different wavelengths or frequencies. The first recorded observation of a spectrum was made by Isaac Newton who, in 1666, demonstrated that when light passes through a prism, it is dispersed into its constituent colours. He documented his findings in his book *Opticks*, published in 1704 [56]. This work became the foundation of spectroscopy and opened a pathway for future discoveries in the field. In 1814, Fraunhofer detected absorption lines in the spectrum of the Sun. Fifty years later, in 1864, Giovanni Battista Donati observed the first spectrum from a comet. He detected broad emission bands in the spectrum of Comet C/1864 N1 Tempel, which are labelled γ, β, and α in his drawing, shown in Figure 4.1. Donati pioneered in the field of cometary spectroscopy and he correctly predicted in the same year that "by comparing the spectra of different comets between them and with the spectra of other light sources, we should be able to know more on the still mysterious nature of these objects" (Donati, 1864). This work is what we try to continue in this master's thesis 160 years later as those comet spectra still keep some mysteries.

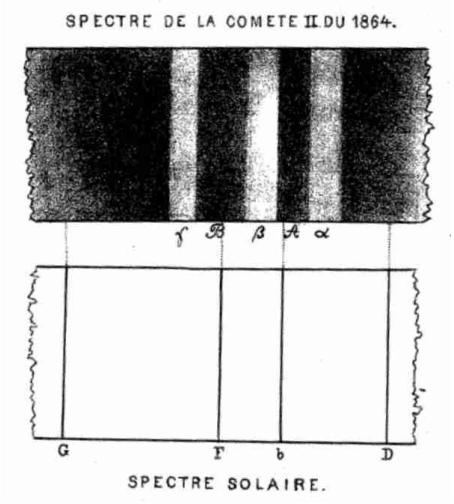


Figure 4.1: Drawing of the first cometary spectrum by Donati in 1864 (top), compared with the spectrum of the Sun (bottom). Source: James Lequeux, taken from [17].

In 1868, William Huggins was able to relate the emission lines observed by Donati to the C_2 diatomic molecule. He discovered the Fraunhofer absorption lines in a cometary spectrum, confirming the idea that the light from the comet was reflected sunlight on dust particles mixed with the comet's own emission. In the following years, many new cometary spectra were observed and new laboratory measurements were conducted, allowing researchers to identify new molecules in optical spectra of comets, like the highly poisonous CN radical. Although Johannes Kepler had described the structural features of comets, including the coma and tail, as early as 1625, the materials forming these components remained unknown. It was not until cometary spectroscopy had matured that it was discovered that the coma and principal tail are made of dust grains, responsible for the reflected sunlight, as well as gas molecules.

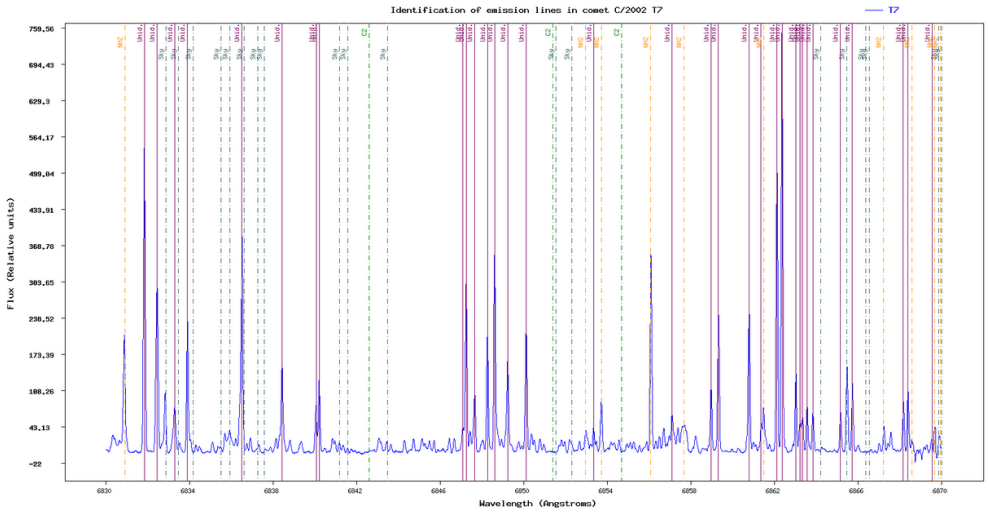
Spectre de la comète Morehouse (1908 *c*).

Int.		Description	Identification
1	7027	Tête	Non identifiée.
1	6848	Tête	Non identifiée.
3	6254	Tête avec queue	C 3e gr. nég. 6245-6196 (C, + 2).
2	6024	Tête diffuse, large	C I, 6191-5958.
1	6020	Tête avec queue	C 3e gr. nég. 6021-5976 (B, + 3) (?).
2	5900	Tête avec queue	C 3e gr. nég. 5906-5862 (D, 0).
1	5563	Tête avec trace de queue	C II, 5635-5470.
1	5482	Tête avec trace de queue	C 3e gr. nég. 5504-5467 (B, + 2).
1	5369	Tête très faible	Non identifiée.
1	5260	Tête avec queue	C 3e gr. nég. 5248-5214 (C, 0) (?).
2	5163 à 5107	Tête diffuse	C III, 5165-5082.
3	5107	Tête avec queue	C 3e gr. nég. 5076 (B, + 1).
3	5021	Tête avec queue	C 3e gr. nég. 5044 (B, + 1).
1	4987	Tête avec trace de queue	Non identifiée.
1	492	Tête avec trace de queue	C 3e gr. nég. 4915-4884 (A, + 2).
3	4879	Tête avec queue	C 3e gr. nég. 4869 (C, — 1)
3	4846	Tête avec queue	C 3e gr. nég. 4840 (C, — 1)
1	4772	Tête avec trace de queue	Non identifiée.
8	4746 à 4637	Tête	C IV, 4737-4680.
2	4722 à 4523	Queue	Az β, 4709-4515.

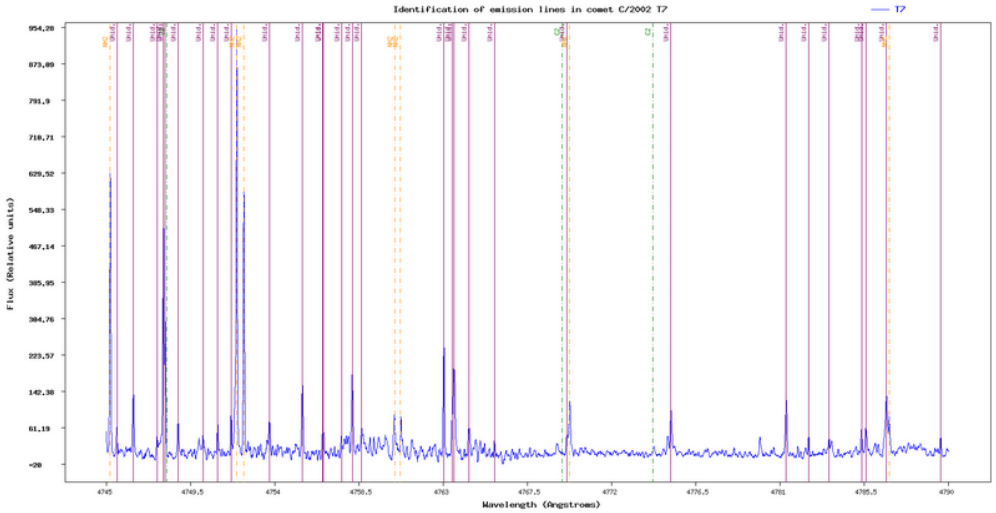
Figure 4.2: Emission lines of Comet Morehouse, as reported by de la Baume Pluvinel and Baldet in 1911 [57].

Scientists continued to capture cometary spectra of increasing quality because, as Donati indicated, one can gain valuable information by comparing the spectra of different comets. In 1908, de la Baume Pluvinel and Baldet [57] observed a spectrum Comet Morehouse (1908c). Figure 4.2 lists the emission lines they were able to detect. They primary species they observed was C_2,

one of the few identified molecules in comets at that time. Rather than examining the identified lines, it is more interesting to explore the unidentified ones. The authors were unable to classify 5 regions: 7027, 6848, 5369, 4987, and 4772 Å. Since the publication of this old article, a lot of progresses have been made and a variety of new molecules have been detected in comets, allowing the region at 7027 to be attributed to NH_2, a molecule that was yet to be detected at that time. The indication "tête", means that the emission band at 7027 Å was only detected in the spectrum of the inner coma (the head of the comet), and not in that of the tail. This is in agreement with NH_2, which is a neutral molecule known to fade very rapidly with increasing distance to the nucleus. The idea of comparing the line intensities for spectra taken at different positions in the comet will be addressed in more detail in Section 4.1.4. The regions 5369 and 4987 Å are associated with C_2 based on modern catalogues of lines. The identified lines in these zones are all found in our Atlas of T7. The two remaining regions, at 6848 and 4772 Å, however, still remain unidentified in our Atlas, 115 years after their detection. The two unidentified regions are plotted in Figure 4.3. In Figure 4.3a, a bundle of unidentified lines is located at 6848 Å, precisely where de la Baume Pluvinel and Baldet reported their emission band. Apart from this bundle, other rays in this region are also yet to be identified. The other region at 4772 Å turns out to be heavily populated with unidentified lines in our spectra, as indicated by Figure 4.3b. This example highlights that, despite the significant progress already achieved, for instance the identification of atomic FeI and NiI lines in all the comets observed by the ULiège team with UVES, and even for comets far from the Sun in cold environment (Manfroid *et al.*, 2021 [58]), many questions remain to be answered.



(a) Unidentified lines in 6848 Å in the spectrum of T7. Plot from the Atlas.



(b) Unidentified lines in 4772 Å in the spectrum of T7. Plot from the Atlas.

Figure 4.3: Two regions that were labelled as unidentified in 1911 in the spectrum of Comet Morehouse and are still unidentified today in the Atlas.

4.1.2 Spectrographs

Having discussed the first cometary spectra and their development over the years, it is now time to address how these optical spectra are recorded. In 1814, Fraunhofer placed a prism in front of the lens of his telescope to discover the absorption lines from the Sun, building the first spectroscope. A spectroscope is a device that spreads incoming light into its constituent wavelengths. Since 1814, the technology has advanced substantially, leading to the instrument that recorded the spectra of T7, 16R2, X5, and 8P/Tuttle called the Ultraviolet and Visual Echelle Spectrograph (UVES) (Dekker *et al.*, 2000 [59]). A spectrograph is a spectroscope combined with a recording device [60].

The most important components of a spectrograph are a slit, a grating, and a camera. The slit only lets light from a certain region of the comet pass. This light hits a grating that decomposes

the light into wavelengths. The decomposed beam is then captured by a detector. UVES is based on this principle but its design is heavily refined and optimised. It is installed at the Very Large Telescope (VLT) on Unit Telescope 2. The layout of UVES is shown in Figure 4.4. It is a two-arm echelle spectrometer, with one arm dedicated to the blue region of the spectrum (300–500 nm) and the other to the red region (420–1100 nm). For legibility, the principal components described above are highlighted in red and blue for each arm, respectively. The components are explained as follows, in order of the sequence in which light entering the telescope interacts with them:

- The first element is the beaming slit. As there is a separate slit for each arm, the "arm selector" redirects the light to the chosen slit. The light can either be fed to the red arm, the blue arm, or both simultaneously.
- Then, the collimator ensures that the light rays that reach the grating are all parallel. The grating is a component with grooves that diffract the light similarly to how a prism refracts light. The grating equation is used to compute the diffraction angle:

$$n\lambda = d(\sin\theta + \sin\theta') \tag{20}$$

  where d is the distance between the grooves, θ is the incidence angle, θ' is the diffraction angle, and n is the diffraction order. "Echelle" gratings are those in which the grooves are far apart and the "blaze angle" θ, the incident angle at which the efficiency is maximal, is extremely high. From Equation 20, it can be seen that since the groove distance d and the incident angle θ are large, the energy is concentrated at high orders n. This results in high angular separation between wavelengths. Since the goal of UVES is to produce high-resolution spectra, echelle gratings are ideal. The price to pay for this high resolution is the reduced maximal spectral bandwidth, the "free spectral range", which is given by

$$\lambda_2 - \lambda_1 = \frac{\lambda_1}{n} \tag{21}$$

  where λ_1 and λ_2 are the lower and upper limits of the free spectral range. Due to the large values of n, the difference $\lambda_2 - \lambda_1$ is small (Thorlabs, [61]).
- Finally, the diffracted light is captured by one charge-coupled device (CCD) camera for the blue arm and two for the red arm.

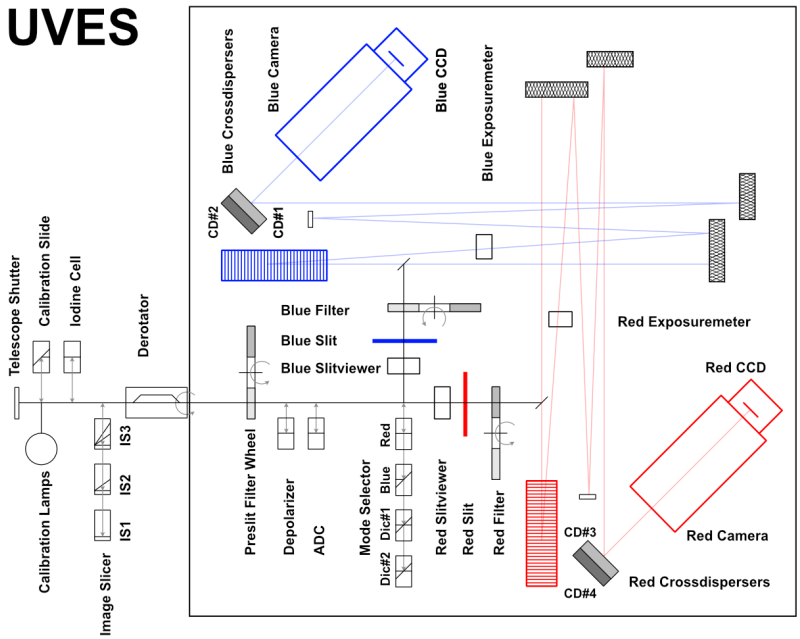


Figure 4.4: Overview of UVES. The thin light red and light blue lines indicate the paths of the light in the red and blue arms. The red and blue instruments are the slit, echelle grating, and CCD for the red and blue arm. Figure adapted from the UVES user manual [62].

4.1.3 Data reduction and preprocessing

The data reduction and preprocessing methods presented in this section were all completed by others before I received the data. For completeness, however, I will discuss the principal steps that have been implemented.

The UVES setup provides raw images of the comet, which then need to be reduced using the UVES data reduction pipeline, as described in the UVES data reduction cookbook [63]. A short description of the main points follows. The raw data retrieved from the CCD cameras are two-dimensional spectra like the one illustrated in Figure 4.5. The echelle orders n run vertically, and the wavelength increases from top to bottom and from left to right. In addition to the CCD data, images needed for calibration are taken to correct for instrumental effects and other noise sources. These correction images include a dark field, a flat field, and a bias field. The dark and bias field are subtracted from the raw file, and the result is divided by the flat field. This results in a reduced 2D spectrum that is then collapsed to obtain a more useful 1D spectrum. The first step is to generate a 1D spectrum in counts per pixel position and then to generate the final 1D spectrum of intensity versus wavelength. For the first step, all the counts from one CCD row are combined, resulting in a 1D spectrum of counts per row. In the second step, the pixels need to be assigned to the correct wavelengths. For this purpose, lamps with narrow emission lines are used. The exact positions of the lines are known from laboratory measurements. From there, it is possible to attribute each pixel to a specific wavelength and obtain the spectrum of intensity versus wavelength. UVES uses thorium–argon (ThAr) lamps for wavelength calibrations [64].

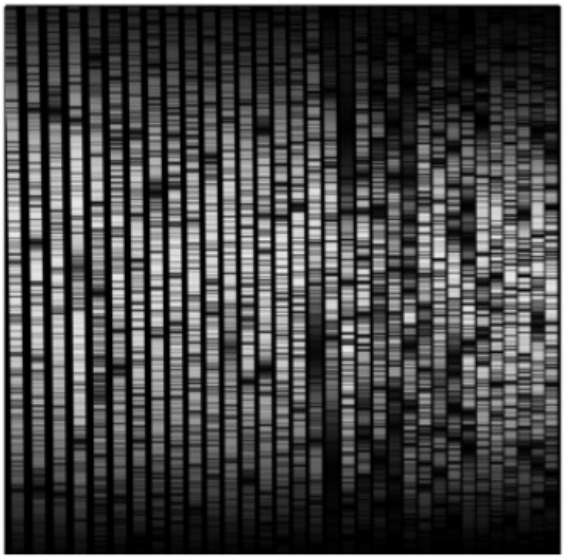


Figure 4.5: Example of a 2D echelle spectrum. The orders move from bottom to top. Source: [65]

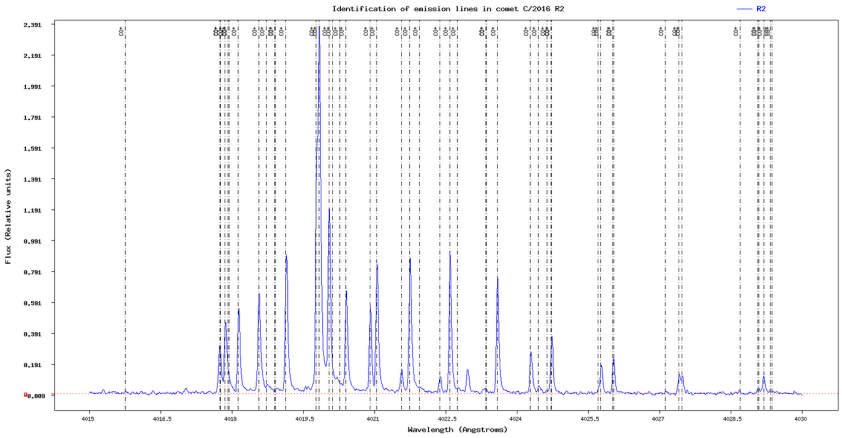
When a comet is observed, its motion relative to Earth causes its spectrum to undergo a Doppler shift. If the comet is approaching Earth, the spectrum shifts towards the blue end of the spectrum (blueshift), and if the comet is moving away, it shifts towards the red end (redshift). To compare spectra to laboratory lists of lines to identify them, all spectra are adjusted to the Earth's rest frame by applying a shift that counters the comet's geocentric velocity.

The emission lines of the atoms and molecules of the comet's coma are the centre of interest, but they can be masked by the continuum due to reflected sunlight on the dust particles that populate the coma. Further contamination may come from sunlight reflected from the Moon or from twilight. To counteract this undesirable part of the spectrum, the Sun's continuum is subtracted, leaving only the cometary spectrum. This subtraction was done using Kurucz *et al.*'s (2005 [66]) spectrum of the Sun. Furthermore, the use of ground-based telescopes implies that the cometary light has to travel through the atmosphere, leading to additional *telluric* lines, either emission lines caused by the emission processes in the atmosphere or absorption lines caused by the molecules in the atmosphere. The lines also have to be corrected for, using, for example, the Molecfit software. This action requires careful identification of the emission lines in order to not confuse them with cometary lines.

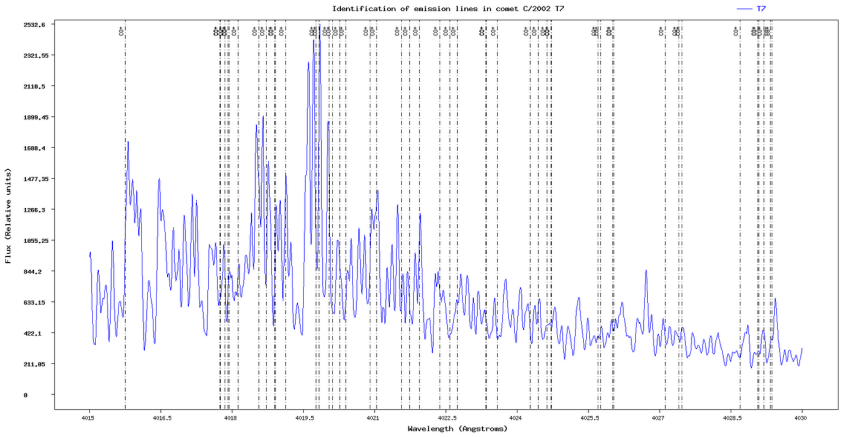
4.1.4 Noteworthy features of cometary spectra

Once the data reduction and the preprocessing have been completed, the spectra are ready for analysis. To achieve the results presented in Section 5, we made use of spectra from the four comets discussed in Section 2.6.

It has already been mentioned that the advantage of analysing and comparing emission spectra from several comets lies in their differing compositions: it can be easier to detect a molecule in a comet where that molecule is present at a higher concentration. Figure 4.6 illustrates this for the ion CO^+, known to be abundant in 16R2 [67]. Figure 4.6a shows the well-defined peaks of CO^+ in 16R2, whereas in Figure 4.6b it is hard to identify the CO^+ in Comet T7 because it is weak and hidden by lines from other species.



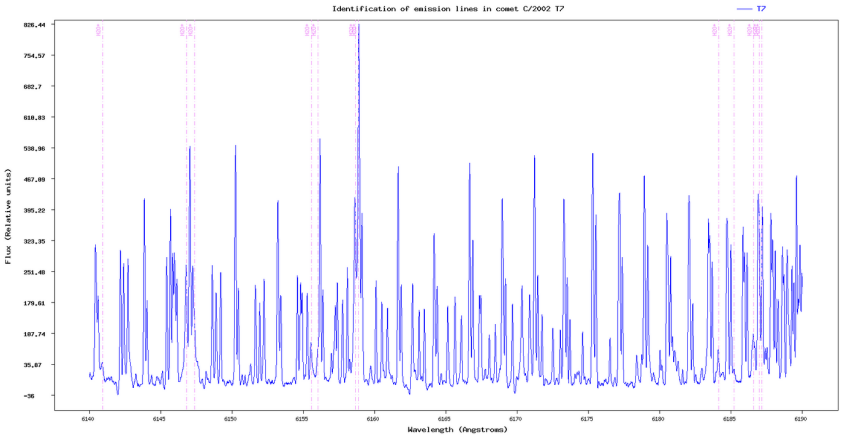
(a) Distinct emission CO^+ lines in R2.



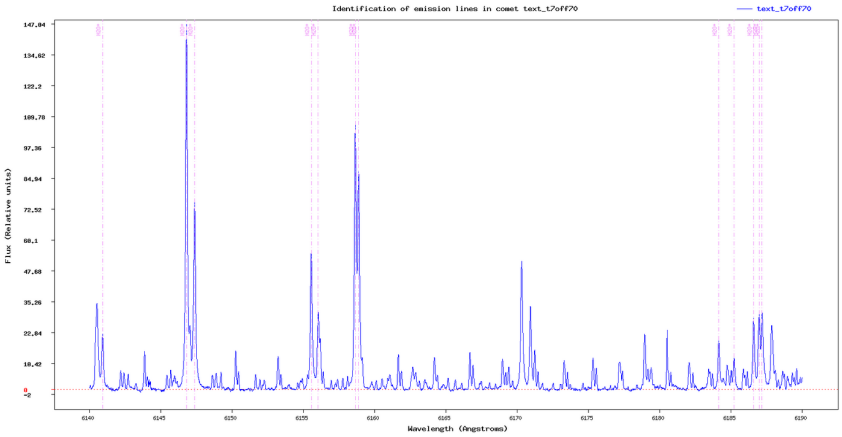
(b) Unclear emission CO^+ lines in T7.

Figure 4.6: Comparison between the emission lines of CO^+ in 16R2 and T7.

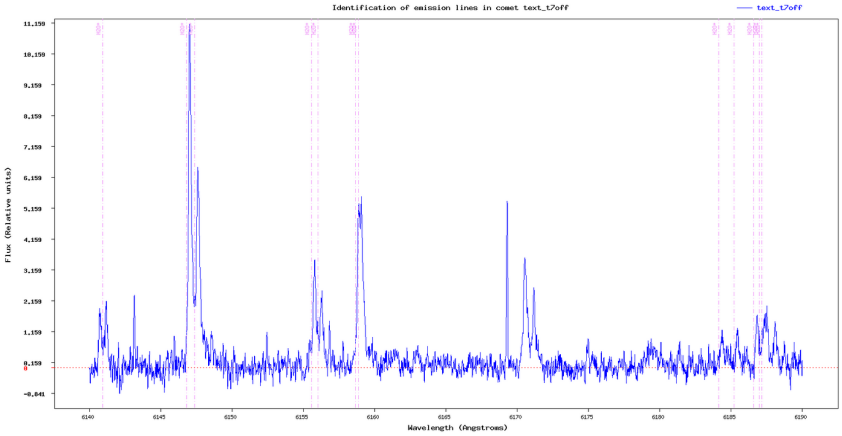
In Section 2.6, Table 2 displays three different slit positions for T7 indicated as "offset". The 5″ offset captures the inner coma, dominated by neutrals, and will be considered as the centred spectrum from here on. The reason for offsetting the slit by 5″ instead of perfectly centring it is to reduce the contamination by the solar spectrum caused by the reflection of the sunlight by dust particles. Indeed, a slight offset decreases significantly the dust continuum because it is very concentrated on the nucleus. The observations offset by 70″ and 245″ are aligned with the plasma tail, populated by ions. Comparing centred and offset spectra can provide information about the nature of the molecules: if the lines appear intense in the centred spectrum and dampened in the offset spectra, the molecule is probably neutral. In contrast, if they are weak in the centred spectrum and intense in the offset spectrum, the lines are likely caused by an ion. Figure 4.7 illustrates the case of an ion, here H_2O^+, whose emission lines do not appear clearly in the centred spectrum of T7, as shown in Figure 4.7a, because of the cluster of other lines in the region. The two offset spectra of T7 in Figures 4.7b and 4.7c showcase distinct lines of H_2O^+ present in the plasma tail.



(a) Crowded region in the centred spectrum of T7 . H_2O^+ lines are not clearly visible.



(b) The same region in the spectrum of T7 offset by 70″ in the tail direction. H_2O^+ lines are clearly visible.



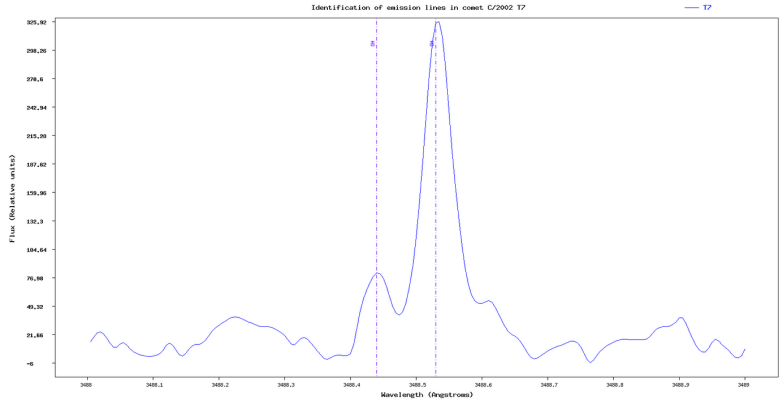
(c) The same region in the spectrum of T7 offset by 245″ in the tail direction. H_2O^+ lines are clearly visible.

Figure 4.7: Comparison of the emission lines of H_2O^+ (pink) between 6140 and 6190 Å in spectra of T7 that are centred, offset by 70″, or offset by 245″ in the tail direction.

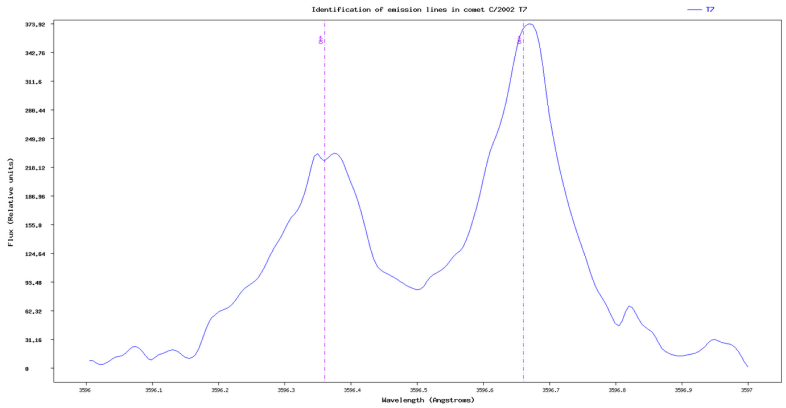
The emission lines of the H_2O^+ appearing in the spectrum are not perfect peaks but rather have a certain width. This line broadening can be caused by several phenomena, including Doppler, natural, and collisional broadening [68]:

- Doppler broadening occurs because the velocities of the molecules are not identical but follow a distribution. Molecules with slightly different velocities will produce emission lines that have a slightly different Doppler shift, causing lines to be broadened.
- Natural broadening is due to Heisenberg's uncertainty principle. Since the lifetime of an excited state is uncertain, the wavelength of the emitted photon is also uncertain.
- Collisional broadening occurs when the emitting molecule collides with another particle. This shortens the lifetime of the excited state and increases the uncertainty in the wavelength.

Ions have higher velocity than neutral molecules because they are pushed by the solar wind at about 400 km/s compared to 1 km/s for neutrals. Therefore, the Doppler broadening for ion lines is also stronger than that of the neutral lines. Broad emission lines in a spectrum can, therefore, hint towards ionic particles and help the identification process. Figure 4.8 shows two emission lines of OH and 2 emission lines of OH^+. Both show a 1 Å portion of the spectrum so that it is apparent that the ion indeed presents broader lines.



(a) OH emission lines in T7.



(b) OH^+ emission lines in T7.

Figure 4.8: Comparison of ionic and neutral emission lines, demonstrating the greater Doppler broadening observed for ions.

4.2 NH fluorescence model

As described, the second part of my master's thesis involves the development of a fluorescence model of the NH molecule. The purpose of this model is to not only determine the line position but also the line intensity of NH. In addition, the model also allows for a more theoretical approach in this work.

For the development of this model, I was welcomed by Prof. Philippe Rousselot from the Astro group at UTINAM[6] (Univers - Théorie - Interfaces - Nanostructures - Atmosphère et environnement - Molécules) at the Université de Franche-Compté from 4–15 March. During my two-week stay, Prof. Rousselot guided me through the various steps involved in creating a model. In addition to the development of the model, I was fortunate to learn more about the historical background of the Besançon Observatory, built in 1884 for the acclaimed clock fabrication in the city. Indeed, the observatory allowed to calibrate the clocks and watches [69].

The development of the fluorescence model can be divided into three parts: the computation of the transition probabilities, of the populations of the different levels, and of the line intensities. The detailed Python code of the model can be found in Appendix B.

4.2.1 Transition probabilities

The starting point for computing the transition probabilities was to obtain the energy of each considered level. The quantum numbers used here to describe the energy levels are N, the rotational angular momentum quantum number, J, the resultant angular momentum quantum number, F, the total angular momentum quantum number, and v, the vibrational quantum number. For this model, the energy levels for v' and v'' ranging from 0 to 2, for N' ranging from 1 to 10, and for N'' ranging from 0 to 10 were taken from Fernando *et al.* [70]. Considering the NH $^3\Pi_i$-$^3\Sigma^-$ transition for this model, there are three J values for each pair of v and N. In addition, the Π state has a splitting of each energy level due to the symmetry, f or e. In total, there are 93 levels for the Σ state and 180 levels for the Π state. Figure 4.9 illustrates the arrangement of the first energy levels as well as some transitions.

[6] https://www.utinam.cnrs.fr/

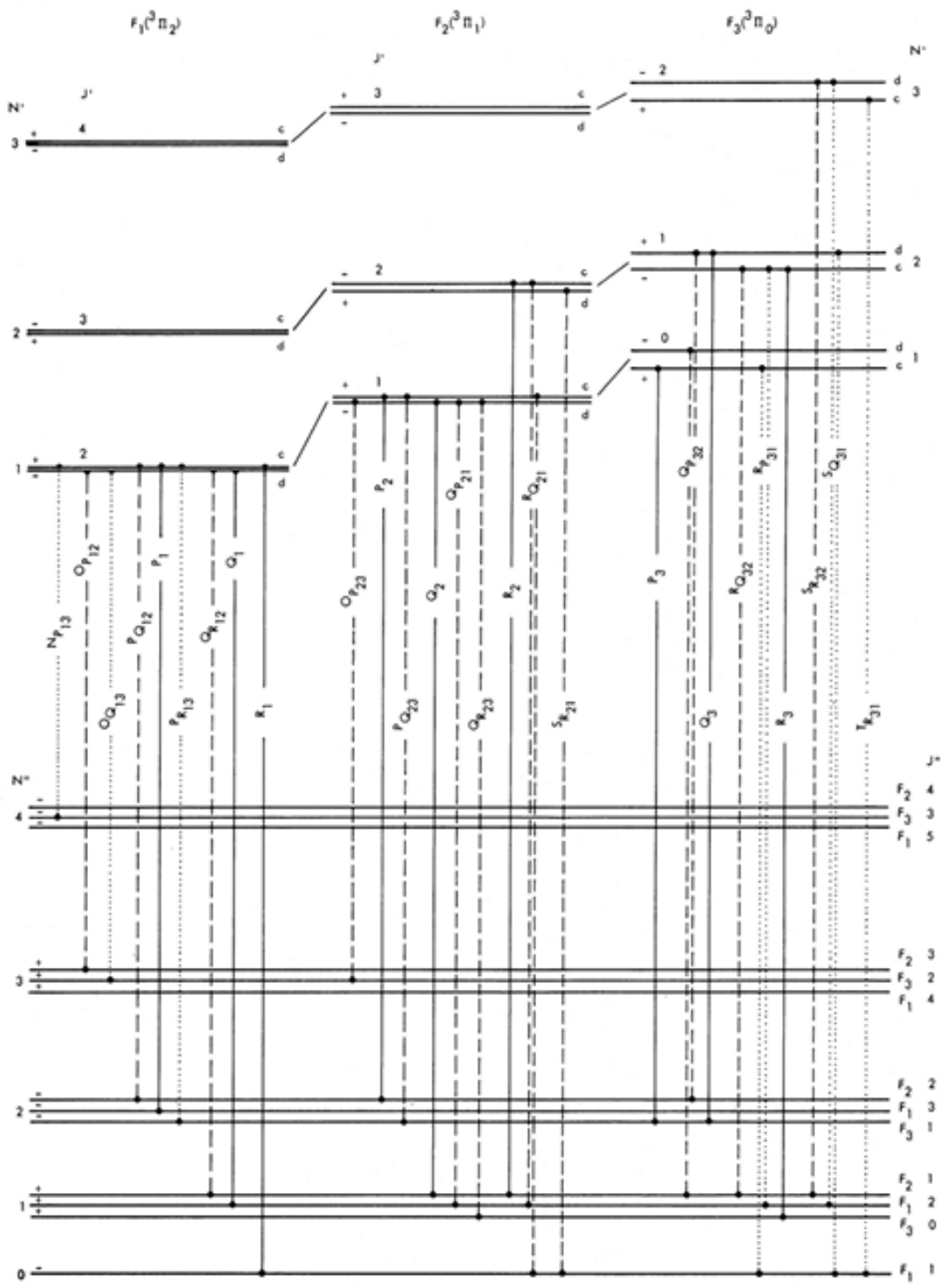


Figure 4.9: Energy level diagram and first lines of the NH $^3\Pi_i$-$^3\Sigma^-$ transition. Source: Litvak and Kuiper , 1981 [71].

The transition probabilities are given by A for emission and Bρ_ν for absorption, respectively ([72],

[48]), where A and B are the Einstein coefficients, and ρ_ν is the density of the solar spectrum at a given frequency ν. The formulae to compute A and B, given by Schleicher and A'Hearn (1982) [73], are as follows:

$$A_{ij} = \frac{2-\delta_{0,\Lambda'}}{2-\delta_{0,\Lambda'+\Lambda''}} \frac{\nu_{line}^3}{\nu_*^3} A_{v'v''} \frac{S_{J'J''}}{2J'+1} \quad (in \quad \mathrm{s}^{-1})$$

$$B_{ji} = \frac{1}{8\pi h} \frac{c^3}{\nu_{line}^3} \frac{2J'+1}{2J''+1} A_{ij} \quad (in \quad \mathrm{cm}^3\,\mathrm{Hz}\,\mathrm{erg}^{-1}\,\mathrm{s}^{-1})$$

where i and j represent the upper and lower states, respectively, δ is the Kronecker delta, Λ is the projection of the total orbital momentum along the nuclear axis, ν_{line} is the wavenumber of the considered transition, ν_* is the wavenumber of the band head[7], $A_{v'v''}$ is the transition probability of a given band $v'-v''$, $S_{J'J''}$ are the Hönl–London coefficients, J' is the total angular momentum of the initial state, h is the Planck constant expressed in erg Hz^{-1}, and c is the speed of light in vacuum.

For the case of a $^3\Pi$-$^3\Sigma^-$ transition, Λ' is 1 and Λ'' is 0. The first term in the definition of A_{ij} consequently reduces to

$$\frac{2-\delta_{0,1}}{2-\delta_{0,1}} = 1$$

The values for transition probabilities $A_{v'v''}$ were taken from Meier *et al.* (1998) [74].
The formulae for the Hönl–London coefficients $S_{J'J''}$ are given by Herzberg (1950) [48] and can be reduced in this case to the following:

- P line: $S_{J'J''} = \frac{3}{2}J'$
- Q line: $S_{J'J''} = \frac{3}{2}(2J'+1)$
- R line: $S_{J'J''} = \frac{3}{2}(J'+1)$

To obtain the absorption transition probabilities from a lower energy level to a higher one, the B_{ji} values need to be multiplied by ρ_ν, which is given by the high-resolution spectrum of Kurucz *et al.* (2005) [66]. The provided density is give in 10^{-20} erg cm^{-3} Hz^{-1}, which was measured for a heliocentric distance of 1au. Therefore, it needs to be divided by the square of the heliocentric distance of the comet (in au) at the time of the observation. Then, the wavelengths of the fluorescence model need to be interpolated to the solar spectrum. Before doing that, the model wavelengths need to be converted from wavelengths in vacuum λ_{vacuum} to wavelengths in the air λ_{air}. Additionally, the wavelengths of the solar spectrum need to be corrected for a Doppler shift caused by the heliocentric velocity of the comet v_h.These wavelength corrections can be written as

$$\lambda_{air\_model} = \frac{\lambda_{vacuum\_model}}{n_{air}}$$

$$\lambda_{air\_corrected\_solar} = \lambda_{air\_solar}(1 - \frac{v_h}{c})$$

The refraction index of the air n_{air} was computed using the formula from Morton (1991) [11], which is the IAU standard.

4.2.2 Population calculations

This model is based on fluorescence calculations published by Zucconi and Festou (1985) [12], who present a new set of equations. A peculiarity is that it consists of N_X equations, where

[7] The band head is the transition between the lowest energy level of the considered v' and the lowest energy level of the considered v''.

N_X represents the number of electronic sub-state ground levels, whereas in the past, a set of N equations was used, with N the total number of molecular levels. This consequently reduces the number of equations and makes them easier to solve. For this particular model for the NH molecule, N is 273, whereas N_X is 93. This simplification is justified because the excited levels have much shorter lifetimes than the ground levels. The formulae used in this model, taken from Zucconi and Festou (1985), are

$$\dot{n}_i = -n_i \sum_{j=1}^{N} p_{ij} + \sum_{j=1}^{N} n_j p_{ij} \quad i = 1, N \tag{22}$$

where p_{ij} are the probabilities to transition from an i level to a j level, computed in Section 4.2.1, and n_i is the population of the i^{th} level of the molecule. Equation 22 gives the temporal evolution of the i^{th} level and expresses between emissions and adsorptions from that level. This formula is often unsolvable because of the large value of N. Therefore, Zucconi and Festou established the following set of formulae:

$$T_{ij} = p_{ji} - \delta_{ij} \sum_{k=1}^{N} p_{ik} \quad i, j = 1, N \tag{23}$$

where δ_{ij} is the Kronecker delta. Here, $\mathbf{T}$ is an $N \times N$ matrix that can be decomposed into four sub-matrices,

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_{XX} & \mathbf{T}_{XE} \\ \mathbf{T}_{EX} & \mathbf{T}_{EE} \end{pmatrix} \begin{pmatrix} \mathbf{n}_X \\ \mathbf{n}_E \end{pmatrix} \tag{24}$$

where X refers to the electronic ground state Σ and E refers to the excited electronic state Π. Using this decomposition,

$$\begin{aligned} &\mathbf{x} = (n_1, ...n_{N_X - 1}) \\ &b_i = (\mathbf{T}_{XX} - \mathbf{T}_{XE}\mathbf{T}_{EE}^{-1}\mathbf{T}_{EX})_{iN_X} \quad i = 1, N_X - 1 \\ &A_{ij} = (\mathbf{T}_{XX} - \mathbf{T}_{XE}\mathbf{T}_{EE}^{-1}\mathbf{T}_{EX})_{ij} - b_i \quad j = 1, N_X - 1 \\ &\mathbf{x} = -\mathbf{A}^{-1}\mathbf{b} \end{aligned}$$

Here, $\mathbf{x}$ is the equilibrium solution, obtained when solving Equation 22 and considering the fluorescence equilibrium $\dot{n}_i = 0$, and contains the population of the $N_X - 1$ lowest levels of the ground state. The equilibrium assumption is justified if the integration time steps are much longer than the lifetimes of the levels.

The only population of the ground state that is not yet determined is the one with the highest energy n_{N_X}, which can be obtained via the conservation equation $\sum_{i=1}^{N_X} n_i = 1$. The population of the excited electronic levels n_i can be determined by incorporating relation 22 into the conservation, which gives

$$n_i = \frac{\sum_{j=1}^{N_X} n_j B_{ji} \rho_\nu}{\sum_{j=1}^{N_X} A_{ij}} \tag{25}$$

4.2.3 Intensities of the lines

Having computed the transition probabilities and the population of each level, the last step of the model is to compute the intensities I of each line which can be done using

$$I_{ij} = x_i p_{ij} \frac{hc}{\lambda_{vacuum}} \quad (in \quad erg\, molecule^{-1} s^{-1}) \tag{26}$$

Note that the position of a given line is given by it's wavelength in the air (see 4.2.1) while the intensity of the line is computed using the wavelength in the vacuum (if c is given in the vacuum). This is due to the fact that the transition energy gets released in the vacuum in space.

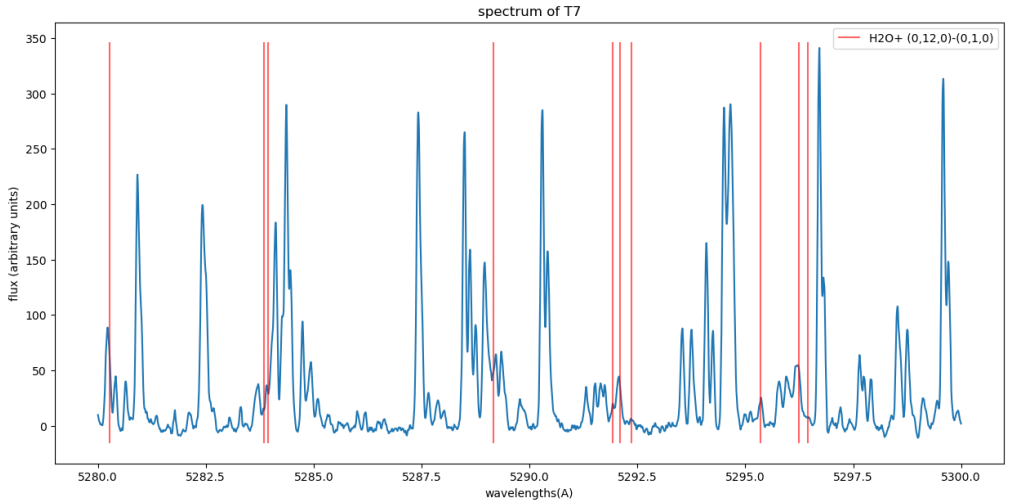
5 Results

5.1 Cometary Atlas

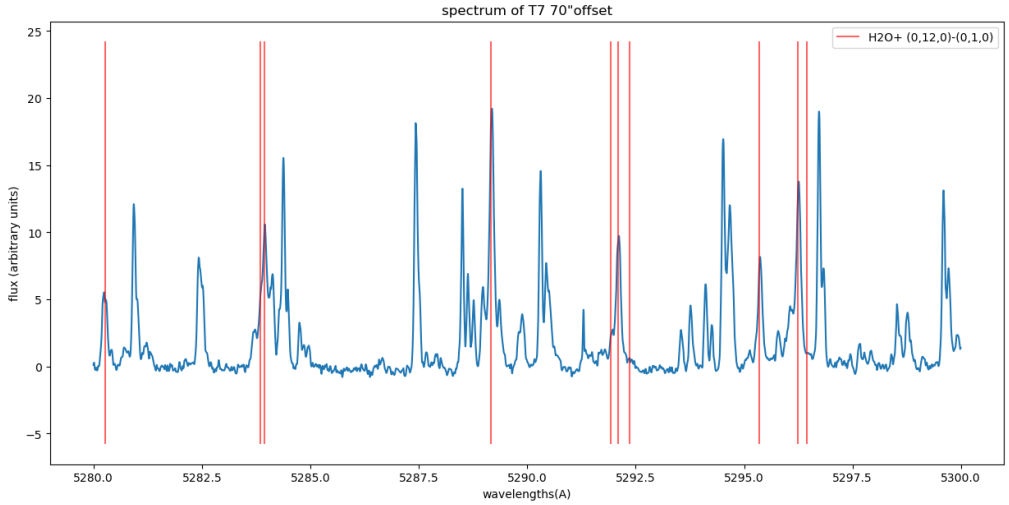
5.1.1 Oxoniumyl H_2O^+

Water is the by far the most dominant molecule in comets, so it comes to no surprise that H_2O^+ is present in most spectra and especially tail spectra. Hardy had already identified 77 H_2O^+ lines in 2022. These lines stem from the electronic transition $\tilde{A}^2A_1 - \tilde{X}^2B_1$ and the vibrational bands (0,3,0)-(0,0,0) and (0,2,0)-(0,0,0). Given the availability of other spectra for this work, including those of the tail of T7, it can be fruitful to look for additional bands in these tail spectra. In 1976, Lew [5] tabulated many lines belonging to the $\tilde{A}^2A_1 - \tilde{X}^2B_1$ transition. However, these lines are not accessible as downloadable data. We have transcribed the tables and made them available for anyone wanting to use Lew's work in Appendix A. In addition to the original data from 1976, we used the updated formula from Morton, 1991 [11] to provide recalculated wavelengths taken in the air λ_{air} *recalculated*. For layout reasons, only the recalculated wavelengths are displayed in the dark green column in the Appendix. The tables containing both the recalculated and original wavelengths can be accessed via the link in Appendix A.

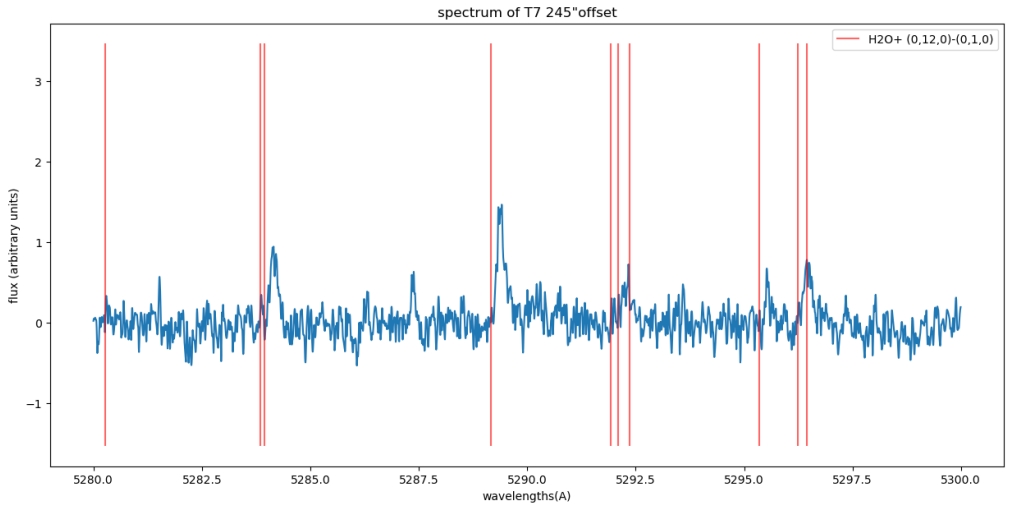
In the "High-Resolution Catalogue of Cometary Emission Lines" of Comet Swift-Tuttle and Brorsen-Metcalf published by Brown *et al.* in 1996 [75], they present a line list for H_2O^+ containing only one line for the band (0,12,0)-(0,1,0). This line at 5289.21Å perfectly matches a peak in the T7 offset tailward by 70″ spectrum. Upon investigation, other lines of the (0,12,0)-(0,1,0) band from Lew's table also match the spectrum, justifying the addition of this band in the Atlas. The (0,11,0)-(0,1,0) band was also added to the Atlas, showing satisfactory results in matching with the spectrum. In total, 55 new H_2O^+ lines were added to the Atlas but only two unidentified lines were removed from the new version of the Atlas. This low number is explained by the use of the centered spectrum of T7 for the detection of the unidentified lines. As Figure 5.1a shows, the H_2O^+ lines get drowned by the myriad of (neutral) lines in the spectrum. However, the ionic lines come forward in the spectra of the ionic tail shown in Figure 5.1b and their identification becomes possible.



(a)



(b)



(c)

Figure 5.1: Some of the H_2O^+ emission lines that were added to the Atlas. The lines are shown in 3 spectra of T7: the spectrum that is centered (top), offset by 70″ (middle), and offset by 245″ (bottom).

During the analysis of the H_2O^+ lines, it has been noticed that in Figure 4.7c (see Section 4.1.4), these lines appear slightly shifted to the right with respect to the expected position from laboratory tables of the H_2O^+ lines in pink. We believe that this position shift of the lines can be attributed to a Doppler shift with respect to Earth, not due to the movement of the comet, but due to the fast movement of the ions in the tail. In fact, the motion of the comet with respect to Earth has already been corrected for during the preprocessing, as explained. However, the ions in the tail have an additional velocity component with respect to the comet that has not been corrected for. It would be quite challenging to rectify this shift by taking the velocity of the ions into account because each ion has a unique velocity. It may be surprising that this effect is only visible in the spectrum offset by 245″, but this can be explained by the unusual configuration on the nights when the spectrum offset by 70″ was taken. During the nights around 28 April 2004, T7, Earth, and the Sun almost formed a right angle. Therefore, once the velocity of the comet is subtracted, the plasma tail pointing in the opposite direction of the Sun only has a perpendicular velocity component with respect to the Earth's rest frame. Since only radial velocity components cause a Doppler shift, there is no shift in the peaks of the spectrum offset by 70″. During the 15 days that separated the measurements of the 70″ and the 245″ offset spectra, the configuration changed significantly. T7, Earth, and the Sun no longer formed a 90-degree angle, as shown in Figure 5.2, and the plasma tail had a radial velocity component inducing a Doppler shift. The ions of the plasma tail moved away from the Sun, redshifting the spectrum to longer wavelengths (to the right in Figure 4.7c). A shift of 0.26 Å was measured at the H_2O^+ doublet at 6519 Å. The velocity of the ions projected to the Earth direction can then be estimated using the Doppler shift formula:

$$\begin{aligned} \frac{\Delta\lambda}{\lambda_0} &= \frac{v}{c} \\ \Leftrightarrow v &= \frac{0.26}{6159} c \\ &\approx 12.664\,\mathrm{km/s} \end{aligned}$$

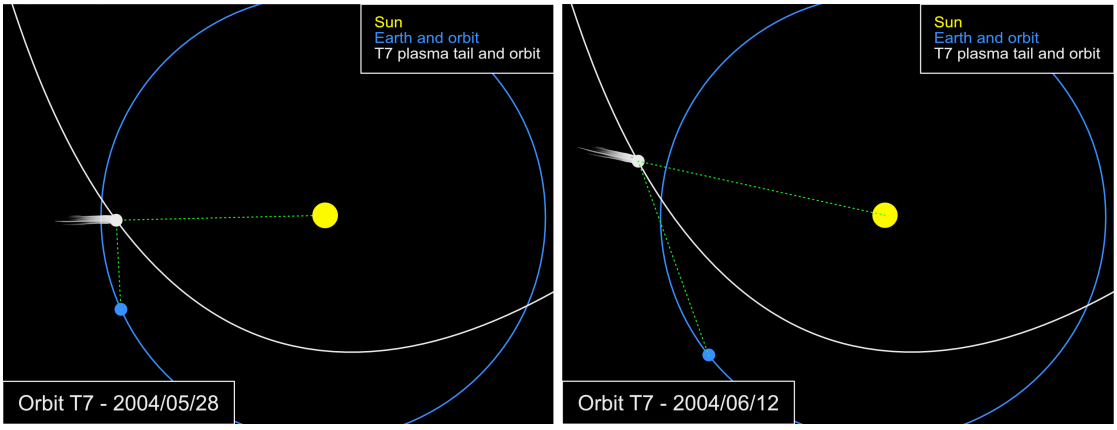


Figure 5.2: Configuration of the Sun, T7, and Earth on 28 May and 12 June 2004. Ephemerides calculated by NASA's Jet Propulsion Laboratory.

Note that the same shift is also visible in the Figure 5.1c.

5.1.2 Methylidyne radical CH

The spectral region ranging from 3630–3660 Å regrouped many still unidentified lines in the Atlas, which turned out to be the CH band B-X 1-0, shown in Figure 5.3. This band is only clearly visible

in the centered spectrum of T7. It has never been identified in a comet before for several reasons: the CH B-X 1-0 band could not be identified in the many atlases, like the one of Comet 122P/ de Vico, for the simple reason that their spectra only start after 3630 Å. There are however some atlases that begin at lower wavelengths, for instance, the atlas of Comet C/1996 B2 (Hyakutake) shown in Figure 5.4, published by A'Hearn, Wellnitz, and Meier in 2015 [76]. However, the B-X 1-0 band was not detected in this spectrum. This illustrates the diversity of cometary spectra that depend on various parameters, for instance the composition of the comet, the Sun-comet distance at the time of the observation, and the position of the slit on the comet (centred or not). The emission lines in Figure 5.3 are well defined and separated, which is characteristic of CH seen in other regions of the Atlas, where the detection of CH is already confirmed.

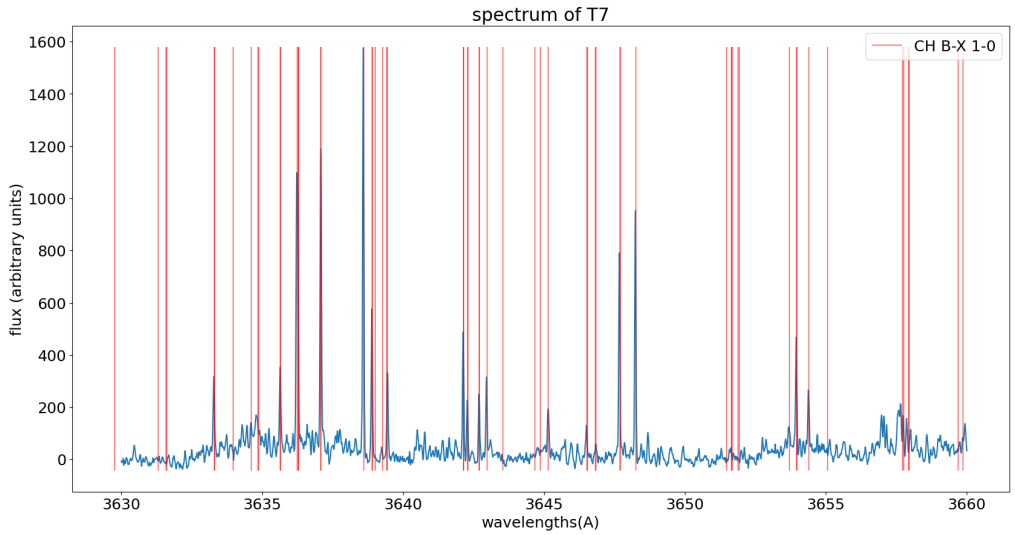


Figure 5.3: CH band B-X 1-0 (red) in the spectrum of T7 centred on the nucleus (blue).

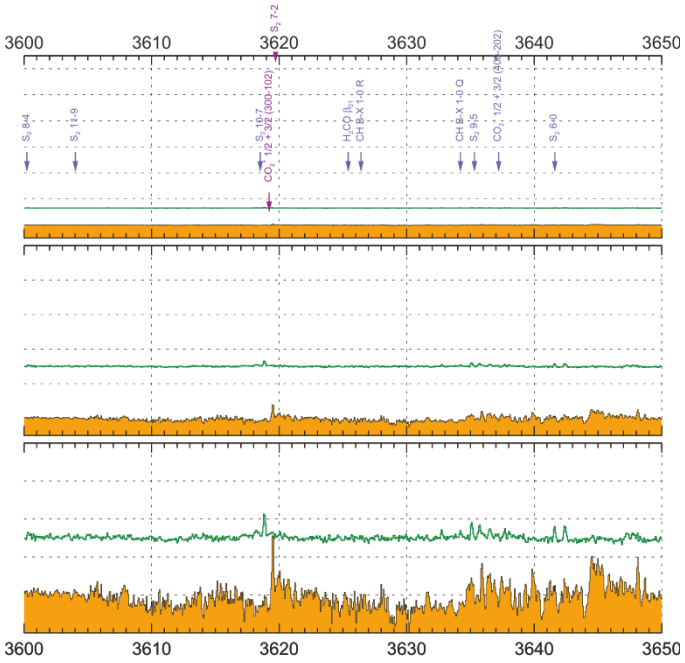


Figure 5.4: 3600–3650 Å region of the spectrum of Comet C/1996 B2 (Hyakutake). The nucleus-centred extract is in orange and the spectrum extracted 10 arcseconds off the nucleus is in green (offset vertically by 1 Å). The three vertical panels represent a 1×, 12×, and 60× zoom respectively. The spectral lines in purple indicate firm identifications, whereas the blue lines may match the observed peaks or not.

To crosscheck this identification, the B-X 1-0 band is also plotted in the spectrum of 8P/Tuttle in Figure 5.5 where the lines correspond well to the spectrum's peaks, confirming this identification.

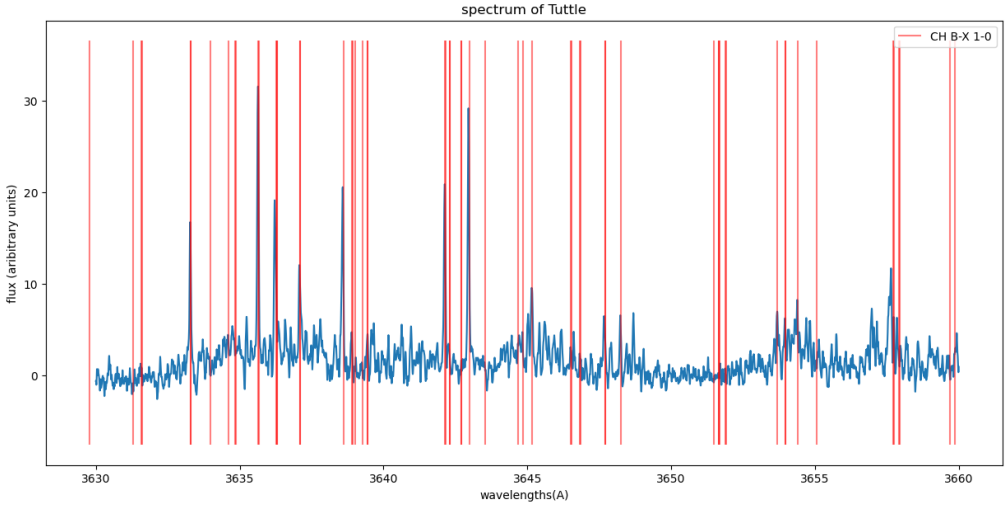


Figure 5.5: CH band B-X 1-0 (red) in the spectrum of 8P/Tuttle (blue).

The 190 lines belonging to the B-X 1-0 band that were added to the Atlas, have been complied by Jorgenson *et al.*, 1996 [77] and were taken from Kurucz's website http://kurucz.harvard.edu/.

Through the addition of this nice and well defined band, 27 unidentified lines in the Atlas were then removed.

5.1.3 Carbon Dioxide Cation CO_2^+

The carbon dioxyde ion was first observed in Comet Bester in 1950 by Swings&Page [78]. The Atlas already contained the 3-3 and 3-1 bands of CO_2^+ at 3503.07–3510.06 Å and 3511.14–3516.98 Å, respectively. In 2019, Opitom, ULiège team *et al.* [41] detected supplementary CO_2^+ bands in the spectrum of the very peculiar Comet 16R2 shown in Figure 5.6.

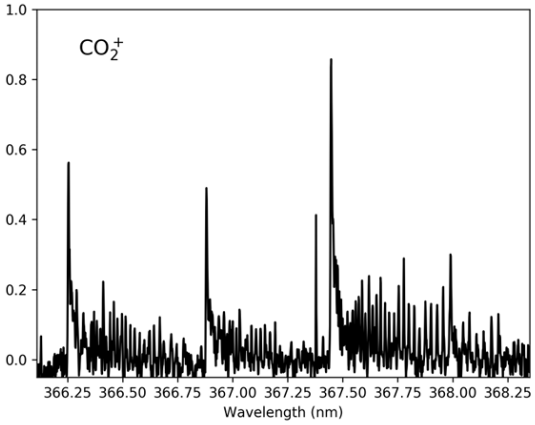
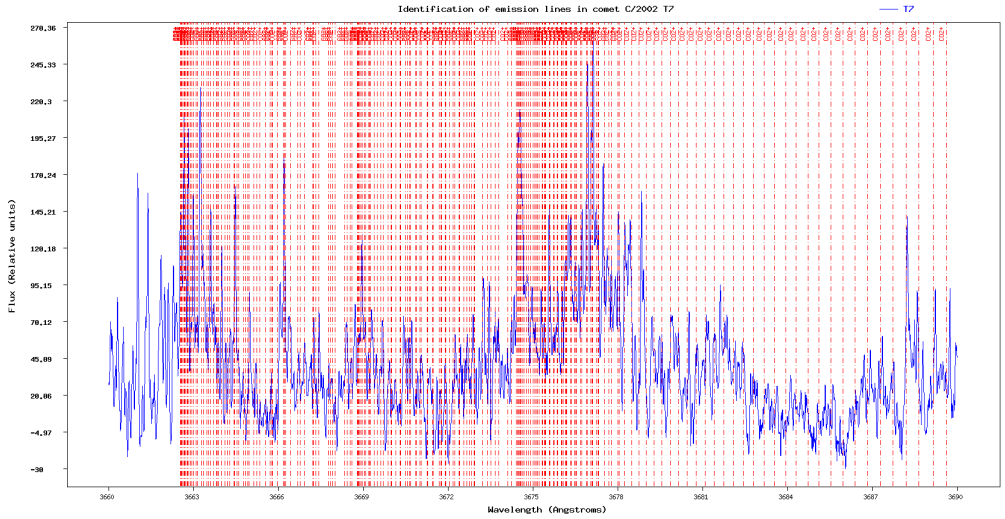
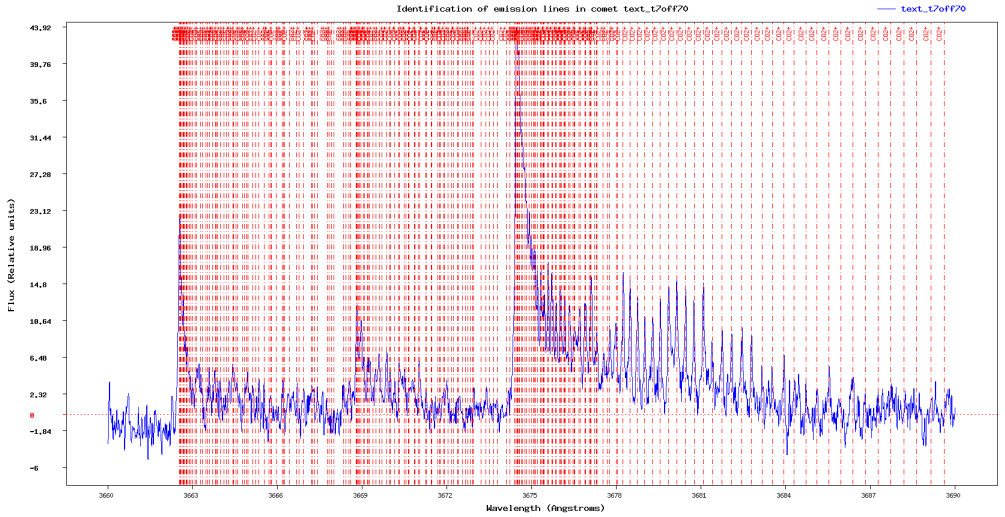


Figure 5.6: CO_2^+ bands in the spectrum of 16R2: $(0,0,0)\,\tilde{A}^2\Pi_{u,3/2} \rightarrow (1,0,0)\,\tilde{X}^2\Pi_{g,3/2}$ at 366.3 nm and 366.9 nm and $(0,0,0)\,\tilde{A}^2\Pi_{u,1/2} \rightarrow (1,0,0)\,\tilde{X}^2\Pi_{g,1/2}$ at 367.4 nm. Source: Opitom *et al.*, 2019 [41].

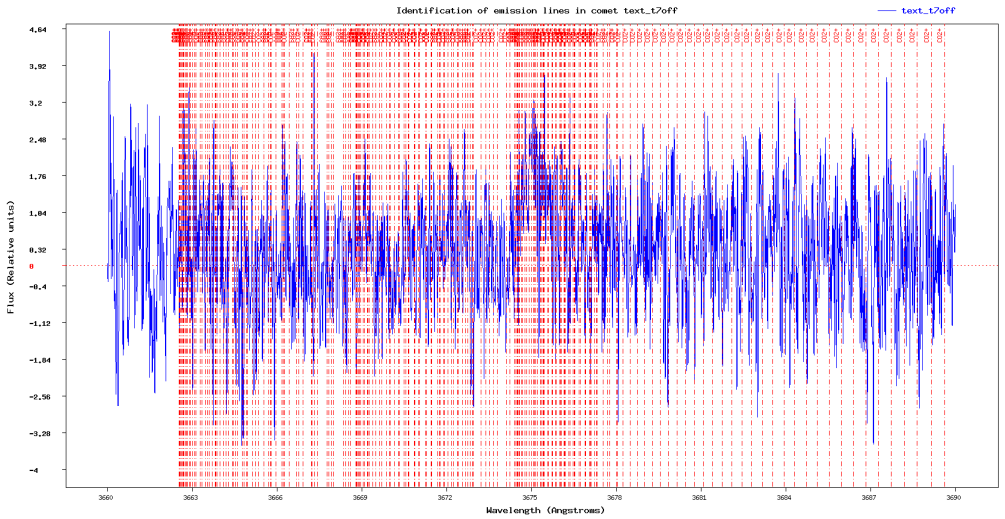
Between 1941 and 1947, Mrozowksi published the four-part series "On the $^2\Pi_u \rightarrow ^2\Pi_g$ Bands of CO_2^+" [79]. The line positions of the bands presented in Figure 5.6 are taken from these articles. Thanks to the identification of these bands in the spectra of 16R2 and the spectra of T7 offset tailwards, 314 lines are added to the Atlas but no unidentified lines are removed. Indeed, despite the presence of unidentified lines between 3660 Å and 3690 Å, the centred spectrum of T7 shown in Figure 5.7a does not display the CO_2^+ lines as nicely as 16R2 does (Fig. 5.6). It is therefore not clear whether the unidentified peaks are caused by CO_2^+ or another molecule that would present emission lines in the same region. Additionally, the presence of unidentified lines right before and after the bands of CO_2^+ reinforces this doubt.



(a) $CO_2{}^+$ lines (red) added to the Atlas in the centred spectrum of T7.



(b) $CO_2{}^+$ lines (red) added to the Atlas in the spectrum of T7 offset by $70''$.



(c) $CO_2{}^+$ lines (red) added to the Atlas in the spectrum of T7 offset by $245''$.

Figure 5.7: $CO_2{}^+$ added to the Atlas.

As presented in Figure 5.7b, the spectrum of T7 offset by 70″ shows the most intense CO_2^+ lines. It is not surprising to find CO_2^+ in both T7 and 16R2 as CO_2 was a common molecule in the presolar nebula (Gicquel *et al.*, 2023 [80]). CO_2^+ is mainly produced by photoionization of CO_2 (Leach, 1988 [81]). We were able to add the bands in the Atlas by means of the presence of CO_2^+ lines in the spectra of 16R2 and the spectra of T7 offset by 70″. Besides the spectral region 3662.5–3682.5 Å presented by Opitom and the ULiège team, other regions might also host CO_2^+ lines and remain a source of further investigation.

In 1986, Wyckoff tentatively[82] detected CO_2^+ in the tail spectrum of Comet Halley at around 4130 Å. This region will be examined in the following section.

5.1.4 4310 Å region

Based on laboratory measurements made by Smyth in 1931 [83], Wyckoff was able to detect CO_2^+ in the spectrum of Comet Halley's tail at 4130 Å. This region is filled with unidentified lines in the Atlas, raising the possibility that it might correspond to CO_2^+ as detected by Wyckoff. Even though CO_2^+ is only weakly detected in the centred spectrum of T7, the possibility that the unidentified band corresponds to a CO_2^+ band is investigated because of the close match between the position of the unidentified band in the centred spectrum of T7 and the CO_2^+ band detected by Wyckoff. To detect CO_2^+, it is best to use the spectra of 16R2 and T7 offset by 70″ as shown in the previous section. However, Symth's lines do not align with the emission lines of the spectra as shown in Figure 5.8. The uncertainty on the wavelengths of the article is only 0.1 Å, meaning that the mismatch between the laboratory lines and the spectrum cannot be due to uncertainty. However, when plotting the lines from Symth in the region 3660–3690 Å, where the presence of CO_2^+ was confirmed in the previous section, they do not match with the spectrum either. The appearance of the unidentified band at 4130 Å matches the ones observed for CO_2^+, however, it is important to keep in mind that CO_2^+ bands are barely detected in the centred spectrum of T7. A more recent line list needed to confirm or reject this suspicion was not found in the literature.

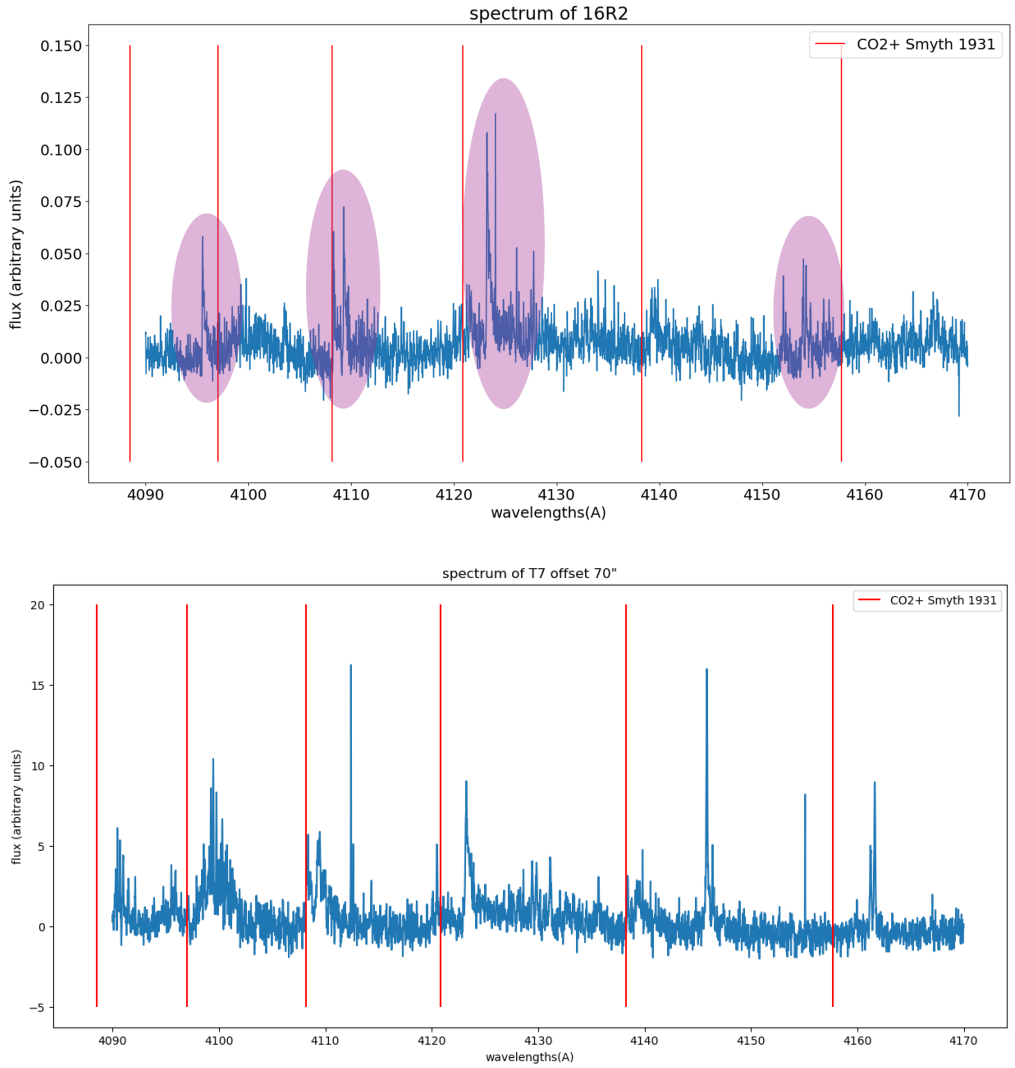


Figure 5.8: $CO_2{}^+$ lines from Smyth superposed on the spectrum of 16R2 and T7 offset 70″. Broad unidentified lines are visible in the spectra, resembling ionic ones. Their origin still remains unclear: it could be $CO_2{}^+$ or another ion that still need to be identified.

Due to C_3 crowding the region around 4130 Å in the spectrum of T7, this molecule is considered another candidate besides $CO_2{}^+$. However, there is only few C_3 in 16R2 and C_3 intensity drops fast with increasing distance from the nucleus. Since the unknown band is clearly visible in the spectrum of 16R2 and in the tail of T7, this provides evidence that these bands are not C_3.

According to Rousselot *et al.* (2023 [84]), the (4,1) band of CO^+ is located around 4100 Å. The intensity of this band is however very weak. The bands (3,0) and (2,0) that are located before and after the (4,1) band are already in the Atlas, raising the possibility that the (4,1) band is also present. However, the appearance of the band resembles that of $CO_2{}^+$, compact band, whereas CO^+ presents bands where the lines are well separated.

The detection of O_2 by Rosetta (Bieler *et al.*, 2015 [85]) in Comet 67P/Churyumov–Gerasimenko with a mean abundance of $3.8 \pm 0.85\%$ introduces the possibility for the unknown band at 4130 Å to be $O_2{}^+$. However, the lack of extensive line lists prevents further investigation.

Besides the strongest band at 4125 Å, other weaker bands having a similar shape were discovered. They are marked by the pink ellipses in Figure 5.8 and are assumed to have the same origin as the principal peak at 4125 Å.

5.1.5 Carbon Monoxide ion CO^+

Rousselot *et al.* (2023, [84]) present a fluorescence model of CO^+ based on the ULiège UVES spectrum of the peculiar CO and CO_2 rich 16R2 Comet. Some of the bands that were identified in this article are still missing in the Atlas, including the bands (classified in ascending order of wavelengths)

(5,0)
(5,1)
(4,1) (possible candidate for the unidentified band at 4125 Å)
(3,1)
(3,2)
(0,0)
(5,4)
(2,2)
(3,3) (shown in Figure 5.9)
(4,4)

These are only the bands that are bright enough to be detected in 16R2, though some might be superposed with emission lines from another molecule. Since the line list of the missing band computed by Rousselot *et al.* is not available, other references were searched for. Brown, Dittman, and McGilvery [86] measured the (0,0) band in 1984 . While these measurements lie in the spectral region Rousselot *et al.* detected the (0,0) band (4875–4960 Å), the exact line positions do not match those of the 16R2 spectrum. Since Rousselot *et al.*'s model presented in their article fits the observational spectrum much better, we suppose Brown *et al.*'s lines to be slightly offset. Nevertheless, the identification of new bands in the Atlas is firm thanks to Rousselot *et al.*'s article and will be added upon obtention of the line list. Most of the bands are not clearly present in either of the T7 spectra, meaning that not many unidentified lines will be removed by this addition. 16R2 spectrum is indeed unique with its very strong CO^+, $N_2{}^+$ and $CO_2{}^+$ lines and they helped us to enrich the cometary line Atlas, with even more lines not present in the Atlas yet but of cometary interest.

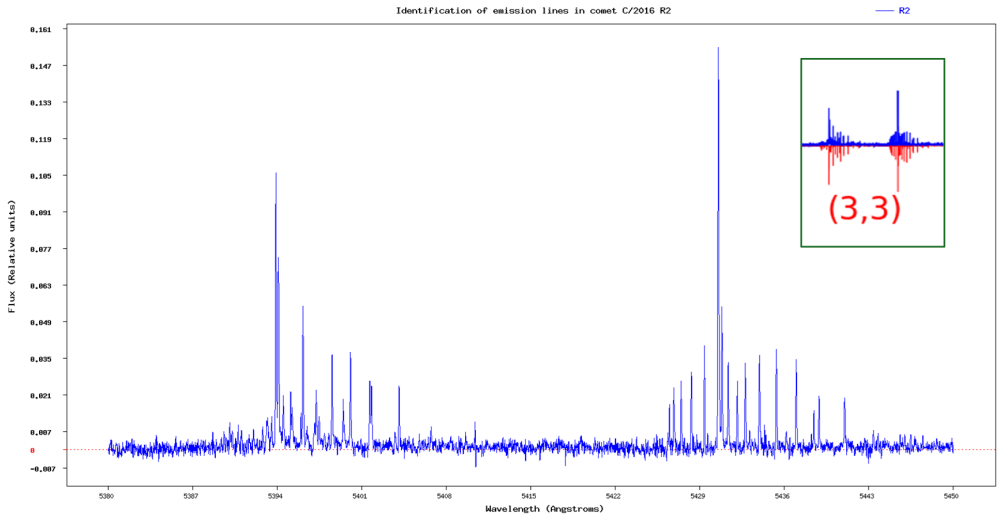


Figure 5.9: (3,3) band of CO^+ in the spectrum of 16R2. The results of Rousselot *et al.*'s model are added in the green frame for comparison. The red spectrum indicates the (3-3) band computed by Rousselot *et al.*. This plot is from the Atlas and will need to be updated to include the CO^+ lines (work in progress).

5.1.6 Disulfur S_2

Disulfur was first observed in Comet IRAS-Araki-Alcock (C/1983 H1) by A'Hearn *et al.* in 1983 [87]. In 2003, Kim *et al.* [88] detected S_2 in Comet Hyakutake by comparing the spectrum of this comet to a fluorescence model of S_2. On this basis, it was investigated whether S_2 is present in one of the spectra used for this work. S_2 is known to have a very short lifetime, meaning that it is only found at short distances from the nucleus. Therefore, the spectra of T7 offset by $70''$ and $245''$ are not good candidates to detect S_2. To attempt the detection of S_2, the new line list of sulfur dimer published in 2024 by Gomez *et al.* [89] was used.

Several bands have been looked for in the centred spectra of T7, 16R2, X5, and 8P/Tuttle, but none showed convincing proof for the presence of S_2 in the spectrum of the respective comet. Two of the bands that were analysed, yet could not be definitively identified in any of the spectra, are presented hereafter. The comparison between the B-X 8-0 band found in Hyakutake, shown in Figure 5.10, and the B-X 8-0 line positions plotted over the centred spectra, shown in Figure 5.11, provides no convincing evidence for the presence of the 0-8 band. Note that the peak at 3849 Å has been attributed to CO^+.

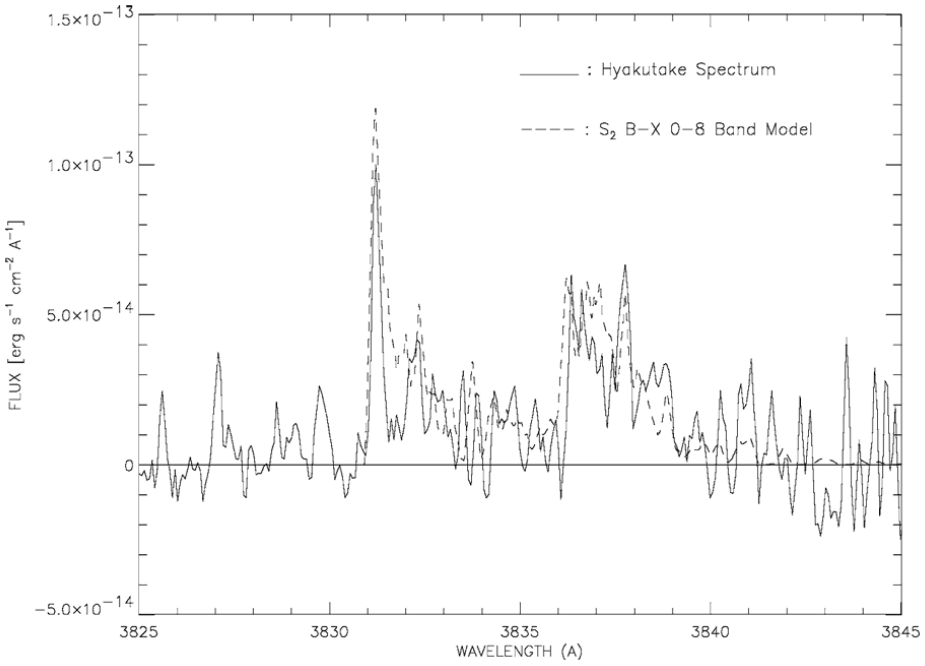


Figure 5.10: Comparison between the fluorescence model of the S_2 1-4 band (dashed line) and the spectrum of Comet Hyakutake (solid line). The spectrum exhibits prominent S^2 emission lines due to Comet Hyakutake's close approach to Earth at a distance of only 0.1 au. This proximity enabled the capture of a high spatial resolution spectrum in the region near the comet's nucleus. Source: Kim *et al.* (2003, [88]).

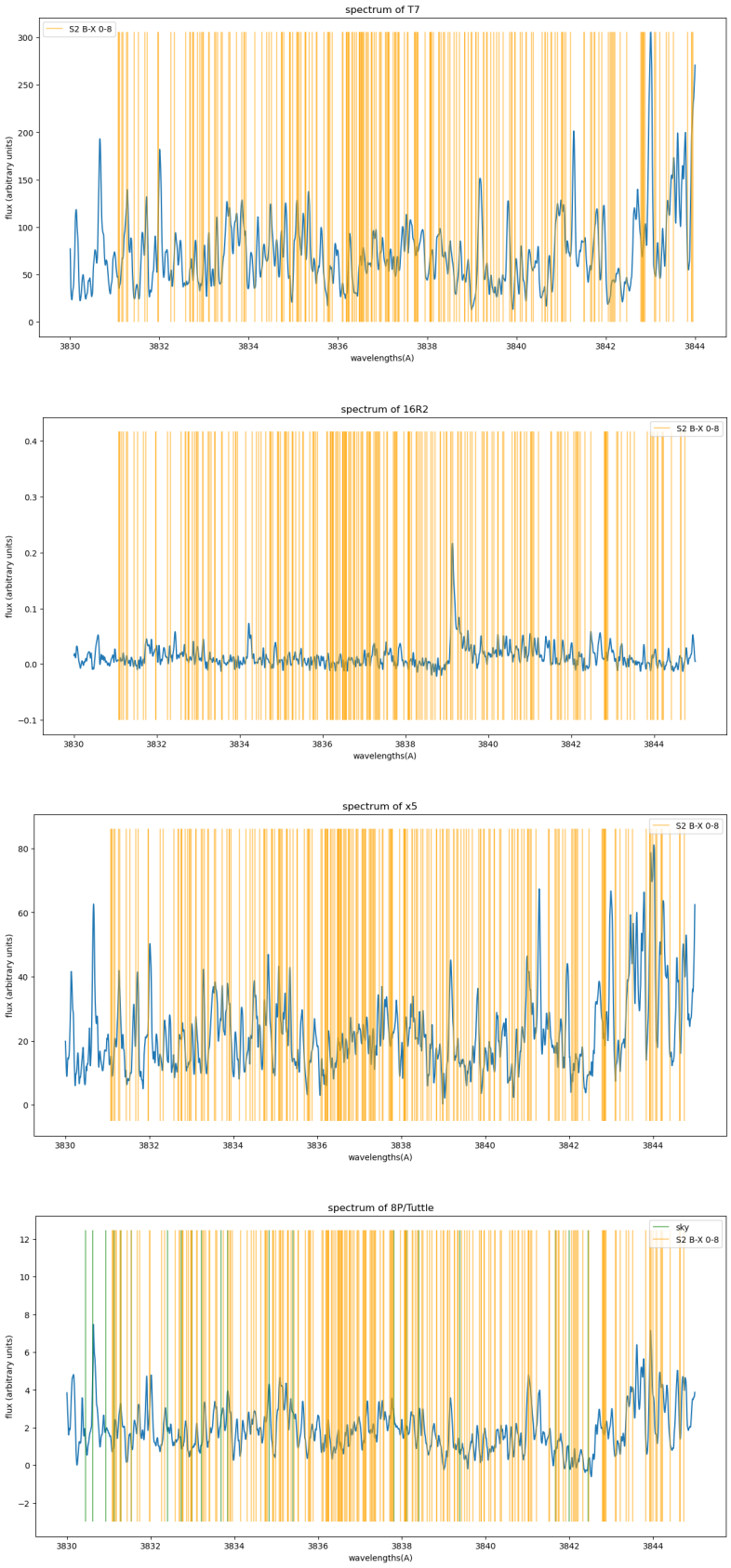


Figure 5.11: The S_2 0-8 band plotted in our spectra (T7, 16R2, X5, and 8P/Tuttle).

The following figures compare the B-X 1-4 band of S_2 in Comet Hyakutake, shown in Figure 5.12, and the B-X 1-4 band of S_2 in our spectra, shown in Figure 5.13. Note that the line marked by the question mark in Kim *et al.*'s article has been identified as NiI in our Atlas using Manfroid *et al.*'s publication in 2021 [58]. The last image showing the spectrum of Comet 8P/Tuttle is the only one displaying possible traces of S_2. However, the search for other S_2 bands, including the 0-8 band, in the spectrum of 8P/Tuttle, showed inconclusive results.

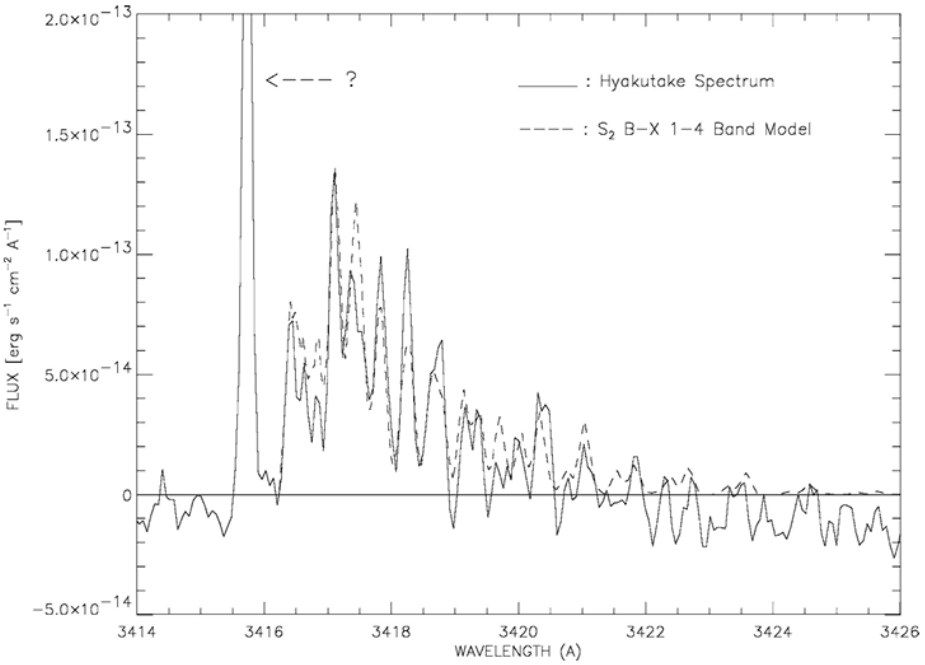


Figure 5.12: Comparison between the fluorescence model of the S_2 0-8 band (dashed line) and the spectrum of Comet Hyakutake (solid line). Source: Kim *et al.* (2003, [88]).

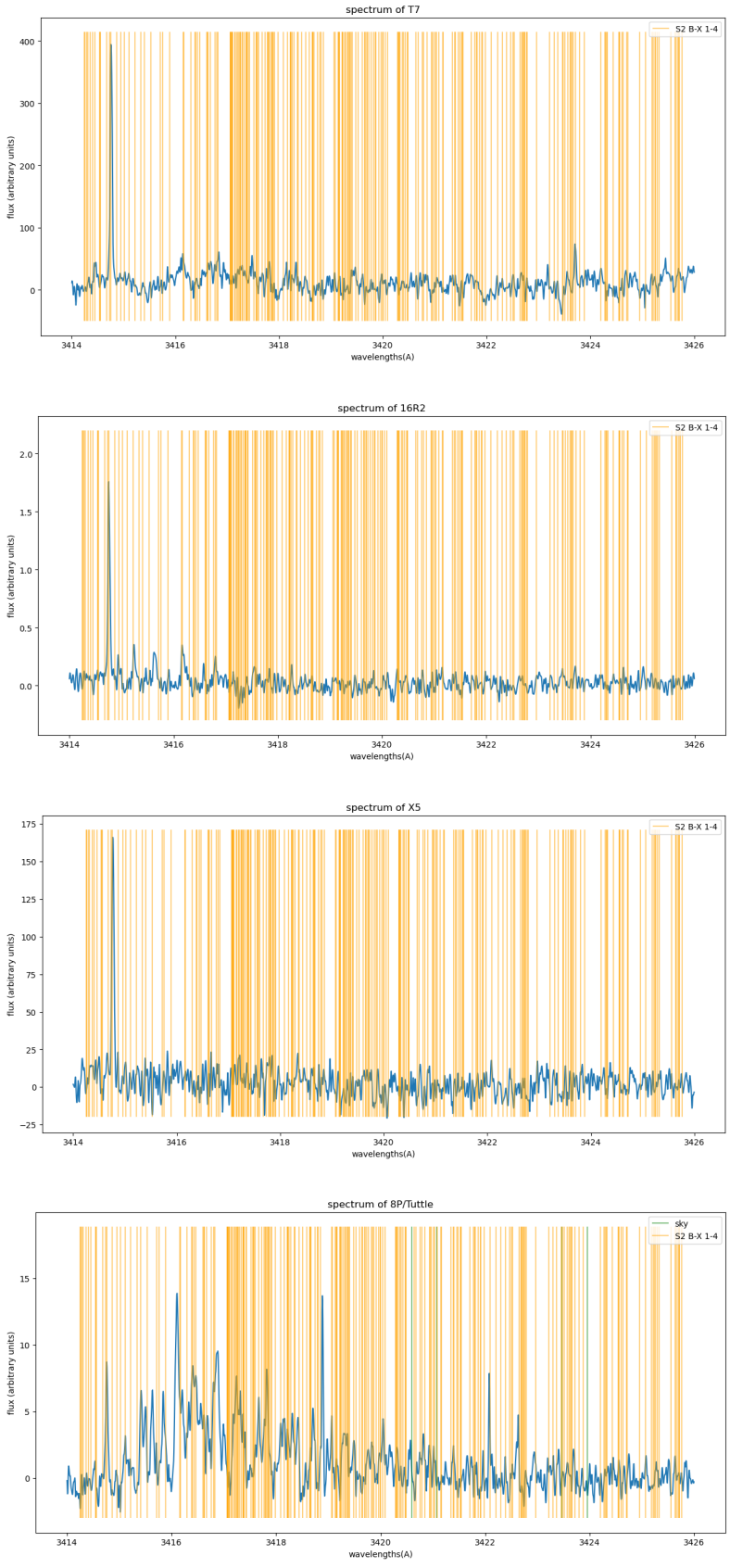


Figure 5.13: The S_2 1-4 band plotted in our spectra (T7, 16R2, X5, and 8P/Tuttle).

5.1.7 New sky emission lines

Emissions of the night sky produce lines in the spectrum that are superposed to the emission lines of the comet. ESO has measured the night-sky emission using UVES and the obtained spectrum is available on their website. A 160 Å chunk of this spectrum is shown in Figure 5.14, the blue lines indicate the lines that have been identified through the criteria set by ESO. Hardy has inserted the night-sky lines in the Atlas and denoted them as "sky".

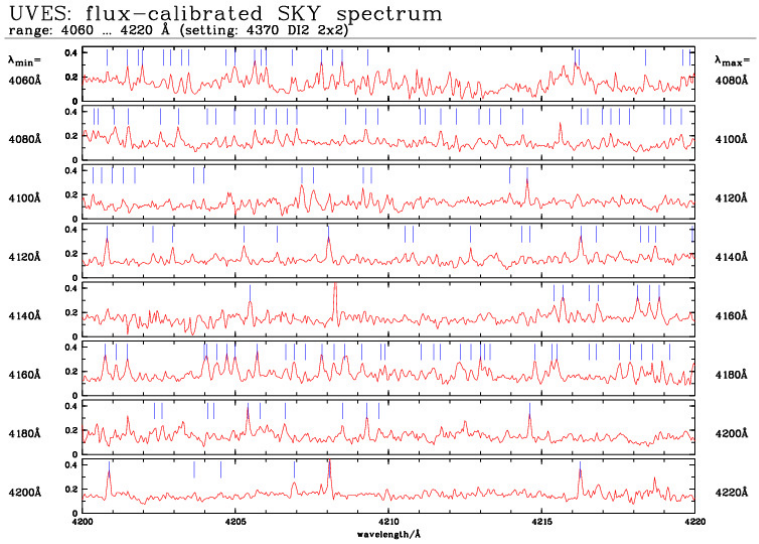
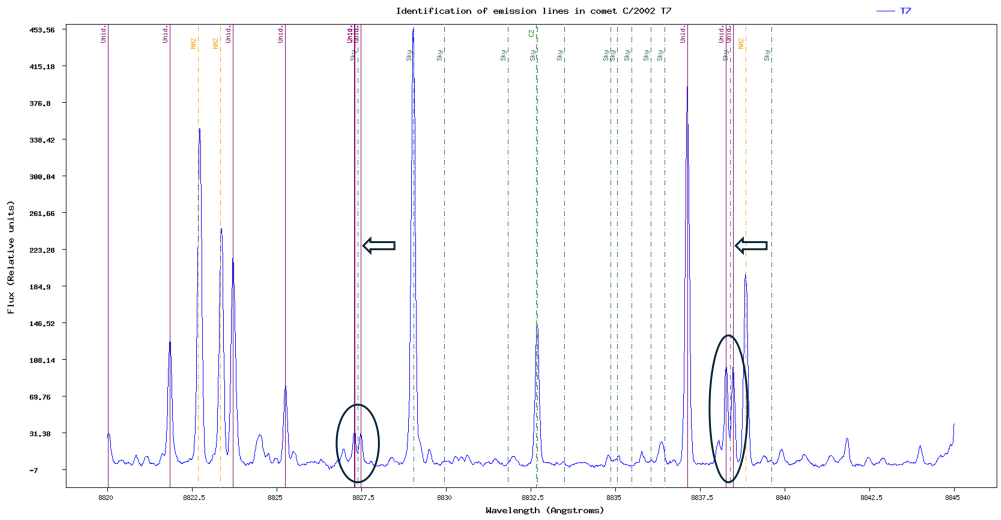
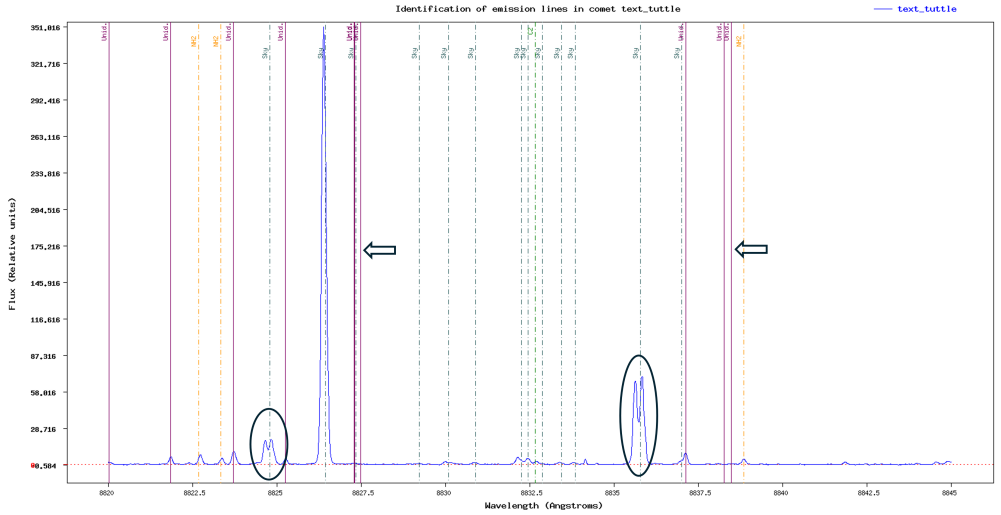


Figure 5.14: Night-sky emission measured by UVES. Blue bars indicate the measured lines. Source: https://www.eso.org/observing/dfo/quality/UVES/uvessky/sky_3460_1.html

The resolving power of this spectrum is $R = 45,000$ in the blue and $R = 43,000$ in the red, whereas $R = 80,000$ in the blue and $R = 110,000$ in the red for the spectrum of T7. T7 being of higher resolution implies that it contains more refined sky emissions than the reference sky-emission spectrum. This can lead to sky doublets in the spectrum of T7 whereas ESO's spectrum only detected one sky line, resulting in unidentified lines in the Atlas. The sky doublets appear as two "Unid." lines separated by a "sky" line. This pattern may hint towards potential errors. One has to be careful however, because it is possible that the two unidentified lines that surround the sky line are actual cometary lines. It is therefore most prudent to compare those lines in another cometary spectrum in order to verify whether the unidentified lines are emitted by the sky or the comet. Sky emissions are not located at the same wavelengths for spectra taken at different geocentric velocities, while cometary emission lines stay fixed. Consequently, a sky line that is encompassed by two unidentified peaks in two different spectra confirms the atmospheric origin of the unidentified lines. Such an example is shown in Figure 5.15 where the suspected sky emissions in T7 were confirmed in the spectrum of Comet 8P/Tuttle.



(a) Four suspected sky emission, still labeled as "Unid." in purple and marked by the arrows with ESO's sky emission lines in the middle marked as "sky" in green in the Atlas.

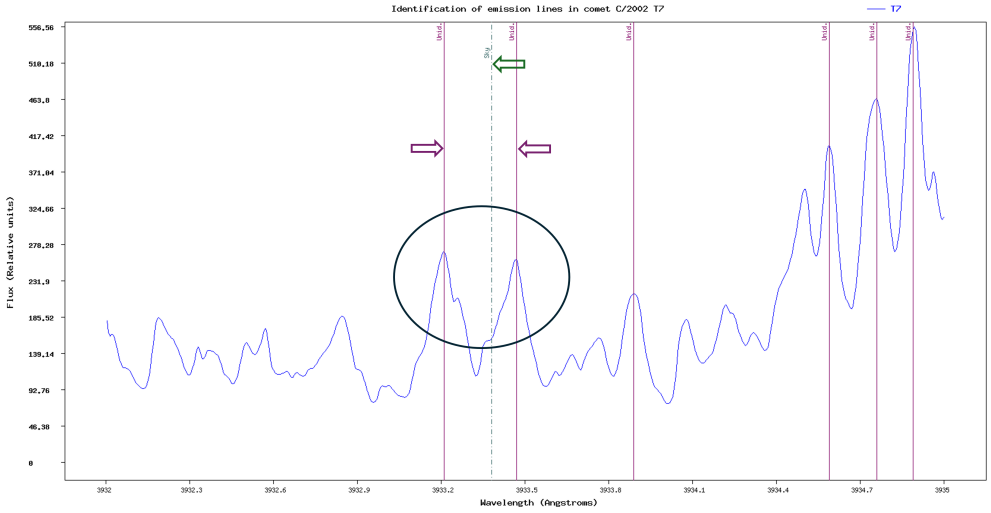


(b) The four "Unid." lines in purple marked by the arrows remain fix at 8827 Å and 8838 Å whereas the sky emission lines (green dotted line inside the ellipse) have shifted to 8824.5 Å and 8836 Å in the spectrum of Tuttle. The four unidentified peaks marked by the ellipses have shifted with the "sky" line, confirming their atmospheric origin.

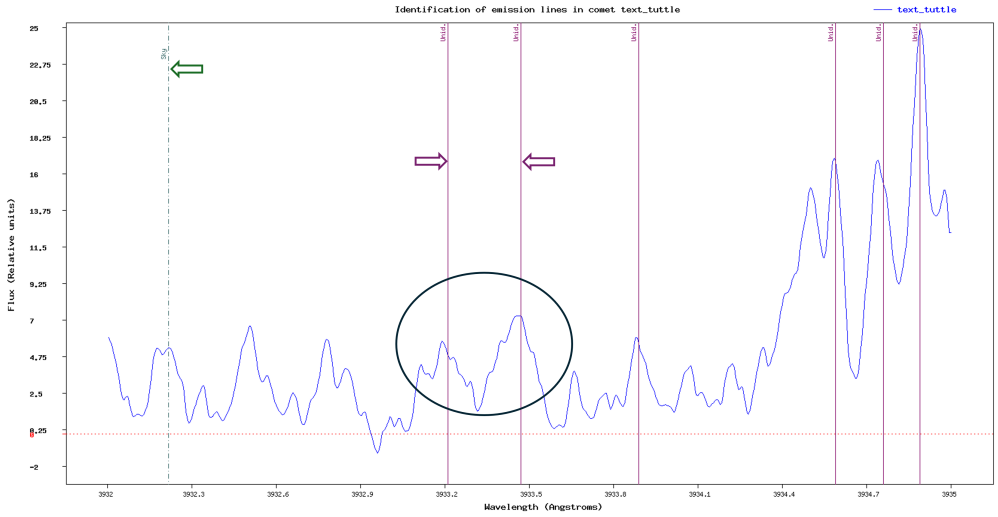
Figure 5.15: Two examples where the four purple "Unid" lines are atmospheric lines.

However, as described, two unidentified lines of cometary origin can coincidentally surround a sky line in T7's spectrum. This scenario is described below and in Figure 5.16. To avoid classifying cometary lines as sky lines it is necessary to crosscheck these lines in another spectrum, for instance that of Tuttle. If the two unidentified peaks bordering the sky line in the spectrum of T7 remain at the same wavelengths in Tuttle's spectrum while the sky emission shifts position, then these "Unid." lines are of cometary origin. Figure 5.16 illustrates this, showing two unidentified lines at 3933 Å in T7's spectrum that were initially suspected to be sky emissions. However, upon verification with Tuttle's spectrum, this hypothesis was disproved. As demonstrated in Figure 5.16, the two unidentified peaks around 3933 Å remain fixed, while the sky emission shifts to the left. Consequently, these two unidentified lines are indeed of cometary origin and demand further

investigation to determine their nature.



(a) 2 suspected sky emission, labelled as "Unid." in purple with ESO's sky emission line in the middle marked as "sky" in green in the spectrum of T7.



(b) The 2 "Unid." lines in purple remain fixed at 3933 Å whereas the sky emission line in green has shifted to 3932 Å in the spectrum of Tuttle. The two unidentified peaks marked by the ellipse remain fixed, disproving their celestial origin.

Figure 5.16: Examples where the two purple "Unid" lines are cometary lines. Their origin remains unknown and still needs investigation.

Five sky doublets have been detected in T7 and confirmed in Tuttle. To incorporate these in the Atlas, the following steps are implemented:

- The Atlas uses a file containing the position of the sky lines in the reference frame of Earth. When users load a spectrum, they have to indicate the geocentric velocity of the comet to allow the algorithm behind the Atlas to compute and plot the sky lines in the frame of the comet.
- The position of the "Unid." lines have been measured in the reference frame of T7.

- It is therefore not possible to simply move the two bordering lines marked initially as “Unid.” into the “sky” file. Instead, a conversion of the “Unid.” lines from T7’s frame to Earth’s frame is needed.
- This is achieved by applying a Doppler shift to the ”Unid.” lines that counteracts the Doppler shift induced by the T7’s geocentric velocity. Given that the geocentric velocity of T7 on the observation dates was 65.6 km/s, a Doppler shift of -65.6 km/s is applied to the “Unid.” lines.
- The “Unid.” lines are removed from the Atlas and the converted lines are added to the “sky” file of the Atlas.
- Lastly, the initial central sky emission line is removed from the “sky” file.

This results in ten unidentified lines being deleted from the Atlas.

5.1.8 Artefacts in the spectra

It was discovered that a few of the unidentified lines in the spectrum of T7 were not comet or sky lines but rather artefacts. These artefacts can be produced when a very bright band hitting the CCD camera in the red arm gets reflected back to the grating. Upon the second reflection on the grating, the secondary band gets captured by the CCD camera as indicated in the UVES user manual [90]. It is possible to detect such reflected bands by carefully looking at the 2D spectrum of the comet. In Figure 5.17, the *ghost lines* are marked by the red ellipses. One notices that these lines do not follow an order and can therefore not be real lines. This is characteristic for superposed reflected lines. In the spectrum of T7, ghost lines appear at 3248–3252 Å and at 3274–3286 Å and are due to the reflection of the bright CN band.

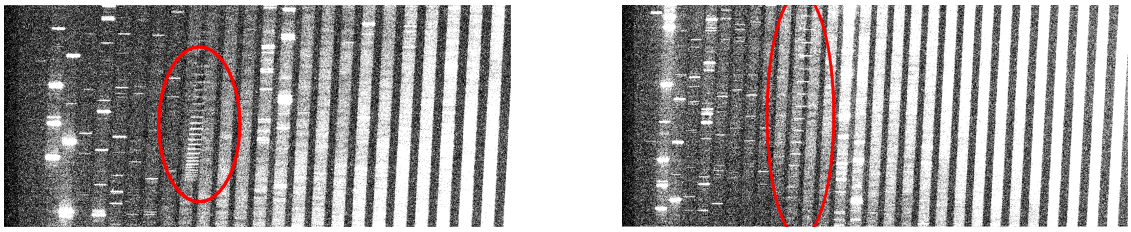
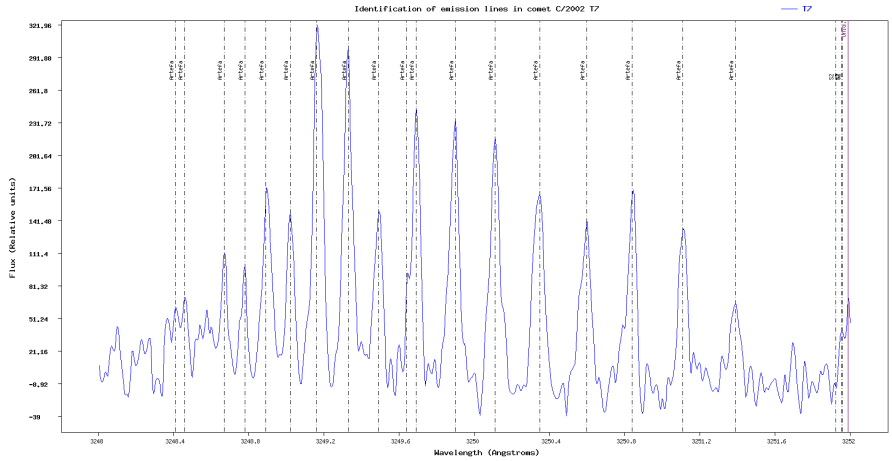
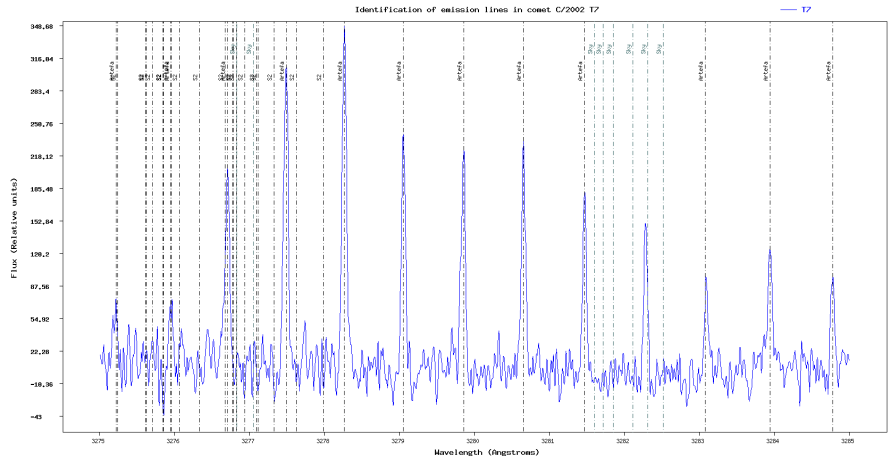


Figure 5.17: Parts of the 2D spectrum of T7. The red ellipses indicate the ghost lines which are superimposed on the orders.

These ghost lines have been removed from the list of unidentified lines, and instead, they are now named “Artefact” in the Atlas. This eliminated 18 “Unid.” lines in Region (a) shown in Figure 5.18 and 13 in Region (b).



Region (a) from 3248-3252 Å



Region (b) from 3275-3285 Å

Figure 5.18: Artefacts due to reflection

The ghost lines, referring to artificial lines caused by reflections, are not the only artefacts that can be present in a spectrum. Indeed, it is possible that during the observation, cosmic rays hit the CCD camera. Cosmic rays are high-energy particles originating from the Milky Way or other galaxies and can be as energetic as 10^{21} eV. Cosmic rays are frequent with an incidence rate of 50 per cm^2 per hour [64]. They get recorded and appear point-like in the 2D spectrum, as shown in Figure 5.19. Thankfully, the UVES pipeline can eliminate some of those artefacts during the data reduction. Additionally, taking the median of several observed spectra of the same comet erases even more artefacts as the cosmic rays never fall two times at the same place. The spectrum used for T7 is the median of several spectra, explaining the absence of the cosmic ray traces. Unfortunately, if only one spectrum is available and it is consequently impossible to compile the median, the artefact might remain in the 1D spectrum. Luckily, it is possible to distinguish false lines caused by cosmic rays from real molecular lines in the 1D spectrum: Real molecular lines have a certain width as explained in Section 4.1.4 and show a Gaussian profile, whereas cosmic lines are very thin (usually only one pixel). Such artefacts were detected in the spectrum of Comet C/2002 X5 (Kudo-Fujikawa) indicated by the pink lines in Figure 5.20b. Figure 5.20 compares the same spectral region in Comet T7, where the median removed the cosmic rays, and in Comet X5, where no median was taken, leaving the cosmic rays not removed by the UVES pipeline.

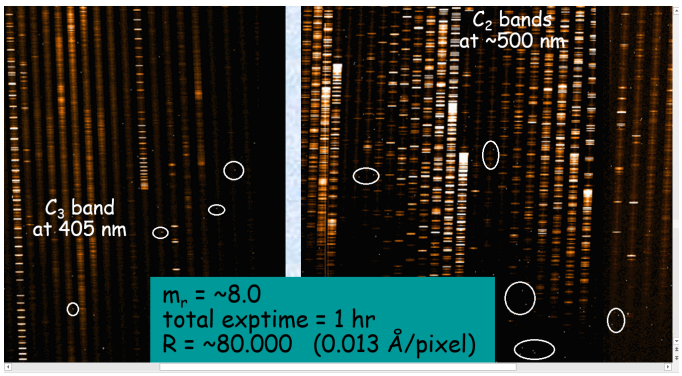
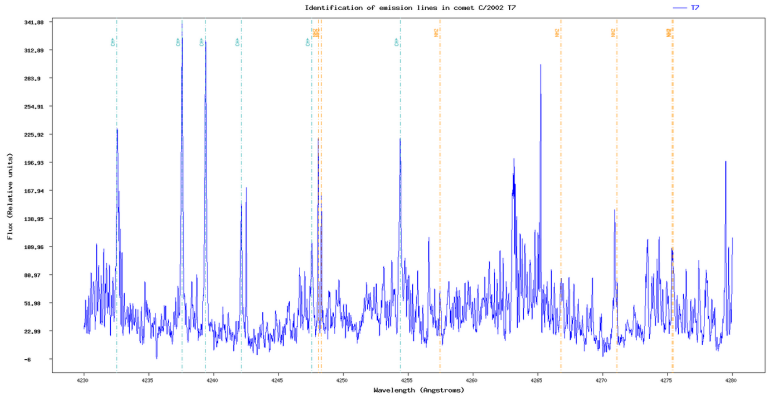
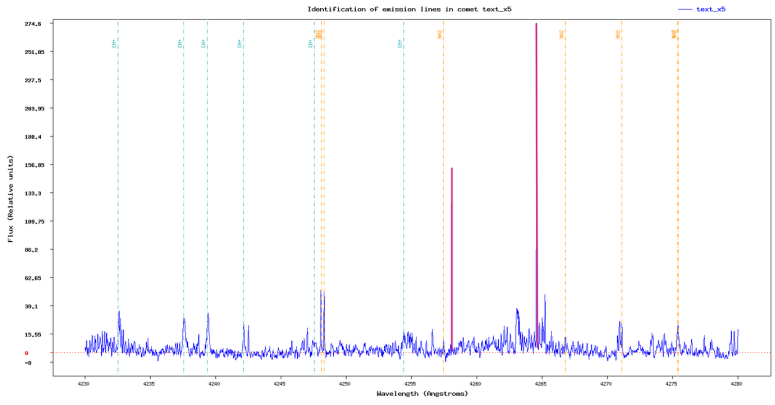


Figure 5.19: 2D spectrum of Comet C/2002 T7. The white ellipses indicate the numerous traces left by cosmic ray impacts.



(a) Spectrum of T7, where the cosmic rays are removed.



(b) Spectrum of X5, where the cosmic rays are not removed (pink).

Figure 5.20: Comparison between the same region in T7 (without cosmic rays) and X5 (with cosmic rays.)

5.2 NH fluorescence model

As described in Section 4.2, where the methods for the fluorescence model of imidogen are presented, the model can be divided into three major parts. It is beneficial to find means to verify

each step, in order to pinpoint errors that might come up. These verification methods and their outcome are explained hereafter, before presenting the final product of the model.

The first step is computing the transition energies for each possible transition. This is done by subtracting the energy of the lower level from the energy of the higher level. The article from Fernando *et al.* [70] that provides the energy levels, also indicates the transition energies. We were therefore able to confirm that the transition energies computed by the model are correct. This is the base for a model as the transition energies give the position of the lines. A greater transition energy is synonym of a greater wavenumber resulting in a smaller wavelength. The computing of Einstein coefficients A and B is the next step in the model conception. The results are compared to the Einstein coefficient available on the Exomol website [91]. Since the Hönl-London coefficients, needed to compute the Einstein coefficients, were calculated using simplified formulae from Herzberg, our results do not match Exomol's coefficients perfectly. Nevertheless, our results seem to roughly coincide with the Einstein coefficients found online.

Once the transition probabilities are computed, the population of each energy level needs to be calculated. The population of the excited state should be much lower than that of the ground state due to shorter lifetimes. This is verified for the model where the ratio between the population of the excited state and the population of the ground state is of the order of 10^8.

The last step of the model is the computation of the line intensities. The verification of this last step is straightforward and consists in plotting and comparing the model to prior knowledge of NH. The lines obtained have the form of a Dirac peak. To compare the result with other spectra, it is therefore useful to convolve the model lines with a Gaussian. This procedure is also known as applying a Gaussian filter. The full width at half maximum (FWHM) of the Gaussian was adapted to fit the FWHM of the instrument response function of the UVES instrument to which the model is compared to. Figure 5.21 and Figure 5.22 show the model output before and after applying the Gaussian filter, for illustration purpose only.

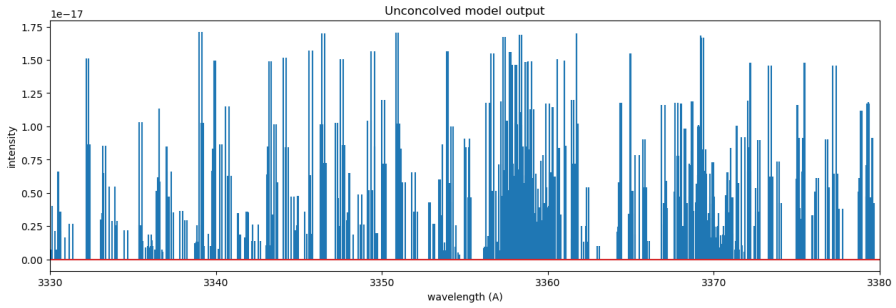


Figure 5.21: Unconvolved output of the NH model.

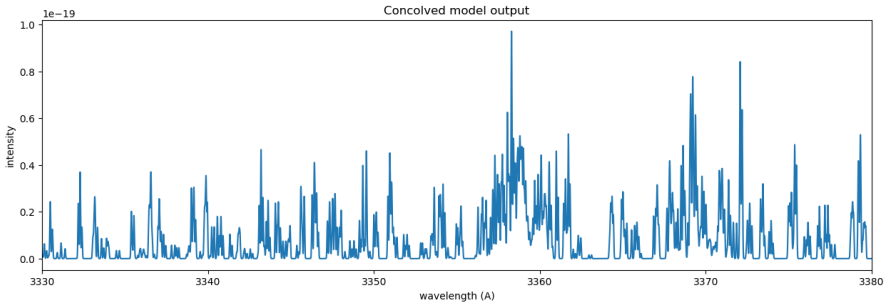


Figure 5.22: Convolved output of the NH model.

The work done in this thesis lays the foundation for refining the model in the future. Indeed, the model contains rotational transition probabilities only for low values of J since no fitting formula was found in the literature. Therefore, the rotational energy levels of high J are too populated compared to reality, causing wrong intensities. To nevertheless get a better grasp of the quality of the model, the satellite lines where $\Delta J \neq \Delta N$ are removed to compare the model to the T7 spectrum. T7 was chosen for this comparison as NH bands are clearly identified in its spectrum. Figure 5.23 shows the 0-0 band of NH in the spectrum of T7 in blue and the superposed model in pink. The transitions that are too intense in the model appear as extra lines compared to the spectrum of T7.

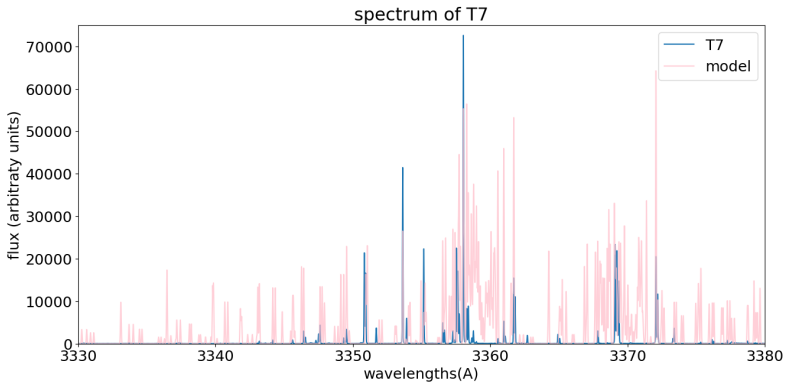


Figure 5.23: Comparison between the NH model and the 0-0 NH band in the spectrum of T7.

6 Conclusion and Scientific Perspectives

This dissertation is divided into five sections, followed by this concluding section, which outlines research perspectives. The first section is devoted to the concept of the atlas. Unlike an atlas used in geography, a cometary atlas is distinguished by its objective of cataloguing both the identified and unidentified emission lines of molecules in comets. This is followed by a section on comets to contextualise the topic. In this section, a description of the perception of comets over time, from Antiquity to the present day, is first given to place the scientific research on comets in its historical context. Next, a general description of the characteristics of comets, including both the classification of comet orbits and their anatomy and composition, is provided to place comets in the context of celestial bodies. Finally, a description is given of the comets that are the subject of this master's thesis: T7, 16R2, X5, and 8P/Tuttle. The third section presents the theoretical framework of the dissertation: spectroscopy and fluorescence. Section 4 describes the methodology of this study by introducing cometary spectra; it explains what they are, how they are taken, and some of their characteristics. The results of this study are presented in Section 5.

The aim of this master's thesis is to expand the current ULiège cometary line Atlas, which was developed to bring together and organise all the identified and unidentified molecular lines of comets. The Atlas is a reference for cometary scientific research and also serves as a tool for researchers to identify lines in their spectra as this is not an easy task due to the many lines in comets as well as to compare different spectral regions and spectra from different comets. It plots the emission lines onto the spectra of comets, serving as visual support for their identification. One of the objectives of this master's thesis is to investigate the many unidentified lines still present in the spectrum of T7 and detect new lines using spectra from different comets. The second objective is to build a fluorescence model for imidogen (NH).

The key findings of the work are the identification of 579 molecular lines and nine molecular bands of CO^+, for which the exact number of lines is unknown. These 579 lines have been added to the Atlas, whereas 57 unidentified lines have been removed because we were able to identify or discard them. Table 4 summarises the new lines that have been added to the Atlas and the unidentified lines that have been removed. The number of added lines is much larger than the number of unidentified lines removed, as most bands are ionic and not present in the T7 central spectrum that was used for the original spectrum of the Atlas. These ionic bands have not always been detected in the centred spectrum of T7, as they can be hidden by the large number of other lines crowding a spectral region. In addition to the molecules indicated in Table 4, other molecules have been investigated as possible emitters of unidentified lines. The identification has not always been conclusive but has reduced the number of possibilities.

Table 4: Summary of the lines added to the Atlas

Molecules	**Lines Identified**		**Unidentified Lines Removed**
	Lines in Atlas	Bands not yet in Atlas	
H_2O^+	55		2
CH	190		27
CO_2^+	314		0
CO^+		9	0
Sky doublets	20		10
Ghost lines	0		18
Total	579	9	57

These findings improve the quality of the Atlas and contribute to the field of cometary research because the Atlas consolidates the cometary lines into a single place, easing the work for researchers.

The first objective of this thesis of growing the Atlas by identifying new lines (579) is accomplished even though a large number of unidentified lines still remain in the Atlas (57 out of 4928 were identified). Despite all our efforts, it is challenging to assign the unidentified lines in the spectrum of T7 to molecules. The second objective of building a fluorescence model could not be developed to its full potential. However, this thesis lays the groundwork for future improvements to the model. As discussed in Section 5.2, the inaccuracies in the intensities of the fluorescence model are attributed to the overpopulation of high-energy rotational levels. The root cause of this issue is the current absence of a formula for calculating the rotational probabilities for the $^3\Pi_i - {}^3\Sigma^-$ transition.

In terms of scientific perspectives, the Atlas should be considered a *living document*, which means that it is continually evolving and can be updated as new discoveries are made. There are many possibilities for further developing the Atlas, some of which are presented in the following.

To expand the Atlas in the future, it is essential to monitor newly published cometary line lists. The lack of recent line lists for certain molecules complicates their definitive identification, underscoring the importance of incorporating updated data as they become available. Candidate molecules that could potentially be detected in one of the spectra taken by the ULiège team are molecules that have already been detected in optical spectra of other comets, for example, NCN and H_2CO detected in Comet Hyakutake. Other possibilities include daughter species of the molecules detected by Rosetta (Fig. 2.6) like CS^+ or SH. The analyses carried out in this thesis suggest that most of the unidentified lines in the Atlas are neutrals, because most of the ionic lines added did not match the unidentified ones.

So far, the list of unidentified lines in our Atlas refers only to those in the centred T7 spectrum. Applying Hardy's line detection code from 2022 to other spectra would produce different lists of unidentified lines. By comparing these unidentified lines across various spectra, we can gain further insights into their origins. Below are some examples of comparisons between unidentified lines from different spectra and the insights they may provide:

- By comparing the unidentified lines in the centred spectrum of T7 with those in the spectra offset by $70''$ and $245''$ in the tail direction, we can obtain information on the nature of the molecule producing the lines. This concept has already been introduced in Section 4.1.4,

where the comparison was performed manually. However, having the list of unidentified lines compiled by the code simplifies the task. Indeed, it would be straightforward to check whether the emission band is only visible in the centred spectrum, only visible in the offset spectra, or visible in both. The presence of the emission band only in the centred spectrum suggests a neutral molecule, wheareas its presence in the offset spectra indicates an ion.

- By comparing the unidentified lines of different comets, for example, T7 and 16R2, additional information can be retrieved. It is known that the composition varies among comets. Also getting spectra of comets at different observing circumstances, for example at different distances from the Sun, may reveal new species that are not observed at certain heliocentric distances. However, care is required, as some emission bands might be hidden by lines from another molecule in a crowded spectral region.
- In addition to comparing the positions of the unidentified bands, it can be beneficial to analyse their fluxes. If the fluxes from unidentified lines are available for only one spectrum, it is difficult to detect any correlation between the unidentified bands. However, if the fluxes of different spectra are available, it is possible to detect certain patterns that assist the identification process. Indeed, the same ratio between the flux of one region and that of another region in several comets indicates that one molecule is responsible for the lines in both regions, as illustrated in Figure 6.1. Imagine, for example, that the spectral regions 4000–4020 Å and 5000–5020 Å both contain unidentified bands in the comets T7 and 16R2. Then a constant ratio $\frac{\text{flux in region } 4000-4020 \text{ Å}}{\text{flux in region } 5000-5020 \text{ Å}}$ between T7 and 16R2 would indicate that the bands at 4000–4020 Å and 5000–5020 Å originate from the same molecule.

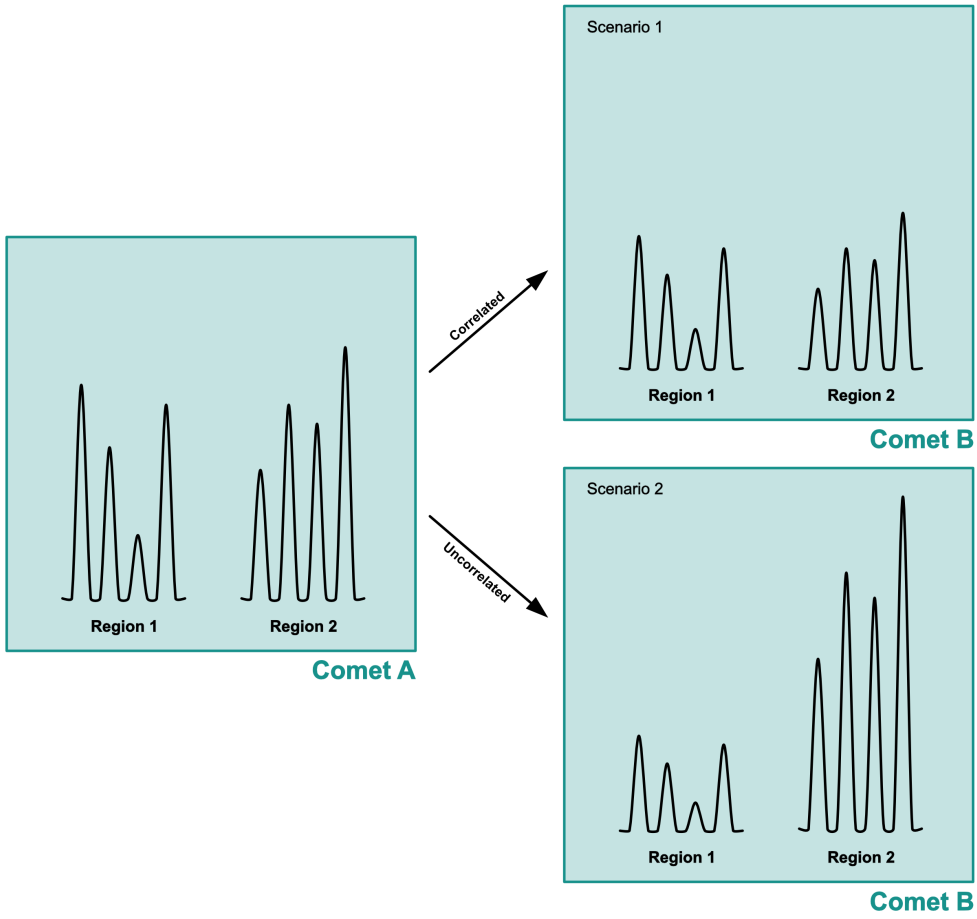


Figure 6.1: Two scenarios of variations in fluxes in spectral regions 1 and 2.

In the above, three ways in which the comparison of fluxes can help in the identification process are listed. Using a combination of these methods can further narrow down the range of possibilities for molecules, thus increasing the precision of identification.

Section 4 describes how a centred spectrum is optimal for detecting the emission lines of neutral molecules, whereas a spectrum offset towards the tail is more effective in revealing ionic emission lines. A potential approach to isolate ionic emission lines is to subtract the centred spectrum from the offset spectrum. This subtraction could help distinguish between ionic and neutral emission lines, thereby facilitating their identification. However, it is crucial to assess the feasibility of this method, as it may lead to potential inaccuracies or misinterpretations.

In our Atlas, the lines designated as "sky" correspond to emissions from molecules in Earth's atmosphere. However, such molecules also create absorption lines, called *telluric* lines. The Molecfit script is used to correct for these telluric lines; however, this script may overcorrect some absorption lines, resulting in false positives in the final cometary spectrum. To identify these false positives in the Atlas, it is beneficial to overplot the telluric spectrum on the cometary spectrum. The telluric spectrum can be obtained by two methods:

1. Subtract the spectrum corrected by the Molecfit script from the raw, uncorrected spectrum. Theoretically, this subtraction should yield the telluric spectrum. This method has already been applied by the ULiège team. However, it is not ideal, as it relies on the observed

cometary spectrum. Figure 6.2 illustrates the result of this subtraction, highlighting the prominent telluric lines.

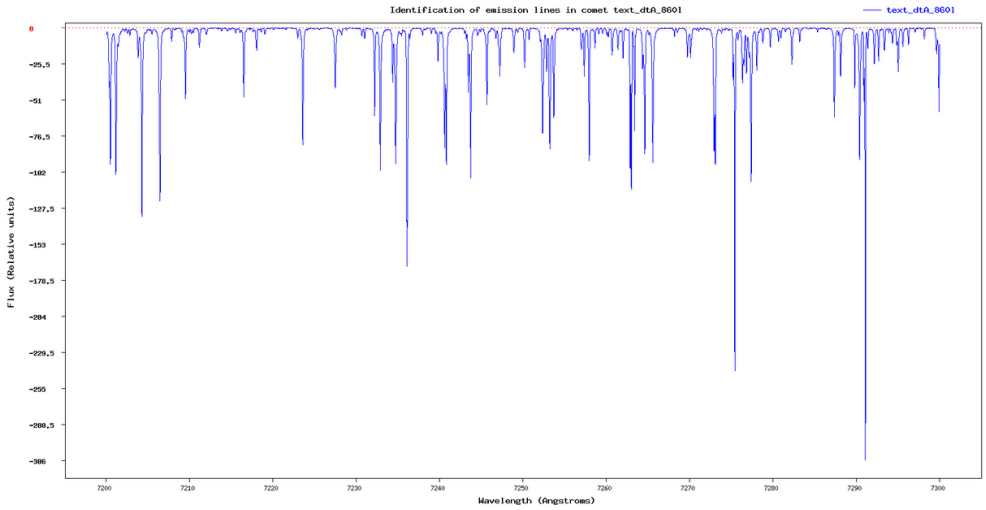
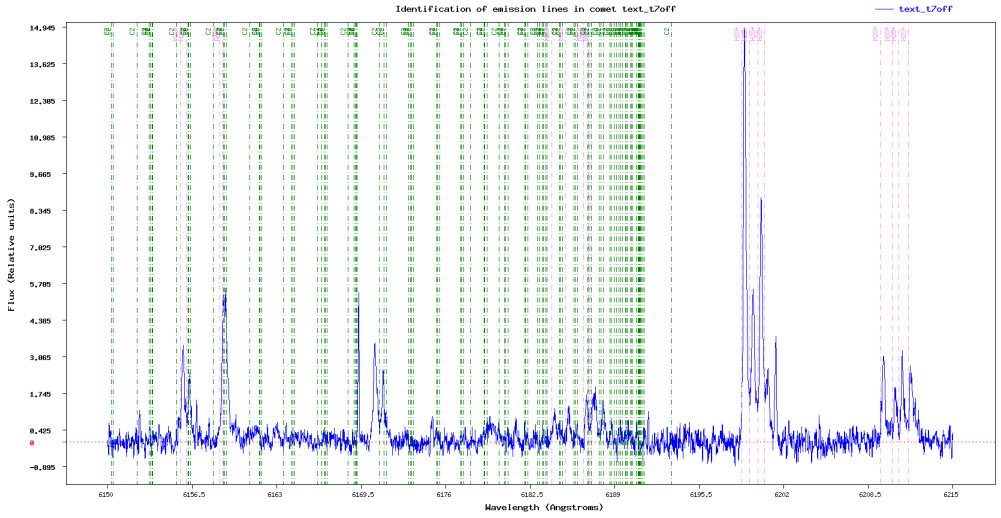


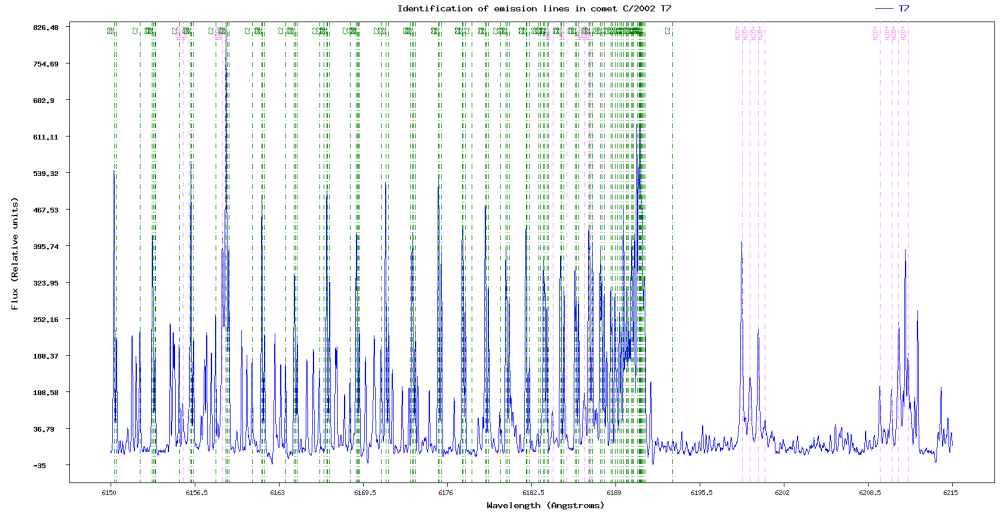
Figure 6.2: Telluric spectrum at 7200–7300 Å obtained by the ULiège team.

2. Modify the Molecfit script to directly retrieve the telluric spectrum that it has computed.

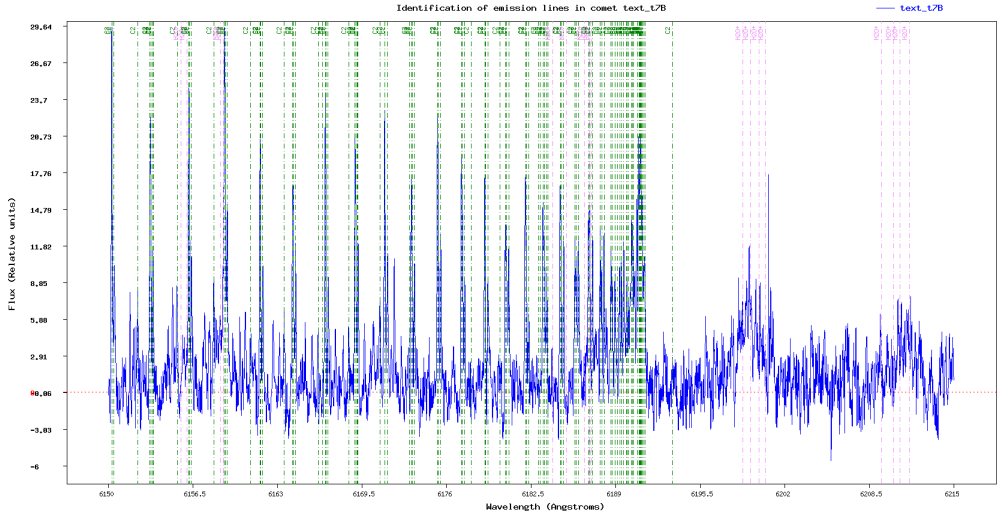
The ULiège team also possesses a spectrum of Comet T7 offset by 110″ in the direction of the Sun, opposite to the plasma tail. Consequently, spectra oriented towards the Sun contain fewer ionic lines than those centred on the coma. These spectra can be used in combination with other spectra to help determine whether the unidentified lines originate from ions or neutrals. Figure 6.3 shows the region of 6150–6215 Å in the spectra of T7 offset by 245″ towards the tail, centred, and offset 110″ towards the Sun. This region hosts C_2 (neutral) lines and H_2O^+ (ion) lines. In the centred spectrum (Figure 6.3b), both ion and neutral lines are visible. Upon moving the slit towards the Sun (Figure 6.3c), the ion lines become less defined and only form a weak bump, whereas the neutral lines remain clearly defined. In contrast, when the slit is moved in the direction of the tail (Figure 6.3a), the ion lines are intense and well-defined, whereas the neutral lines are very weak.



(a) Spectrum of T7 offset by 245″ in the tail direction. The H_2O^+ ion lines are strong and the CN neutral lines are weak.



(b) Centred spectrum of T7. The neutral lines are strong and ionic lines are faint.



(c) Spectrum of T7 offset by 110″ in the Sun direction. The ionic lines disappear.

Figure 6.3: Comparison of the emission lines in the region of 6150–6215 Å in the spectra of T7 with three different slit positions: 245″, centred, and 110″.

Currently, we have access only to high-resolution spectra of the tail of Comet T7. A valuable next step would be to obtain high-resolution spectra of the tails of other comets to facilitate a comparative analysis of their ionic emission lines. Achieving this would be challenging due to the difficulty of securing observation time on large telescopes, such as the Very Large Telescope (VLT). However, such a comparison could provide significant insights into the ionic compositions of cometary tails.

To tackle the problem of the missing formula for the fluorescence model, an alternative approach involves adapting an existing formula from a different transition to suit the model. This adaptation requires a deep understanding of the theoretical background used to establish these formulas. Once the transition intensities are accurately determined, the abundance of the molecule can be calculated. The ultimate objective is to develop fluorescence models for the major cometary molecules. The next molecule of interest for constructing a fluorescence model is CH, which shares similarities with NH. The primary modification required will involve adapting the constants used in the model.

Comets hold one of the keys to unlocking the mysteries of the formation of the Solar System and ultimately the origins of life on Earth. It's a privilege to contribute - modestly- to the advancement of scientific knowledge in this field.

Appendix A

Click here to access the tables containing the digitised wavelength (H_2O^+ lines) tables as well as my recalculation of λ_{air}, termed: "λair recalculated".

TABLE 1. Wavelengths and frequencies of transitions to $v_2" = 0$: (A) Σ subbands; (B) Δ subbands; (C) Π subbands; (D) Φ subbands [ν are in cm–1 and λ are in Å]

(A) Σ SUBBANDS

Branch	Rot. Trans.	L. Str[†] (Int.)	(0,5,0)-(0,0,0)		(0,7,0)-(0,0,0)		(0,9,0)-(0,0,0)		(0,11,0)-(0,0,0)		(0,13,0)-(0,0,0)		(0,15,0)-(0,0,0)	
			ν_{vac}	λ_{air} recalculated	ν_{vac}	λ_{air} recalculated	ν_{vac}	λ_{air} recalculated	ν_{vac}	λ_{air} recalculated	ν_{vac}	λ_{air} recalculated	ν_{vac}	λ_{air} recalculated
$^{P}R_{1,N-1}$	202-110	0.50			15314.594	6527.916	17274.961	5787.120	19285.299	5183.853	21340.471	4684.621	23423.506	4268.014
		(0.43)			15313.173	6528.521	17273.865	5787.487	19283.192	5184.420	21339.635	4684.805	/	/
	303-211	2.96	13424.825		15321.621	6524.922	17282.576	5784.570	19293.034	5181.775	21350.684	4682.380	23433.368	4266.218
		(2.18)	13424.825	7446.836	15320.868	6525.242	17281.614	5784.892	19288.966	5182.868	21350.147	4682.498	23432.481	4266.380
	404-312	1.46	/		15325.641	6523.210	17286.268	5783.335	19296.144	5180.940	21359.572	4680.432		
		(0.87)	13426.809	7445.735	15325.106	6523.438	17285.444	5783.611	/	/	21359.142	4680.526		
	505-413	5.71	13425.245	7446.603	15322.831	6524.406	17287.997 [b]	5782.756	19296.801	5180.763	21369.765	4678.199	23449.624	4263.261
		(2.66)	13425.245	7446.603	15323.442	6524.146	17286.537 [b]	5783.245	19301.666	5179.457	21367.465	4678.703	23448.323	4263.497
	606-514	2.28			15322.457	6524.566	17287.997 [b]	5782.756						
		(0.76)			15322.174	6524.686	17286.537 [b]	5783.245						
$^{P}Q_{1,N}$	101-111	4.48	13389.375	7466.552	15282.846	6541.477	17243.510	5797.676	19253.413	5192.438	21308.672	4691.612	23391.292	4273.892
		(4.26)	13387.204	7467.763	15281.982	6541.846	17242.655	5797.963	19254.557	5192.130	21307.748	4691.815	/	/
	202-212	2.47	13384.859	7469.071	15280.628	6542.426	17240.957	5798.534	19251.284	5193.013	21306.458	4692.100		
		(2.12)	13384.526	7469.257	15279.589	6542.871	17240.273	5798.764	19249.508	5193.492	/	/		
	303-313	10.28	13380.201	7471.672	15277.007	6543.977	17237.911	5799.559	19248.394	5193.792	21306.014	4692.197	23388.732	4274.360
		(7.58)	13380.449	7471.533	15276.459	6544.212	17237.250	5799.781	19244.559	5194.827	/	/	23388.085	4274.478
	404-414	4.33	13376.027	7474.003	15274.559	6545.026	17235.157 [b]	5800.485	19245.296	5194.628				
		(2.59)	13376.027	7474.003	15274.267	6545.151	17234.545	5800.691	19246.899	5194.196				
	505-515	15.5	13371.638	7476.456	15269.218	6547.315	17234.391	5800.743	19243.202	5195.194	21316.165	4689.963		
		(7.22)	13371.815	7476.357	15270.105	6546.935	17233.222	5801.137	19248.343	5193.806	21324.127	4688.212		
	606-616	5.91			15269.569	6547.165	17235.157 [b]	5800.485						
		(2.01)			15269.569	6547.165	17233.908	5800.906						
$^{P}P_{1,N-1}$	000-110	0.99			15261.345	6550.693	17222.001	5804.917	19230.962	5198.500	21287.207	4696.343		
		(0.99)			15260.435 [s]	6551.083	17221.102 [s]	5805.220	19230.092 [s]	5198.735				
	101-211	4.44	13339.723	7494.344	15233.187 *	6562.802	17193.886	5814.409	19203.752	5205.866	21259.090	4702.554	23341.650	4282.982
		(4.22)	13337.742	7495.457	15232.484 *	6563.104	17193.152	5814.657	19205.061	5205.511	21258.219	4702.747	23340.975	4283.106
	202-312	1.95	13305.096	7513.848	15200.837	6576.768	17161.198	5825.484	19171.498	5214.625	21226.708	4709.728	23309.694	4288.854
		(1.68)	13304.600	7514.128	15199.697	6577.262	17160.375	5825.763	19169.616	5215.136	21226.194	4709.842	/	/
	303-413	7.14	13267.127	7535.352	15163.894	6592.791	17124.831	5837.855	19135.354	5224.474	21192.951	4717.230	23275.580	4295.140
		(5.26)	13267.127	7535.352	15163.127	6593.125	17123.895	5838.174	19131.257	5225.593	21192.441	4717.344	23274.716	4295.299
	404-514	2.73	13227.070	7558.172	15125.607	6609.479	17086.192	5851.057	19096.350	5235.145	21159.561	4724.674		
		(1.64)	13226.778	7558.339	15125.074	6609.712	17085.343	5851.348	19097.662	5234.786	21159.026	4724.794		
	505-615	8.89	13185.105	7582.228	15082.801	6628.238	17047.932	5864.189	19056.806	5246.009	21129.749	4731.340	23209.569	4307.356
		(4.14)	13185.105	7582.228	15083.300	6628.018	17046.417	5864.710	19061.527	5244.709	21127.313	4731.886	23206.282	4307.966
		3.03			15045.205	6644.801	17010.816	5876.984						
		(1.03)			15044.794	6644.982	17010.103	5877.230						

[†] L. Str. = Calculated line strength, including nuclear statistical weight; Int. = Calculated relative line intensities for T=500 K and rotational energies of v'_2=7 level.
*Calculated from $^{P}Q_{1,N}$ (1) using known combination differences. Lines obscured by H_2. [b]Blended lines. [s]Satellite lines.

Figure 6.4: One of the ten tables that were manually transcribed from Lew 1976 [5]. The dark green columns indicate the recalculated wavelengths using Morton's formula [11]. Wavenumbers ν are in cm^{-1} and wavelengths λ are in Å.

Appendix B

The Python codes used for the development of the NH fluorescence model are presented in this appendix. They are divided into four sections:

1. Calculation of the energy levels
2. Calculation of the transition probabilities
3. Calculation of the population of the energy levels
4. Calculation of the intensity and plotting of the final model

Energy Levels

In [1]:

```
import numpy as np

#upload the constants from Ram Bernath (2010)
constSigma = np.loadtxt("constantes X3sigma.txt")
constPi = np.loadtxt("constantes A3Pi.txt")

#calculation of the energy levels of 3Sigma (formules from the site
https://physics.nist.gov/PhysRefData/MolSpec/Diatomic/Html/sec3.html)
F1Sigma = np.empty(shape = (11,3))   #N varies from 0 to 10 and v from 0 to 2
F2Sigma = np.empty(shape = (11,3))   #F2 and F3 do not have a level N=0, I will put "0" for these levels
F3Sigma = np.empty(shape = (11,3))
energiesSigma = np.empty(shape = (99,4))
for v in range(3):
    energiesSigma[v*33:v*33+33,0] = v
    energiesSigma[v*33:v*33+33,1] = np.repeat([0,1,2,3,4,5,6,7,8,9,10], 3)    #gives the N
    for N in range(11):
        energiesSigma[v*33+N*3, 2] = 1    #gives the F
        energiesSigma[v*33+N*3+1, 2] = 2
        energiesSigma[v*33+N*3+2, 2] = 3
        J1 = N+1   # for F1, J=N+1
        J2 = N
        J3 = N-1
        if N==0:
            F1Sigma[N,v] = constSigma[0,v]+constSigma[1,v]*(J1**2+J1+1)-constSigma[2,v]*
(J1**4+2*J1**3+7*J1**2+6*J1+2)-(3/2)*constSigma[6,v]-(1/3)*constSigma[9,v]-np.sqrt(((2*J1+1)*
(constSigma[1,v]-2*constSigma[2,v]*(J1**2+J1+1)-(1/2)*constSigma[6,v])-(3*constSigma[9,v])/(3*
(2*J1+1)))**2+4*J1*(J1+1)*(constSigma[9,v]/(2*J1+1))**2)
            F2Sigma[N,v] = 0
            F3Sigma[N,v] = 0
        else:
            F1Sigma[N,v] = constSigma[0,v]+constSigma[1,v]*(J1**2+J1+1)-constSigma[2,v]*
(J1**4+2*J1**3+7*J1**2+6*J1+2)-(3/2)*constSigma[6,v]-(1/3)*constSigma[9,v]-np.sqrt(((2*J1+1)*
(constSigma[1,v]-2*constSigma[2,v]*(J1**2+J1+1)-(1/2)*constSigma[6,v])-(3*constSigma[9,v])/(3*
(2*J1+1)))**2+4*J1*(J1+1)*(constSigma[9,v]/(2*J1+1))**2)
            F2Sigma[N,v] = constSigma[0,v]+(constSigma[1,v]-constSigma[2,v]*J2*(J2+1))*J2*(J2+1)-constSigma[6,v]
+ (2/3)*(constSigma[9,v])
            F3Sigma[N,v] = constSigma[0,v]+constSigma[1,v]*(J3**2+J3+1)-constSigma[2,v]*
(J3**4+2*J3**3+7*J3**2+6*J3+2)-(3/2)*constSigma[6,v]-(1/3)*constSigma[9,v]+np.sqrt(((2*J3+1)*
(constSigma[1,v]-2*constSigma[2,v]*(J3**2+J3+1)-(1/2)*constSigma[6,v])-(3*constSigma[9,v])/(3*
(2*J3+1)))**2+4*J3*(J3+1)*(constSigma[9,v]/(2*J3+1))**2)

        energiesSigma[v*33+N*3, 3] =  F1Sigma[N,v]  #gives the energies
        energiesSigma[v*33+N*3+1, 3] = F2Sigma[N,v]
        energiesSigma[v*33+N*3+2, 3] = F3Sigma[N,v]

np.savetxt("lower energies", energiesSigma)
```

the constants for 3Sigma are:

1. T
2. B
3. D
4. H
5. L
6. M
7. gamma
8. gammaD
9. gammaH

10. lamda

the constants for 3Pi are:

1. T
2. B
3. D
4. H
5. L
6. M
7. A
8. gamma
9. gammaD

In [15]:

```
#calculation of the nergy levels of 3Sigma (formulas from
https://physics.nist.gov/PhysRefData/MolSpec/Diatomic/Html/sec3.html)
F1Sigma = np.empty(shape = (11,3))   #N goes from 0 to 10 and v from 0 to 2
F2Sigma = np.empty(shape = (11,3))   #F2 and F3 do not have a level N=0, I will put "0" for these levels
F3Sigma = np.empty(shape = (11,3))
energiesSigma = np.empty(shape = (99,4))
for v in range(3):
    energiesSigma[v*33:v*33+33,0] = v
    energiesSigma[v*33:v*33+33,1] = np.repeat([0,1,2,3,4,5,6,7,8,9,10], 3)    #gives the N
    for N in range(11):
        energiesSigma[v*33+N*3, 2] = 1    #gives the F
        energiesSigma[v*33+N*3+1, 2] = 2
        energiesSigma[v*33+N*3+2, 2] = 3
        J1 = N+1   # for F1, J=N+1
        J2 = N
        J3 = N-1
        if N==0:
            F1Sigma[N,v] = constSigma[0,v]+constSigma[1,v]*(J1**2+J1+1)-constSigma[2,v]*
(J1**4+2*J1**3+7*J1**2+6*J1+2)-(3/2)*constSigma[6,v]-(1/3)*constSigma[9,v]-np.sqrt(((2*J1+1)*
(constSigma[1,v]-2*constSigma[2,v]*(J1**2+J1+1)-(1/2)*constSigma[6,v])-(3*constSigma[9,v])/(3*
(2*J1+1)))**2+4*J1*(J1+1)*(constSigma[9,v]/(2*J1+1))**2)
            F2Sigma[N,v] = 0
            F3Sigma[N,v] = 0
        else:
            F1Sigma[N,v] = constSigma[0,v]+constSigma[1,v]*(J1**2+J1+1)-constSigma[2,v]*
(J1**4+2*J1**3+7*J1**2+6*J1+2)-(3/2)*constSigma[6,v]-(1/3)*constSigma[9,v]-np.sqrt(((2*J1+1)*
(constSigma[1,v]-2*constSigma[2,v]*(J1**2+J1+1)-(1/2)*constSigma[6,v])-(3*constSigma[9,v])/(3*
(2*J1+1)))**2+4*J1*(J1+1)*(constSigma[9,v]/(2*J1+1))**2)
            F2Sigma[N,v] = constSigma[0,v]+(constSigma[1,v]-constSigma[2,v]*J2*(J2+1))*J2*(J2+1)-constSigma[6,v]
+ (2/3)*(constSigma[9,v])
            F3Sigma[N,v] = constSigma[0,v]+constSigma[1,v]*(J3**2+J3+1)-constSigma[2,v]*
(J3**4+2*J3**3+7*J3**2+6*J3+2)-(3/2)*constSigma[6,v]-(1/3)*constSigma[9,v]+np.sqrt(((2*J3+1)*
(constSigma[1,v]-2*constSigma[2,v]*(J3**2+J3+1)-(1/2)*constSigma[6,v])-(3*constSigma[9,v])/(3*
(2*J3+1)))**2+4*J3*(J3+1)*(constSigma[9,v]/(2*J3+1))**2)

        energiesSigma[v*33+N*3, 3] =  F1Sigma[N,v]  #gives the energies
        energiesSigma[v*33+N*3+1, 3] = F2Sigma[N,v]
        energiesSigma[v*33+N*3+2, 3] = F3Sigma[N,v]

np.savetxt("lower energies", energiesSigma)
```

In [ ]:

```
#calculation of theenergy differences (transitions)
indexS = []
indexP = []
for s,NS in enumerate(energiesSigma[:,1]):
    if energiesSigma[s,2] == 1:   #F1
        JS = NS+1
    elif energiesSigma[s,2] == 2:   #F2
```

```
        JS = NS
    else:                           #F3
        JS = NS-1
    for p,NP in enumerate(energiesPi[:,1]):
        if energiesPi[p,2] == 1:    #F1
            JP = NP+1
        elif energiesPi[p,2] == 2:   #F2
            JP = NP
        else:                           #F3
            JP = NP-1
        if np.abs(JS-JP) in [-1,0,+1]:    #selection rule
            if not JS==JP==0:
                if NS==0 and energiesSigma[s,2] == 1:       #attention au cas quand N=0 pour 3sigma, il ne faut
que prendre F1
                    indexS.append(s)
                    indexP.append(p)
                elif NS != 0:
                    indexS.append(s)
                    indexP.append(p)

print(len(indexS))
transitions = np.empty(shape=(len(indexS),7))
transitions[:,0] = energiesPi[indexP,0].astype(int)
transitions[:,1] = energiesPi[indexP,1]
transitions[:,2] = energiesPi[indexP,2]
transitions[:,3] = energiesSigma[indexS,0]
transitions[:,4] = energiesSigma[indexS,1]
transitions[:,5] = energiesSigma[indexS,2]
transitions[:,6] = energiesPi[indexP,3]-energiesSigma[indexS,3]

np.savetxt("transitions", transitions, fmt ='%1.3f', header = "V_pi , N_Pi , F_Pi , V_Sigma , N_Sigma , F_Sigma
, Delta_E(cm-1)" )
```

Format of the transitions file:

0	1	2	3	4	5	6
V_pi	N_Pi	F_Pi	V_Sigma	N_Sigma	F_Sigma	Delta_E(cm-1)

Transition probabilities

In [1]:
```
import numpy as np
from scipy import constants
from matplotlib import pyplot as plt
```

Get the Einstein coefficients A & B

In [2]:
```
energies = np.loadtxt("all_energies.dat", dtype="str")
#---transition proba A_v'v'' from Meier 1998---
A0=2.359e6  #A_v'
A1=2.169e6
A2=2.090e6
A_vv = np.array([[0.9938 , 0.006 , 0.0002], [0.035 , 0.9500 , 0.015], [0.0002, 0.05, 0.9498]])  #v' for rows and
v'' for columns
A_vv[0,:] *= A0
A_vv[1,:] *= A1
A_vv[2,:] *= A2


#---frequency of the band head (v*), meaning from the lowest levels of the considered v----
band_head = np.array([[29770.6092,32803.64825, 35639.64625],
                      [26645.03703,29678.07608,32514.07408],
                      [23675.73496,26708.77401, 29544.77201]]) #v'rows and v''columns

#---Einstein Aij and Bij coeff. for electronic transitions---
Aij = np.empty(shape=(273,273)).astype(object)
Bji = np.empty(shape = np.shape(Aij)).astype(object)
delta_E = np.empty(shape=(273,273))    #will contain the frequ/delta_E, will be useful later
for i in range(273):
    for j in range(273):
        Jprime = energies[i,2].astype(int)-energies[i,3].astype(int)+2   #formule for J=N-F+2
        Jsecond = energies[j,2].astype(int)-energies[j,3].astype(int)+2
        frequ_line = energies[i,5].astype(float)-energies[j,5].astype(float)    #delta_E (cm-1)
        #the case where we can have electronic transitions from the pi(94-273) to the sigma(1-93) state
        if i in np.arange(93,274) and j in np.arange(0,93):
            vprime = energies[i,1].astype(int)  #v'
            vsecond = energies[j,1].astype(int)  #v''
            frequ_head = band_head[vprime, vsecond]
            Avv = A_vv[vprime, vsecond]    #transition proba
            if Jprime-Jsecond in [-1,0,1]:    #selection rule
                #compute the Hönl-London coeff. (Herzberg)
                if Jsecond == Jprime+1:   #P line
                    S = (3/2)*Jprime
                elif Jsecond == Jprime:   #Q line
                    S = (3/2)*(2*Jprime+1)
                elif Jsecond == Jprime-1: #R line
                    S = (3/2)*(Jprime+1)
                else:
                    print("there is a problem")
                Aij[i,j] = 1*(frequ_line/frequ_head)**3*Avv*S/(2*Jprime+1)   #formula by Schleicher, A'Hearn
1982
            else:
                Aij[i,j] = 0         #forbidden by the selection rules
        #if we are not in the case of electronic transition from a pi to a sigma state
        else:
            Aij[i,j] = 0         #forbidden by the selection rules
        #add the einstein coeff for non-electronic transitions manually from Litvak 1992
        if i+1 == 2 and j+1 == 1:
```

```
            Aij[i,j] = 0.1328e-2
        if i+1 == 5 and j+1 == 4:
            Aij[i,j] = 0.1477e-1
        if i+1 == 9 and j+1 == 5:
            Aij[i,j] = 0.1064
        if i+1 == 9 and j+1 == 6:
            Aij[i,j] = 0.5327e-3
        if i+1 == 10 and j+1 == 7:
            Aij[i,j] = 0.1122
        if i+1 == 95 and j+1 == 94:
            Aij[i,j] = 0.1675e-12
        if i+1 == 105 and j+1 == 94:
            Aij[i,j] = 0.8666e-3
        if i+1 == 101 and j +1== 95:
            Aij[i,j] = 0.3069e-1
        if i+1 == 111 and j+1 == 95:
            Aij[i,j] = 0.3788e-3
        if i+1 == 105 and j+1 == 96:
            Aij[i,j] = 0.129e-1
        if i+1 == 104 and j+1 == 97:
            Aij[i,j] = 0.1137e-1
        if i+1 == 101 and j+1 == 100:
            Aij[i,j] = 0.2304e-9
        if i+1 == 111 and j+1 == 100:
            Aij[i,j] = 0.5823e-3
        if i+1 == 107 and j+1 == 101:
            Aij[i,j] = 0.1193
        if i+1 == 115 and j +1== 101:
            Aij[i,j] = 0.2817e-3
        if i+1 == 111 and j+1 == 102:
            Aij[i,j] = 0.19e-1
        if i+1 == 110 and j+1 == 103:
            Aij[i,j] = 0.1754e-1
        if i+1 == 111 and j+1 == 105:
            Aij[i,j] = 0.8401e-1
        if i+1 == 115 and j+1 == 106:
            Aij[i,j] = 0.41e-3
        if i+1 == 114 and j+1 == 107:
            Aij[i,j] = 0.4037e-3
        if i+1 == 112 and j+1 == 108:
            Aij[i,j] = 0.2631
        if i+1 == 113 and j+1 == 109:
            Aij[i,j] = 0.2648
        if i+1 == 114 and j+1 == 110:
            Aij[i,j] = 0.2542
        if i+1 == 117 and j+1 == 112:
            Aij[i,j] = 0.3023e-1
        if i+1 == 116 and j+1 == 114:
            Aij[i,j] = 0.5464
        if i+1 == 3 and j+1 == 1:
            Aij[i,j] = 0.1649e-1
        if i+1 == 5 and j+1 == 3:
            Aij[i,j] = 0.9591e-3
        if i+1 == 7 and j+1 == 4:
            Aij[i,j] = 0.2731e-1
        if i+1 == 10 and j+1 == 6:
            Aij[i,j] = 0.1434e-1
        if i+1 == 11 and j+1 == 8:
            Aij[i,j] = 0.3750e-3
        if i+1 == 11 and j+1 == 9:
```

```
            Aij[i,j] = 0.2843
        if i+1 == 13 and j+1 == 11:
            Aij[i,j] = 0.5842
        if i+1 == 97 and j+1 == 94:
            Aij[i,j] = 0.1515e-2
        if i+1 == 109 and j+1 == 94:
            Aij[i,j] = 0.2552e-2
        if i+1 == 103 and j+1 == 95:
            Aij[i,j] = 0.1009e-1
        if i+1 == 97 and j+1 == 96:
            Aij[i,j] = 0.2807e-7
        if i+1 == 110 and j+1 == 96:
            Aij[i,j] = 0.2961e-2
        if i+1 == 111 and j+1 == 97:
            Aij[i,j] = 0.2816e-2
        if i+1 == 103 and j+1 == 100:
            Aij[i,j] = 0.1534e-3
        if i+1 == 113 and j+1 == 100:
            Aij[i,j] = 0.3518e-2
        if i+1 == 109 and j+1 == 101:
            Aij[i,j] = 0.1563e-1
        if i+1 == 103 and j+1 == 102:
            Aij[i,j] = 0.3437e-7
        if i+1 == 114 and j+1 == 102:
            Aij[i,j] = 0.3414e-2
        if i+1 == 115 and j+1 == 103:
            Aij[i,j] = 0.3255e-2
        if i+1 == 107 and j+1 == 106:
            Aij[i,j] = 0.2389e-4
        if i+1 == 116 and j+1 == 106:
            Aij[i,j] = 0.1405e-3
        if i+1 == 117 and j+1 == 107:
            Aij[i,j] = 0.1807e-3
        if i+1 == 115 and j+1 == 108:
            Aij[i,j] = 0.2478e-1
        if i+1 == 114 and j+1 == 109:
            Aij[i,j] = 0.2340e-1
        if i+1 == 115 and j+1 == 111:
            Aij[i,j] = 0.2516
        if i+1 == 115 and j+1 == 113:
            Aij[i,j] = 0.9686e-6
        if i+1 == 117 and j+1 == 115:
            Aij[i,j] = 0.5425
        if i+1 == 4 and j+1 == 1:
            Aij[i,j] = 0.2121e-2
        if i+1 == 5 and j+1 == 2:
            Aij[i,j] = 0.1824e-1
        if i+1 == 6 and j+1 == 3:
            Aij[i,j] = 0.3642e-1
        if i+1 == 12 and j+1 == 8:
            Aij[i,j] = 0.1967e-1
        if i+1 == 11 and j+1 == 10:
            Aij[i,j] = 0.2457e-1
        if i+1 == 100 and j+1 == 94:
            Aij[i,j] = 0.3055e-1
        if i+1 == 110 and j+1 == 94:
            Aij[i,j] = 0.3082e-3
        if i+1 == 104 and j+1 == 95:
            Aij[i,j] = 0.7883e-3
        if i+1 == 98 and j+1 == 96:
```

```
            Aij[i,j] = 0.6885e-3
        if i+1 == 99 and j+1 == 97:
            Aij[i,j] = 0.9161e-3
        if i+1 == 104 and j+1 == 98:
            Aij[i,j] = 0.1131e-1
        if i+1 == 106 and j+1 == 100:
            Aij[i,j] = 0.1192
        if i+1 == 114 and j+1 == 100:
            Aij[i,j] = 0.2261e-3
        if i+1 == 110 and j+1 == 101:
            Aij[i,j] = 5588e-3
        if i+1== 104 and j+1 == 102:
            Aij[i,j] = 0.3036e-4
        if i+1 == 105 and j+1 == 103:
            Aij[i,j] = 0.3624e-4
        if i+1 == 105 and j+1 == 104:
            Aij[i,j] = 0.74e-7
        if i+1 == 109 and j+1 == 106:
            Aij[i,j] = 0.2235e-4
        if i+1 == 108 and j+1 == 107:
            Aij[i,j] = 0.1708e-4
        if i+1 == 109 and j+1 == 108:
            Aij[i,j] = 0.4715e-7
        if i+1 == 116 and j+1 == 108:
            Aij[i,j] = 0.3315e-2
        if i+1 == 117 and j+1 == 109:
            Aij[i,j] = 0.3158e-2
        if i+1 == 113 and j+1 == 112:
            Aij[i,j] = 0.5606e-7
        if i+1 == 116 and j+1== 113:
            Aij[i,j] = 0.2894e-1
        if i+1 == 117 and j+1 == 116:
            Aij[i,j] = 0.265e-8
        if i+1 == 9 and j+1 == 1:
            Aij[i,j] = 0.1076e-4
        if i+1 == 7 and j+1 == 3:
            Aij[i,j] = 0.9414e-2
        if i+1 == 8 and j+1 == 6:
            Aij[i,j] = 0.1259
        if i+1 == 9 and j+1 == 7:
            Aij[i,j] = 0.1933e-1
        if i+1 == 12 and j+1 == 10:
            Aij[i,j] = 0.2896
        if i+1 == 13 and j+1 == 12:
            Aij[i,j] = 0.2976e-1
        if i+1 == 102 and j+1 == 94:
            Aij[i,j] = 0.9498e-2
        if i+1 == 96 and j+1 == 95:
            Aij[i,j] = 0.1416e-2
        if i+1 == 108 and j+1 == 95:
            Aij[i,j] = 0.2502e-2
        if i+1 == 102 and j+1 == 96:
            Aij[i,j] = 0.1759e-1
        if i+1 == 103 and j+1 == 97:
            Aij[i,j] = 0.1789e-1
        if i+1 == 105 and j+1 == 99:
            Aij[i,j] = 0.180e-1
        if i+1 == 108 and j+1 == 100:
            Aij[i,j] = 0.148e-1
        if i+1 == 102 and j+1 == 101:
```

```
            Aij[i,j] = 0.1321e-3
        if i+1 == 112 and j+1 == 101:
            Aij[i,j] = 0.3461e-2
        if i+1 == 108 and j+1 == 102:
            Aij[i,j] = 0.9366e-1
        if i+1 == 109 and j+1 == 103:
            Aij[i,j] = 0.9463e-1
        if i+1 == 110 and j+1 == 104:
            Aij[i,j] = 0.8545e-1
        if i+1 == 112 and j+1 == 106:
            Aij[i,j] = 0.1971e-1
        if i+1 == 113 and j+1 == 107:
            Aij[i,j] = 0.2075e-1
        if i+1 == 110 and j+1 == 108:
            Aij[i,j] = 0.4795e-5
        if i+1 == 111 and j+1 == 109:
            Aij[i,j] = 0.4783e-5
        if i+1 == 111 and j+1 == 110:
            Aij[i,j] = 0.2758e-7
        if i+1 == 114 and j+1 == 112:
            Aij[i,j] = 0.1229e-5
        if i+1 == 115 and j+1 == 114:
            Aij[i,j] = 0.7223e-8

#creation of the B coeff
        #attention to the potential division by 0
        if frequ_line !=0:
            Bji[j,i] = (1/(8*constants.pi*6.626068e-27*frequ_line**3))*((2*Jprime+1)/(2*Jsecond+1))*Aij[i,j]
        else:
            Bji[j,i] = 0
        delta_E[i,j] = frequ_line
np.savetxt("Einstein A coeff matrix", Aij, fmt= "%s")
np.savetxt("Einstein B coeff matrix", Bji, fmt= "%s")
np.savetxt("wnbr_vacuum matrix", delta_E, fmt="%s")


#rotational transitions
#only for the Sigma state-->only until 93
#selection rule: delta_J=+-1
for i in range(93):
    for j in range(93):
        v1 = energies[i,1].astype(int)
        v2 = energies[j,1].astype(int)
        N1 = energies[i,2].astype(int)
        N2 = energies[j,2].astype(int)
        J1 = N1-energies[i,3].astype(int)+2
        J2 = N2-energies[j,3].astype(int)+2
        if v1 == v2:
            if J1 == J2+1:   #only transition J-->J-1
                if Aij[i,j] == 0:
                    Aij[i,j] = 1e-2

                    #formula from Arpigny 1964
#only for the Sigma state-->only until 93
#selection rule: delta_J=+-1
mu = 1.52913964*3.33564e-30  #from Brooke 2014, from r = 1.058, Attention que c'est en debey, je transforme en
cm
#constant from Ram&Bernath 2010, for v = 0,1,2
B0 = 16.343275263
B1 = 15.696418284
```

```
B2 = 15.050471062
h = 6.626068e-27  #in cgs units
for i in range(93):
    for j in range(93):
        v1 = energies[i,1].astype(int)
        v2 = energies[j,1].astype(int)
        N1 = energies[i,2].astype(int)
        N2 = energies[j,2].astype(int)
        J1 = N1-energies[i,3].astype(int)+2
        J2 = N2-energies[j,3].astype(int)+2
        if v1 == v2==0:
            if J1 == J2+1:   #only transition J-->J-1
                if Aij[i,j] == 0:
                    Aij[i,j] = (256*constants.pi**4*mu**2*B0**3)/(3*h)*(J2**4)/(J+1/2)   #formula from Arpigny
1964
        if v1 == v2==1:
            if J1 == J2+1:   #only transition J-->J-1
                if Aij[i,j] == 0:
                    Aij[i,j] = (256*constants.pi**4*mu**2*B1**3)/(3*h)*(J2**4)/(J+1/2)   #formula from Arpigny
1964
        if v1 == v2==2:
            if J1 == J2+1:   #only transition J-->J-1
                if Aij[i,j] == 0:
                    Aij[i,j] = (256*constants.pi**4*mu**2*B2**3)/(3*h)*(J2**4)/(J+1/2)   #formula from Arpigny
1964
```

Multiply B with the radiation density

In [14]:

```
#try the Donald Morton formula
def Morton(array):
    """
    Input: 2D array, wavenumbers in cm-1 in the vacuum
    output: 2D array, wavelength in A in the air
    """
    wvl_air = np.zeros(shape= np.shape(array))
    for i in range(len(array[:,0])):
        for j in range(len(array[0,:])):
            if array[i,j]!=0:
                wvl_vac = 1e8/array[i,j]    #wvl in Angstrom
                s = 1e4/wvl_vac
                n = 1+0.0000834254+0.02406147/(130-s**2)+0.00015998/(38.9-s**2)
                wvl_air[i,j] = (wvl_vac/n)
    return(wvl_air)

d = 0.68 #heliocentric distance of the comet in AU
v = 1 #heliocentric speed of the comet in m/s(v is negative if it gets closer to the Sun)

#convert the delta_energies (cm-1 in vacuum) into wavelength in AA in the air
wvl_air = Morton(delta_E)
wvl_air *= (1-v/constants.c)     #account for the Doppler shift due to the speed of the comet
np.savetxt("wvl_air.dat", wvl_air, fmt = "%s")

#import the solar spectra: 1st colum=wvl, 2nd column=radiation density
sol = np.loadtxt("sol_kur.dat")
sol[:,1] *=1e-20/(d**2) #(*e-20) to have the density inn erg.cm-3.Hz-1, need to divide by the heliocentric
distance^2 in AU

den_trans = np.zeros(shape = (273,273))   #will contain the intensities of the solar spectrum at the wavelength
of the trasitions
```

```
for i in range(273):
    for j in range(273):
        # Interpolate the density for the transition wavelength from the density of the solar spectrum
        den_interp = np.interp(wvl_air[i,j], sol[:, 0], sol[:, 1])
        den_trans[i,j] = den_interp

#multiply the Einstein B factor by the density
Bden = Bji*den_trans

np.savetxt("Bji_density", Bden, fmt = "%s")

#creating one file that combines the A and Brho coefficients
Pij = Aij
for j in range(273):
    for i in range(273):
        if Bden[i,j] !=0.0:
            Pij[i,j] = Bden[i,j]
np.savetxt("Pij_matrix.dat", Pij, fmt="%s")
```

population of the energy levels

In [2]:

```
# Using the formulas from Zucconi and Festou 1985
import numpy as np
Pij = np.loadtxt("Pij_matrix.dat")

#construction of Tij , formula 4
Tij = np.empty(shape=(273,273))
for i in range(273):
    for j in range(273):
        Pik=0
        if i==j:   #terme additionel
            for k in range(273):
                Pik+=Pij[i,k]
        Tij[i,j] = Pij[j,i] - Pik


#decomposition of the matrix according to 6
Txx = Tij[:93,:93]
Txe = Tij[:93,93:]
Tex = Tij[93:,:93]
Tee = Tij[93:,93:]

Tee_inv = np.linalg.inv(Tee)
matrix = Txx - np.matmul(np.matmul(Txe,Tee_inv),Tex)
#b column matrix & A matrix,formula 12
b=matrix[:92,92]
A=np.empty(shape=(92,92))
for i in range(92):
    for j in range(92):
        A[i,j] = matrix[i,j] - b[i]

#get the population, formula 15
A_inv = np.linalg.inv(A)
x = -np.matmul(A_inv,b)

#get n_Nx , formula 10
n_Nx = 1-np.sum(x)
#get n_X, the population of the lower energy levels
n_X=np.append(x, n_Nx)

np.savetxt("n_X.dat", n_X, fmt="%s")
```

Intensities and plot

In [1]:
```
import numpy as np
from scipy import constants
import matplotlib.pyplot as plt
plt.style.use('default')
from scipy import signal
from scipy.ndimage import gaussian_filter1d
from scipy.interpolate import interp1d
```

In [3]:
```
#get the populations from the higher energy levels (use formula 1 from Zucconi)
n_X=np.loadtxt("n_X.dat")
Pij = np.loadtxt("Pij_matrix.dat")
n_E=np.zeros(shape=(180,1))
for i in np.arange(93,273):
    num = 0
    denom = 0
    for j in range(93):
        num += n_X[j]*Pij[j,i]
        denom += Pij[i,j]
    n_E[i-93]=num/denom

print(np.sum(n_X))
print(np.sum(n_E))

#join n_X and n_E
n = np.append(n_X,n_E)
np.savetxt("n.dat", n, fmt="%s")
```

```
1.0
6.1886501628230235e-09
```

with satellite rays

In [ ]:
```
#Load data
wnbr_vacuum = np.loadtxt("wnbr_vacuum matrix")
wvl_air = np.loadtxt("wvl_air.dat")
I = np.zeros(shape=(273,273))  # intensity of the line

# Calculate the intensities
for i in range(93, 273):  # sweep the upper levels
    for j in range(93):  # sweep the lower levels
        I[i,j] = n[i] * Pij[i,j] * 6.626066e-27 * constants.c * wnbr_vacuum[i,j]

#reverse everything from a matrix into a 1d array
I = I.reshape(273*273)
wvl_air = wvl_air.reshape(273*273)

plt.figure(figsize=(15, 10))

# Sort the data based on wavelength
sorted_indices = np.argsort(wvl_air)
I = I[sorted_indices]
wvl_air = wvl_air[sorted_indices]
index = []
#only take the positive wavelength, meaning the emission lines
for i,j in enumerate(wvl_air):
    if j > 0:
        index.append(i)
wvl_air = wvl_air[index]
```

```
I = I[index]
#contstrain the wavelength in a reasonable range : 3000A-10000nm
difference_array = np.absolute(wvl_air-3000)
index_lower = difference_array.argmin()
difference_array = np.absolute(wvl_air-10000)
index_upper = difference_array.argmin()
wvl_air = wvl_air[index_lower:index_upper]
I = I[index_lower:index_upper]


# create a finer grid
step_length = 0.0001  #in Anstrom
fine_wvl = np.arange(wvl_air[0], wvl_air[-1],step_length)
without_I = np.empty (shape = (len(fine_wvl)))   #create an array with zero intensities
fine_wvl = np.append(fine_wvl,wvl_air)    #add the wvl we have to the new grid
with_I = np.append(without_I, I)     #add our intensities to the zero intensity array
index = np.argsort(fine_wvl),
fine_wvl = fine_wvl[index]
fine_I = with_I[index]

# Apply Gaussian filter
FWHM = 0.06*(1/step_length)   #we want the FWHM to be 0.06 Angstrom
convolution = gaussian_filter1d(fine_I, FWHM/2.35)

#plot limits
lower = 3330
upper = 3380
fig, ax = plt.subplots(2, 1, figsize=(15, 10))
ax[0].stem(wvl_air, I, markerfmt = " ")
ax[0].set_xlim([lower, upper])
ax[0].set_xlabel("wavelength (A)")
ax[0].set_ylabel("intensity")
ax[0].title.set_text("Unconcolved model output")

ax[1].plot(fine_wvl, convolution)
ax[1].set_xlim([lower, upper])
ax[1].set_xlabel("wavelength (A)")
ax[1].set_ylabel("intensity")
ax[1].title.set_text("Concolved model output")

plt.show()
```

delete the satellite rays:

In [ ]:

```
#delete the satellite rays to get a better picture of what is happening
energies = np.loadtxt("all_energies.dat", dtype="str")
wvl_air = np.loadtxt("wvl_air.dat")
delta_N = np.empty(shape=(273,273))
delta_N.fill(np.NaN)
delta_J = np.empty(shape=(273,273))
delta_J.fill(np.NaN)
for i in np.arange(93,273):
    for j in range(93):
        delta_N[i,j] = energies[i,2].astype(int)-energies[j,2].astype(int)
        delta_J[i,j] = (energies[i,2].astype(int)-energies[i,3].astype(int)+2)-(energies[j,2].astype(int)-
energies[j,3].astype(int)+2)
        if delta_J[i,j] in [-1,0,1]:   #selection rule
            if delta_J[i,j] != delta_N[i,j]:  #delete the satelite rays
                delta_J[i,j] = np.NaN
        else:
```

```
            delta_J[i,j] = np.NaN  #does not follow the selection rules

#---convulate and plot only the non-satellite rays---

# Calculate the intensities
I = np.zeros(shape = (273,273))
for i in range(93, 273):  # sweep the upper levels
    for j in range(93):  # sweep the lower levels
        if not np.isnan(delta_J[i,j]): #dont't take the satellite rays
            I[i,j] = n[i] * Pij[i,j] * 6.626066e-27 * constants.c * wnbr_vacuum[i,j]

#reverse everything from a matrix into a 1d array
I = I.reshape(273*273)
wvl_air = wvl_air.reshape(273*273)

plt.figure(figsize=(15, 10))

# Sort the data based on wavelength
sorted_indices = np.argsort(wvl_air)
I = I[sorted_indices]
wvl_air = wvl_air[sorted_indices]
index = []
#only take the positive wavelength, meaning the emission lines
for i,j in enumerate(wvl_air):
    if j > 0:
        index.append(i)
wvl_air = wvl_air[index]
I = I[index]
#contstrain the wavelength in a reasonable range : 3000A-10000nm
difference_array = np.absolute(wvl_air-3000)
index_lower = difference_array.argmin()
difference_array = np.absolute(wvl_air-10000)
index_upper = difference_array.argmin()
wvl_air = wvl_air[index_lower:index_upper]
I = I[index_lower:index_upper]

# create a finer grid
step_length = 0.0001  #in Anstrom
fine_wvl = np.arange(wvl_air[0], wvl_air[-1],step_length)
without_I = np.empty (shape = (len(fine_wvl)))   #create an array with zero intensities
fine_wvl = np.append(fine_wvl,wvl_air)    #add the wvl we have to the new grid
with_I = np.append(without_I, I)     #add our intensities to the zero intensity array
index = np.argsort(fine_wvl),
fine_wvl = fine_wvl[index]
fine_I = with_I[index]

# Apply Gaussian filter
FWHM = 0.06*(1/step_length)   #we want the FWHM to be 0.06 Angstrom
convolution = gaussian_filter1d(fine_I, FWHM/2.35)

#plot limits
lower = 3330
upper = 3380
fig, ax = plt.subplots(2, 1, figsize=(15, 10))
ax[0].stem(wvl_air, I, markerfmt = " ")
ax[0].set_xlim([lower, upper])

ax[1].plot(fine_wvl, convolution)
ax[1].set_xlim([lower, upper])
```

```
plt.show()
```

Compare with a real spectrum

In [ ]:

```
#compare the model to the T7 spectrum
import warnings
from datetime import date
from matplotlib.backends.backend_pdf import PdfPages
import datetime

from astropy import units as u
from astropy.io import fits
from astropy.table import QTable
from astropy.visualization import quantity_support
from astropy.wcs import WCS, FITSFixedWarning

from matplotlib import pyplot as plt
from matplotlib.pyplot import figure

import numpy as np

import scipy.signal as sg
from scipy.signal import argrelextrema

from specutils import Spectrum1D, SpectralRegion
from specutils.manipulation import noise_region_uncertainty

import time

plt.rcParams["figure.figsize"] = (20,10)
float_formatter = "{:.4f}".format
np.set_printoptions(formatter={'float_kind':float_formatter})
plt.rcParams.update({'font.size': 18})

f = fits.open('cz_t7A.fits')
flux=f[0].data
header = f[0].header
header[:1200]
quantity_support()   #Enables support for using Astropy quantities in Matplotlib plots.
del header['BUNIT']
fullspec=Spectrum1D.read(f)     #Creates a Spectrum1D object from the FITS file. The Spectrum1D class is part of
the Specutils library, and it is used for representing 1-dimensional spectra.
#plot the entire spectrum
plt.figure(figsize=(15,7))
region=fullspec[lower*u.AA:upper*u.AA]
plt.plot(region.spectral_axis,region.flux, label = "T7")
plt.plot(fine_wvl, convolution*1e24,label = "model", alpha = 0.75, color = "pink")
plt.xlim([lower, upper])
plt.ylim([0,75000])
plt.title("spectrum of T7" )
plt.xlabel("wavelengths(A)")
plt.ylabel("flux (arbitraty units)")
plt.legend()
```

References

[1] P. Hardy, "Atlas of cometary lines obtained from a high-resolution spectrum of comet C/2002 T7 (LINEAR)," Master's thesis, University of Liège, 2022.

[2] A. L. Cochran and W. D. Cochran, "A high spectral resolution atlas of comet 122P/de Vico," *Icarus*, vol. 157, no. 2, pp. 297–308, 2002.

[3] G. Cremonese, M. Capria, and M. De Sanctis, "Catalog of the emission lines in the visible spectrum of comet 153P/Ikeya-Zhang," *Astronomy & Astrophysics*, vol. 461, no. 2, pp. 789–792, 2007.

[4] P. Cambianica, G. Cremonese, G. Munaretto, M. T. Capria, M. Fulle, W. Boschin, L. Di Fabrizio, and A. Harutyunyan, "A high-spectral-resolution catalog of emission lines in the visible spectrum of comet C/2020 F3 (NEOWISE)," *Astronomy & Astrophysics*, vol. 656, p. A160, 2021.

[5] H. Lew, "Electronic spectrum of H2O+," *Canadian Journal of Physics*, vol. 54, no. 20, pp. 2028–2049, 1976.

[6] T. H. Brunet R., Ferras R., *Les mots de la Géographie, dictionnaire critique*. La Documentation Française, 2005.

[7] D. Woon, "Cometary molecules." https://www.astrochymist.org/astrochymist_comet.html. (accessed on 22/07/2024).

[8] P. Swings and L. Haser, *Atlas of representative cometary spectra*. University of Liège, Astrophysical Institute, 1956.

[9] A. B. J. S. C.Arpigny, F.Dossin, "Atlas of cometary spectra," 2012. hhttps://vela.astro.uliege.be/themes/solar/Comets/atlas.html, (accessed on 24/07/2024).

[10] J. Rahe, "Existing cometary data and future needs," *JPL Mod. Observational Tech. for Comets*, 1981.

[11] D. C. Morton, "Atomic data for resonance absorption lines. I-Wavelengths longward of the Lyman limit," *Astrophysical Journal Supplement Series (ISSN 0067-0049), vol. 77, Sept. 1991, p. 119-202.*, vol. 77, pp. 119–202, 1991.

[12] J. Zucconi and M. Festou, "The fluorescence spectrum of the CN radical in comets," *Astronomy and Astrophysics (ISSN 0004-6361), vol. 150, no. 2, Sept. 1985, p. 180-191.*, vol. 150, pp. 180–191, 1985.

[13] J. Romeo, "The long history of comet phobia," September 2020. https://daily.jstor.org/the-long-history-of-comet-phobia/, (accessed on 24/07/2024).

[14] G. Herbert-Brown, "Scepticism, superstition, and the stars: Astronomical angst in Pliny the Elder," *Bulletin of the Institute of Classical Studies. Supplement*, no. 100, pp. 113–126, 2007.

[15] L. A. Seneca, *Des questions naturelles*. Christofle Fourmy, 4BC-64AD.

[16] J. Uri, "955 years ago: Halley's Comet and the Battle of Hastings." https://www.nasa.gov/history/955-years-ago-halleys-comet-and-the-battle-of-hastings/, (accessed on 22/07/2024).

[17] B. de l'Observatoire de Paris. http://cometes.obspm.fr/fr/sitemap.html, (accessed on 31/07/2024).

[18] C. M. O. Alexander, K. D. McKeegan, and K. Altwegg, "Water reservoirs in small planetary bodies: meteorites, asteroids, and comets," *Space science reviews*, vol. 214, no. 1, p. 36, 2018.

[19] A. J. McKay and N. X. Roth, "Organic matter in cometary environments," *Life*, vol. 11, no. 1, p. 37, 2021.

[20] M. De Becker, "Astrochemistry, SPAT0020-2." Lecture notes, Université de Liège, 2023.

[21] F. L. Whipple, "The nature of comets," *Scientific American*, vol. 230, no. 2, pp. 48–57, 1974.

[22] M. Ory, *Chasseur de comètes*. De Boeck Supérieur, 2021.

[23] J. Oort, "The origin and dissolution of comets," *The Observatory, vol. 106, p. 186-193 (1986)*, vol. 106, pp. 186–193, 1986.

[24] G. H. Jones, M. M. Knight, K. Battams, D. C. Boice, J. Brown, S. Giordano, J. Raymond, C. Snodgrass, J. K. Steckloff, P. Weissman, *et al.*, "The science of sungrazers, sunskirters, and other near-sun comets," *Space Science Reviews*, vol. 214, pp. 1–86, 2018.

[25] E. Jehin, "Small Bodies of the Solar System." Lecture notes, Université de Liège, 2023.

[26] H. F. Levison and M. J. Duncan, "The long-term dynamical behavior of short-period comets," *Icarus*, vol. 108, no. 1, pp. 18–36, 1994.

[27] H. F. Levison, "Comet taxonomy," in *Completing the Inventory of the Solar System, Astronomical Society of the Pacific Conference Proceedings, volume 107, TW Rettig and JM Hahn, Eds., pp. 173-191.*, vol. 107, pp. 173–191, 1996.

[28] P. D. Feldman, "Comets," in *Springer Handbook of Atomic, Molecular, and Optical Physics*, pp. 1289–1298, Springer, 1982.

[29] J. Crifo and A. Rodionov, "Modelling the circumnuclear coma of comets: objectives, methods and recent results," *Planetary and space science*, vol. 47, no. 6-7, pp. 797–826, 1999.

[30] ESA, "Structure of a Comet." https://www.esa.in/t/ESA_Multimedia/Images/2023/11/Structure_of_a_comet, (accessed on 18/07/2024).

[31] C. Engrand, J. Duprat, E. Dartois, K. Benzerara, H. Leroux, D. Baklouti, A. Bardyn, C. Briois, H. Cottin, H. Fischer, *et al.*, "Variations in cometary dust composition from Giotto to Rosetta, clues to their formation mechanisms," *Monthly Notices of the Royal Astronomical Society*, p. stw2844, 2016.

[32] F. L. Whipple, "A comet model. I. the acceleration of Comet Encke," *Astrophysical Journal, vol. 111, p. 375-394 (1950).*, vol. 111, pp. 375–394, 1950.

[33] X. Hu, X. Shi, H. Sierks, J. Blum, J. Oberst, M. Fulle, E. Kührt, C. Güttler, B. Gundlach, H. U. Keller, *et al.*, "Thermal modelling of water activity on comet 67P/Churyumov-Gerasimenko with global dust mantle and plural dust-to-ice ratio," *Monthly Notices of the Royal Astronomical Society*, vol. 469, no. Suppl_2, pp. S295–S311, 2017.

[34] A. L. Cochran, A.-C. Levasseur-Regourd, M. Cordiner, E. Hadamcik, J. Lasue, A. Gicquel, D. G. Schleicher, S. B. Charnley, M. J. Mumma, L. Paganini, *et al.*, "The composition of comets," *Space Science Reviews*, vol. 197, pp. 9–46, 2015.

[35] J. Crovisier, "The molecular complexity of comets," *Astrobiology: Future Perspectives*, pp. 179–203, 2005.

[36] ESA, "Gases detected at Rosetta's Comet." https://www.esa.int/ESA_Multimedia/Images/2023/11/Gases_detected_at_Rosetta_s_Comet, (accessed on 11/08/2024).

[37] G. H. Stokes, F. Shelly, H. E. Viggh, M. S. Blythe, and J. S. Stuart, "The Lincoln Near-Earth Asteroid Research (LINEAR) program," *Lincoln Laboratory Journal*, vol. 11, no. 1, 1998.

[38] G. W. Kronk, "Cometography." www.cometoraphy.com, (accessed on 21/07/2024).

[39] N. Kaiser, H. Aussel, B. E. Burke, H. Boesgaard, K. Chambers, M. R. Chun, J. N. Heasley, K.-W. Hodapp, B. Hunt, R. Jedicke, *et al.*, "Pan-STARRS: a large synoptic survey telescope array," in *Survey and Other Telescope Technologies and Discoveries*, vol. 4836, pp. 154–164, SPIE, 2002.

[40] N. Biver, D. Bockelée-Morvan, G. Paubert, R. Moreno, J. Crovisier, J. Boissier, E. Bertrand, H. Boussier, F. Kugel, A. McKay, *et al.*, "The extraordinary composition of the blue comet C/2016 R2 (PanSTARRS)," *Astronomy & Astrophysics*, vol. 619, p. A127, 2018.

[41] C. Opitom, D. Hutsemékers, E. Jehin, P. Rousselot, F. Pozuelos, J. Manfroid, Y. Moulane, M. Gillon, and Z. Benkhaldoun, "High resolution optical spectroscopy of the N2-rich comet C/2016 R2 (PanSTARRS)," *arXiv preprint arXiv:1901.00657*, 2019.

[42] J. K. Harmon, M. C. Nolan, J. D. Giorgini, and E. S. Howell, "Radar observations of 8P/Tuttle: A contact-binary comet," *Icarus*, vol. 207, no. 1, pp. 499–502, 2010.

[43] H. Wilson, "Comet 8P/Tuttle - Hunter Wilson," 2007. https://skyandtelescope.org/online-gallery/comet-8ptuttle-2/, (accessed on 15/08/2024).

[44] N. Jet Propulsion Laboratory California Institue of Technology, "Small-Body Database Lookup." https://ssd.jpl.nasa.gov/sb/orbits.html, (accessed on 14/08/2024).

[45] C. B. K. PAN, "The Catalogue of Near-Parabolic Comets." http://ssdp.cbk.waw.pl/LPCs/near_parabolic_comets_catalogue.html, (accessed on 14/08/2024).

[46] S. Sinha, C. Jeyaseelan, G. Singh, T. Munjal, and D. Paul, "Spectroscopy—Principle, types, and applications," in *Basic Biotechniques for Bioprocess and Bioentrepreneurship*, pp. 145–164, Elsevier, 2023.

[47] M. H. Penner, "Basic principles of spectroscopy," *Food analysis*, pp. 79–88, 2017.

[48] G. Herzberg, *Molecular Spectra and Molecular Structure–I. Spectra of Diatomic Molecules*. Van Nostrand, New York, 1950.

[49] A. Zheng, "Atomic Term Symbol." https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Electronic_Spectroscopy/Spin-orbit_Coupling/Atomic_Term_Symbols, (accessed 15/07/2024).

[50] R. K. Hanson, "Lecture2: Rotational and Vibrational Spectra - Princeton University," 2018. Lecture Notes.

[51] A. Sanli, *Transition Dipole Moment and Lifetime Study of Na2 and Li2 Electronic States via Autler-Townes and Resolved Fluorescence Spectroscopy*. PhD thesis, Temple University, 2017.

[52] V. N. Ochkin, *Spectroscopy of low temperature plasma*. John Wiley & Sons, 2009.

[53] K. Haas, "11.3.1: Selection Rules." Duke University, LibreTextsChemistry.

[54] B. Valeur, "Introduction: On the origin of the terms fluorescence, phosphorescence, and luminescence," in *New trends in fluorescence spectroscopy: applications to chemical and life sciences*, pp. 3–6, Springer, 2001.

[55] J. R. Lakowicz, *Principles of fluorescence spectroscopy*. Springer, 2006.

[56] I. Newton, *Opticks, or a Treatise of the Reflections, Refractions, Inflections and Colours of Light. Also Two Treatises of the Species and Magnitude of Curvilinear Figures*. Royal Society, 1704.

[57] A. de La Baume Pluvinel and F. Baldet, "Spectrum of comet morehouse (1908 c)," *Astrophysical Journal, vol. 34, p. 89*, vol. 34, p. 89, 1911.

[58] J. Manfroid, D. Hutsemékers, and E. Jehin, "Iron and nickel atoms in cometary atmospheres even far from the Sun," *Nature*, vol. 593, no. 7859, pp. 372–374, 2021.

[59] H. Dekker, S. D'Odorico, A. Kaufer, B. Delabre, and H. Kotzlowski, "Design, construction, and performance of UVES, the echelle spectrograph for the UT2 Kueyen Telescope at the ESO Paranal Observatory," in *Optical and IR Telescope Instrumentation and Detectors*, vol. 4008, pp. 534–545, SPIE, 2000.

[60] Britannica, "Advances in auxiliary instrumentation." https://www.britannica.com/science/optical-telescope/Advances-in-auxiliary-instrumentation#ref481238, (accessed on 01/08/2024).

[61] Thorlabs, "Echelle Gratings." https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=1124, (accessed on 02/08/2024).

[62] L. Sbordone, C. Ledoux, A. Escorza, A. Kaufer, and S. Mieske, "Very Large Telescope Paranal Science Operations UV-Visual Echelle Spectrograph User manual," 2018. Karl-Schwarzschild Str.

[63] G. J. C. Ledoux, A. Modigliani, *Very Large Telescope Paranal Science Operations UVES data reduction cookbook*. EUROPEAN SOUTHERN OBSERVATORY.

[64] T. Eversberg and K. Vollmann, *Spectroscopic instrumentation*. Springer, 2015.

[65] C. Fairclough, "Analyzing White Pupil Échelle Spectrographs via Ray Tracing Simulation," January 2018. https://www.comsol.com/blogs/analyzing-white-pupil-echelle-spectrographs-via-ray-tracing-simulation, (accessed on 14/08/2024).

[66] R. L. Kurucz, "Solar Irradiance," 2005. http://kurucz.harvard.edu/sun/irradiance2005/irradthu.dat,(accessed on 25/03/2024).

[67] A. L. Cochran and A. J. McKay, "Strong CO+ and Emission in Comet C/2016 R2 (Pan-STARRS)," *The Astrophysical Journal Letters*, vol. 854, no. 1, p. L10, 2018.

[68] J. Loicq, "Spectroscopy in Astrophysics and Geophysics," 2023. Lecture notes, Université de Liège.

[69] J. Davoigneau, F. Le Guet Tully, L. Poupard, and F. Vernotte, *The Besançon Observatory – Keeping time by the stars*. Édition Lieux Dits, 2009.

[70] A. M. Fernando, P. F. Bernath, J. N. Hodges, and T. Masseron, "A new linelist for the A3π–X3σ- transition of the NH free radical," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 217, pp. 29–34, 2018.

[71] M. Litvak and E. R. Kuiper, "Cometary NH-Ultraviolet and submillimeter emission," *Astrophysical Journal, Part 1, vol. 253, Feb. 15, 1982, p. 622-633. NASA-supported research.*, vol. 253, pp. 622–633, 1982.

[72] A. Einstein, *Zur Quantentheorie der Strahlung*. Verlag u. Druck Gebr. Leemann, 1916.

[73] D. G. Schleicher and M. F. A'Hearn, "OH fluorescence in comets-Fluorescence efficiency of the ultraviolet bands," *Astrophysical Journal, Part 1, vol. 258, July 15, 1982, p. 864-877.*, vol. 258, pp. 864–877, 1982.

[74] R. Meier, D. Wellnitz, S. J. Kim, and M. F. A'Hearn, "The NH and CH bands of comet C/1996 B2 (Hyakutake)," *Icarus*, vol. 136, no. 2, pp. 268–279, 1998.

[75] M. Brown, A. Bouchez, A. Spinrad, and C. Johns-Krull, "A high-resolution catalog of cometary emission lines," *Astronomical Journal v. 112, p. 1197*, vol. 112, p. 1197, 1996.

[76] W. D. M. R. A'Hearn, M.F., "Spectra of C/1996 B2 (Hyakutake) for Multiple Offsets from Photocenter," 2015. https://pds.nasa.gov/ds-view/pds/viewCollection.jsp?identifier=urn%3Anasa%3Apds%3Agbo-kpno%3Ahyakutake_spectra&version=1.0, (accessed on 01/08/2024).

[77] U. Jorgensen, M. Larsson, A. Iwamae, and B. Yu, "Line intensities for CH and their application to stellar atmospheres.," *Astronomy and Astrophysics, v. 315, p. 204-211*, vol. 315, pp. 204–211, 1996.

[78] P. Swings and T. Page, "The Spectrum of Comet Bester (1947k).," *Astrophysical Journal*, vol. 111, 1950.

[79] S. Mrozowski, "On the 2 π u–2 π g Bands of CO2+. Part I to IV," *Physical Review*, vol. 72, no. 8, pp. 691–698, 1941-1947.

[80] A. Gicquel, J. M. Bauer, E. A. Kramer, A. K. Mainzer, and J. R. Masiero, "CO and CO2 Productions Rates of Comets Observed by NEOWISE within Year 1 of the Reactivated Mission," *The Planetary Science Journal*, vol. 4, no. 1, p. 3, 2023.

[81] S. Leach, "Electronic spectroscopy and relaxation of some molecular cations of cometary interest," in *Exploration of Halley's Comet*, pp. 195–200, Springer, 1988.

[82] S. Wyckoff, P. Wehinger, H. Spinrad, and M. Belton, "New molecular ions in spectra of comet P/Halley," in *ESLAB Symposium on the Exploration of Halley's Comet*, vol. 250, 1986.

[83] H. Smyth, "The emission spectrum of carbon dioxide," *Physical Review*, vol. 38, no. 11, p. 2000, 1931.

[84] P. Rousselot, E. Jehin, D. Hutsemékers, C. Opitom, J. Manfroid, and P. Hardy, "Comets $^{12}CO^+$ and $^{13}CO^+$ fluorescence models for measuring the $^{12}C/^{13}C$ isotopic ratio in CO^+," *arXiv preprint arXiv:2311.05700*, 2023.

[85] A. Bieler, K. Altwegg, H. Balsiger, A. Bar-Nun, J.-J. Berthelier, P. Bochsler, C. Briois, U. Calmonte, M. Combi, J. De Keyser, *et al.*, "Abundant molecular oxygen in the coma of comet 67P/Churyumov–Gerasimenko," *Nature*, vol. 526, no. 7575, pp. 678–681, 2015.

[86] R. Brown, R. Dittman, and D. McGilvery, "Remeasurement of the $\tilde{A}2\Pi i \rightarrow \tilde{X}2\Sigma+(0, 0)$ band in 12C16O+ by laser-induced fluorescence excitation spectroscopy," *Journal of Molecular Spectroscopy*, vol. 104, no. 2, pp. 337–342, 1984.

[87] M. F. A'Hearn, D. Schleicher, and P. Feldman, "The discovery of S2 in comet IRAS-Araki-Alcock 1983d," *Astrophysical Journal, Part 2-Letters to the Editor (ISSN 0004-637X), vol. 274, Nov. 15, 1983, p. L99-L103.*, vol. 274, pp. L99–L103, 1983.

[88] S. J. Kim, M. A'Hearn, D. Wellnitz, R. Meier, and Y. Lee, "The rotational structure of the B–X system of sulfur dimers in the spectra of Comet Hyakutake (C/1996 B2)," *Icarus*, vol. 166, no. 1, pp. 157–166, 2003.

[89] F. M. Gomez, R. J. Hargreaves, and I. E. Gordon, "A HITRAN-formatted UV line list of S2-containing transitions involving X3Σg-, B3Σu-, and B3Πu electronic states," *Monthly Notices of the Royal Astronomical Society*, vol. 528, no. 2, pp. 3823–3832, 2024.

[90] A. E. L. Sbordone, C. Ledoux, *Very Large Telescope Paranal Science Operations UV-Visual Echelle Spectrograph User manual*. EUROPEAN SOUTHERN OBSERVATORY.

[91] ExoMol, "High temperature molecular line lists for modelling exoplanet atmospheres." https://www.exomol.com/, (accessed on 01/08/2024).