EMMANUEL DEKEMPER

Development of an AOTF-based hyperspectral imager for atmospheric remote sensing

Ph. D. Thesis

Supervisors

XAVIER URBAIN	UCL
DIDIER FUSSEN	UCL/IASB-BIRA

Jury

ALAIN CORNET	UCL
MICHEL CRUCIFIX	UCL
MARTINE DE MAZIÈRE	IASB-BIRA
VITALY VOLOSHINOV	MSU

Université Catholique de Louvain
Louvain-la-Neuve, Belgium
22nd of October 2014

Contents

Remerciements

Comme étudiant, j'avais souvent tendance à penser qu'un travail ne pouvait être bien fait que si je m'en étais occupé personnellement. Ca pouvait peut-être agacer l'un ou l'autre copain de classe, mais généralement le groupe en était plutôt heureux. Si cette approche a assez bien fonctionné pendant mes études, l'exercice de la thèse m'a montré à quel point les échanges, la collaboration ou les travaux de prédécesseurs sont essentiels. Bref, j'ai pu mesurer à quel point j'étais redevable de la connaissance des autres pour arriver à apporter ma petite contribution originale. Je voudrais donc ici exprimer ma gratitude envers ceux qui m'ont aidé ou soutenu.

Cette thèse est au moins exceptionnelle pour une chose: sa durée! Jamais je ne conseillerais à un jeune de reproduire mon erreur: travailler dans le spatial, oui, mais s'inscrire en thèse dès les balbutiements d'une nouvelle mission, surtout pas! Il y a bien trop de décisions qui se prennent à plus haut niveau que pour espérer une évolution gentillement linéaire... Il aura fallu le double de la durée habituelle pour que je puisse présenter un travail complet et cohérent. Il fallait bien deux promoteurs pour y arriver.

Le goût de la recherche m'est venu avec le Dr. Xavier Urbain dès la dernière année de mes études de physique. Il a su prendre le temps de m'épauler lors de mon mémoire et m'a ouvert au monde délicat de la physique expérimentale au sein du laboratoire de physique atomique et moléculaire du département de physique de l'UCL (maintenant le pôle NAPS de l'IMCN du secteur des sciences et technologies). Un poste d'assistant m'offrait la possibilité de poursuivre des recherches dans ce domaine, mais après une année, un dilemme s'est présenté.

Le Dr. Didier Fussen m'avait en effet proposé de rejoindre son équipe pour travailler sur un projet d'instrument spatial qui démarrait alors. Un petit côté romantique m'a fait accepter, mais surtout fait rester quand

les vents technocratico-politiques nous condamnaient au sur-place. Avec Didier, j'ai vu toutes les facettes d'une mission spatiale: quand il fallait que les choses avancent, j'ai appris à travailler efficacement, quand la situation était bloquée, j'ai pu me replonger dans la recherche fondamentale ou prendre des initiatives et réaliser des mesures au labo ou sur le terrain. En accord avec mon côté autodidacte, Didier m'a souvent laissé le champ libre, mais je dois aussi le remercier pour les remarques décisives et les questions importantes qui n'ont pas manqué d'émailler nos discussions. J'ai sans doute appris beaucoup par moi-même, mais je me serais plusieurs fois perdu en chemin s'il ne m'avait pas recadré.

En cours de route, j'ai reçu des conseils, de l'aide ou des enseignements de toutes parts. Je dois évidemment beaucoup au Pr. Vitaly Voloshinov qui, au cours de ses visites en Belgique et au travers de notre correspondance, m'a beaucoup aidé à comprendre la physique complexe des AOTF, ces petits filtres qui forment une des clefs d'ALTIUS. Je suis aussi très reconnaissant de l'aide technique apportée par Jeroen, Bert et Jurgen. Sans eux, je n'aurais pas pu réaliser une seule mesure expérimentale. Je ne voudrais pas oublier Filip qui a toujours été patient lors de ses explications sur les méthodes d'inversion, ni Nicolas qui a apporté une bonne touche d'humour lors des premières années du projet et qui nous a permis de faire des expériences à bord d'un avion lors d'une mission mémorable à Toulouse.

Et puis, il y a l'excellente bande de collègues qui font que l'institut n'est pas seulement un lieu de travail, mais aussi un endroit très convivial. Parmi eux, je me dois d'insister sur le rôle important d'Alexis qui, bien que la disposition de nos bureaux fasse que l'on se tourne le dos 90% du temps, est toujours là pour échanger des points de vue, partager ses connaissances et détendre l'atmosphère (au sens figuré, pas de manière adiabatique bien entendu). Je m'en voudrais de ne pas citer Fabien pour l'enthousiasme et le dynamisme de cet incroyable organisateur de sorties, soirées ou w-e en tout genre auxquels j'ai pu participer en compagnie de Nuno, Romain et Yannick.

Enfin, si ce manuscript a pu être écrit et présenter quelque qualité, c'est en partie grâce à Gaia qui m'a offert un lieu de retraite à l'écart de l'effervescence de l'institut, et de la vie animée de la maison. Cette animation, j'ai été obligé de m'en soustraire quelques mois, et c'est mon épouse Alexandra qui a mené la barque de ma petite famille. J'espère qu'elle comprendra ce sacrifice et je la remercie pour tout le temps qu'elle y a consacré. Le texte en lui-même a bénéficié des remarques pointues des membres du Jury. Mais je suis également reconnaissant envers Charles, Filip et Nuno pour le temps consacré à la relecture de

certains chapitres, ainsi qu'à Pierre pour m'avoir aidé sur ma lacune de toujours: les statistiques.

Je voudrais terminer par quelques mots pour mes parents qui, je n'en doute pas, sont très fiers de cet accomplissement, et j'aimerais leur dire que je suis bien conscient qu'ils n'y sont pas pour rien, loin de là.

List of acronyms

Acronym	Meaning	Short description
AA	Azimuth Angle	Angle measured from North in the plane perpendicular to the zenith axis
ACE	Atmospheric Chemistry Experiment	Solar occultation experiment onboard SCISAT spacecraft
AIM	Aeronomy of Ice in the Mesosphere	NASA spacecraft
ALTIUS	Atmospheric Limb Tracker for the Investigation of the Upcoming Stratosphere	Spaceborne limb-scattering and occultation experiment
AO	Acousto-Optic	
AOCS	Attitude and Orbit Control System	Spacecraft system
AOTF	Acousto-Optical Tunable Filter	Large aperture tunable spectral filter based on elasto-optic effect
BASCOE	Belgian Assimilation System for Chemical ObsErvations	Assimilation model of atmospheric data
BELSPO	BELgian Science Policy Office	
BEO	Back-End Optics	Group of optical parts placed after the filter in a hyperspectral imaging instrument
BIRA	Belgische Instituut voor Ruimte Aeronomie	Scientific institution (same as IASB and BISA)
BISA	Belgian Institute for Space Aeronomy	Scientific institution (same as IASB and BIRA)
CCM	Chemistry Climate Model	Atmospheric model simulating physical and chemical processes

CFC	ChloroFluoroCarbon	Long-lived chemical compound with high ozone destruction capability
CIRA	COSPAR International Reference Atmosphere	Climatological atmospheric model
COSPAR	Committee On SPAce Research	Orgnization established by the International Council for Science with broad scopes related to spaceborne experimentation
CSA	Canadian Space Agency	
CTM	Chemistry Transport Model	Atmospheric model simulating atmospheric chemical processes and transport
DE	Diffraction Efficiency	AOTF efficiency parameter
DOAS	Differential Optical Absorption Spectroscopy	Method used to derive total or slant columns of absorbers from spectral radiance measurements
ECMWF	European Centre for Medium-range Weather Forecasts	
ENVISAT	ENVIronment SATellite	ESA Earth observation spacecraft
EO	Elasto-Optic	
ESA	European Space Agency	
EURECA	EUropean REtrievable CArrier	ESA multi-payload spacecraft
FEO	Front-End Optics	Group of optical parts placed before the filter in a hyperspectral imaging instrument
FOV	Field Of View	Full acceptance angle of an instrument
F-P	Fabry-Perot	Large aperture spectral filter based on light reflection in a cavity
FTS	Fourier Transform Spectrometer	Type of spectrometer based on light beams interferometry
FWHM	Full Width at Half Maximum	Measure of the spread of a Gaussian distribution
GAW	Global Atmospheric Watch	WMO programme
GOME	Global Ozone Monitoring Experiment	Family of spaceborne nadir sounders
GOMOS	Global Ozone Monitoring by Occultation of Stars	Stellar occultation instrument onboard ENVISAT spacecraft
IASB	Institut d'Aéronomie Spatiale de Belgique	Scientific institution (same as BIRA and BISA)
IGACO	Integrated Global Atmospheric Chemistry Observations	
IGY	International Geophysical Year	

IO3C	International Ozone Commission	
JAXA	Japan Aerospace eXploration Agency	
KDP	KH_2PO_4	Birefringent crystal used as AOTF material for UV spectral imaging
LORE	Limb Ozone Retrieval Experiment	Limb scattering instrument onboard Shuttle flight STS-87
LOS	Line Of Sight	Straight virtual line representing the viewing direction of an instrument
MACC	Monitoring Atmospheric Composition and Climate	Project supported by the European Union for Earth observation data assimilation by atmospheric models
MAP	Maximum A Posteriori	Solution to an inverse problem which maximizes the conditional probability of a given quantity given a set of measurements
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding	Limb emission instrument onboard ENVISAT spacecraft
MLS	Microwave Limb Sounder	Family of spaceborne limb emission instruments
MODTRAN	MODerate resolution atmospheric TRANsmission	Radiative transfer model
MQRE	Mean Quadratic Relative Error	Error index
MSU	Lomonosov Moscow State University	
NASA	National Aeronautics and Space Administration	
NDACC	Network for the Detection of Atmospheric Composition Change	Network of calibrated ground-based instruments
NIR	Near Infra-Red	Light spectrum of wavelengths ranging 800 nm to 2 μm
NLC	NoctiLucent cloud	High-altitude polar cloud (same as PMC)
O3CCI	Ozone Climate Change Initiative	Part of ESA Programme on Global Monitoring of Essential Climate Variables
ODS	Ozone Depleting Substance	
OMPS	Ozone Mapping and Profiler Suite	Nadir and limb instrument onboard SNPP platform (OMPS-LP refers to the Limb Profiler payload specifically)

OMI	Ozone Monitoring Instrument	Nadir sounder onboard Aura spacecraft
ORA	Occultation RAdiometer	Solar occultation instrument onboard EURECA platform
OSIRIS	Optical Spectrograph and InfraRed Imager System	Limb-scattering instrument onboard Odin spacecraft
PA	Principal Axes	Eigenvectors obtained from a PCA
PC	Principal Component	Eigenvalues obtained from a PCA
PCA	Principal Component Analysis	Mathematical method for identification of the natural eigenvectors of a dataset
POAM	Polar Ozone and Aerosol Measurement	Family of spaceborne solar occultation instruments
PMC	Polar Mesospheric Cloud	Cloud found around 80 km in extremely cold air masses (same as NLC)
PRNU	Pixel Response Non-Uniformity	Detector feature assessing the difference in sensitivity among the pixels
PROBA	PRoject for OnBoard Autnonomy	Family of micro-satellites developed by QinetiQ Space
PSC	Polar Stratospheric Cloud	Cloud found in the lower stratosphere of polar regions
RTM	Radiative Transfer Model	Physical model of light propagation through an absorbing and scattering medium
SAGE	Stratospheric Aerosol and Gas Experiment	Family of solar occultation spaceborne instruments
SAM II	Stratospheric Aerosol Measurement	Solar occultation instrument onboard Nimbus VII spacecraft
SBUV	Solar Backscatter Ultra-Violet radiometer	Family of spaceborne nadir sounders
SCD	Slant Column Density	Slant column of an atmospheric absorber generally expressed in molecules per cm^2
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY	Nadir, limb and solar occultation instrument onboard ENVISAT spacecraft
SES	Sun-Earth-Spacecraft	Plane defined by the center of mass of the three bodies
SI2N	SPARC/IO3C/IGACO-O3/-NDACC	Joint initiative for assessment of changes in ozone vertical distribution
SME	Solar Mesospheric Explorer	Limb scattering spaceborne instrument
SNPP	Suomi National Polar-orbiting Partnership	NASA Spacecraft

SNR	Signal-to-Noise Ratio	
SOFIE	Solar Occultation For Ice Experiment	Solar occultation instrument onboard the AIM platform
SOLSE	Shuttle Ozone Limb Sounding Experiment	Limb scattering instrument onboard Shuttle flight STS-87
SPARC	Stratosphere-troposphere Processes And their Role in Climate	Project of the World Climate Research Programme for atmospheric research
STF	Spectral Transfer Function	Spectral response of a filter to a delta function
SVD	Singular Value Decomposition	Mathematical method for eigenvector decomposition
SWIR	Short-Wave InfraRed	Light spectrum of wavelengths ranging 1 to 3 μm
TOA	Top Of Atmosphere	Refers to the altitude below which solar radiation starts being absorbed or scattered (generally 100 km)
TOMS	Total Ozone Mapping Spectrometer	Nadir instrument onboard Nimbus VII spacecraft
TP	Tangent Point	Point of the light trajectory at the shortest distance from the Earth surface
UCL	Université Catholique de Louvain	
UT/LS	Upper-Troposphere / Lower-Stratosphere	Atmospheric region located between 10 and 20 km
UV	Ultra-Violet	Light spectrum of wavelengths shorter than 400 nm
VIS	visible	Light spectrum of wavelengths ranging 400 to 800 nm
WMO	World Meteorological Organization	
ZA	Zenith angle	Angle measured from the local zenith

Introduction

In a pretty near future, presumably before 2020, the atmospheric science community will have to deal with a total lack of high resolution information on the vertical distribution of key atmospheric species. The golden age is clearly behind us, and only a handful of spaceborne instruments are doing this job today.

Yet, a good knowledge of the trace gases concentration profiles is important for many applications in atmospheric chemistry and transport modelling, the radiative energy budget, or the assessment of anthropogenic perturbations. One reason is that the abundance of radiation absorbing and emitting species not only determines the heating and cooling rates of the air mass, but also the depth of penetration of solar and terrestrial radiations. Knowing that, for instance, solar UV-VIS light can trigger chemical reactions through photodissociation or ionization of molecules, the variation of atmospheric species concentration with altitude therefore makes the chemistry and the thermal budget completely vertically-dependent as well.

Among other species, ozone has been granted particular attention for its strong absorption of UV light. When evidences of springtime depletion in the Southern polar stratosphere came, more efforts were made to monitor the depletion trend. In this case, total columns and profile data provide complementary information. Now that the Montreal protocol has been applicable for more than 25 years, all the experts are waiting for clear signs of recovery. A concomitant loss of the vertical information would be highly detrimental in this unique opportunity of validating amospheric models predictions.

The instrumental gap was anticipated by many experts in the early 2000's. But many of them were too busy with operating experiments, while at the same time, space agencies were reluctant on supporting a new round of atmospheric remote sensing instruments. The period 1990-

2005 had seen an unprecedent number of ambitious atmospheric profilers being launched and successfully operated. The Belgian Institute for Space Aeronomy (IASB-BIRA) was among the groups having spotted the forthcoming gap. At the same time, a relatively new remote sensing technique was demonstrated: the limb-scattering geometry.

This technique relies on the measurement of the atmospheric limb-scattered radiance with an instrument line-of-sight grazing the Earth surface at altitudes ranging from the cloud top to the lower mesosphere (approximately 50 km). Clearly, this observation geometry is suitable for the retrieval of vertical information, with the extra advantage of being capable of measurements everywhere in the day side of a low-Earth-orbit.

The accuracy of the retrievals is however dependent on the quality of the numerical radiative transfer model. Such models require high computing ressources due to the very large number of physical processes to be taken into account. With the increase of the computational power, concepts based on the limb scattering geometry started to be proposed and the first instrument completely dedicated to this type of measurements was launched in 2001 (OSIRIS onboard the Odin platform). Other missions followed with a partial commitment to limb scattering (SCIAMACHY, SAGE III) or even a full dedication (OMPS-LP).

The biggest challenge for this technique is the accurate determination of the tangent altitude of the line-of-sight. Depending on the concentration profile, pointing errors of a few hundred meters can induce serious biases (tens of percents). At the exception of OMPS-LP, all the hereabove cited instruments were based on grating spectrometers, asking for a scanning system to capture the entire scene. All of them met difficulties inherited from this approach.

One possible workaround could be to design a *true* hyperspectral imager. An instrument capable of taking instantaneous snapshots of the full limb. Uncertainties between the different atmospheric layers would therefore be removed at the price of acquisitions performed at a single wavelength at a time. Several technologies were capable of spectral imaging, but none of them was really suitable down to the UV domain. Only interference filters had been flown but their performance often proved to be poor.

Yet, having an instrument showing some sensitivity down to 250 nm would allow to measure ozone profiles up to the lower mesosphere. Other UV species like SO_2, BrO, or OClO fall into that range too.

Acousto-optical tunable filters (AOTF) were one of the possible candidates for playing the role of the spectral filter in the remote sensing

imager. But up to then, the dominant material to fabricate such devices was TeO_2, a crystal possessing a transparency edge at 350 nm. However, TeO_2 AOTF had already flown on different space experiments, which provided some confidence on the working principle. In 2004, two papers by Voloshinov and Gupta demonstrated the feasibility of an AOTF down to 200 nm.

Our team realized that this breakthrough was giving rise to a promising concept: a spaceborne multi-channel AOTF-based hyperspectral imager. The potential spectral range could cover the entire scattering domain, i.e. from UV to short-wave infrared (SWIR). On top of that, there were two more good news. First, the imager concept was not only suitable for the limb scattering geometry, it was also powerful for inertial occultations of celestial bodies (Sun, stars, planets), alleviating the need for a complex tracking system provided that a spacecraft capable of switching between both modes existed. Second, such spacecraft was indeed existing: a Belgian company had successfully developed a microsatellite with support from ESA and BELSPO. The satellite was named PROBA, and among other skills, it was excellent in attitude control.

In 2006, the Atmospheric Limb Tracker for the Investigation of the Upcoming Stratosphere (ALTIUS) was proposed to BELSPO by Didier Fussen. Soon after, the preliminary studies were starting.

We have been deeply involved in ALTIUS since its first steps. This dissertation is therefore strongly related to this project. More specifically, we have tried in this work to demonstrate that an AOTF-based hyperspectral remote sensing instrument is a valid concept in atmospheric science. Through several aspects, we have tried to strengthen this concept.

This work starts by outlying the importance of vertically-resolved atmospheric measurements. Although ALTIUS will be capable of multimode observations, we mainly focussed on the limb scattering technique. The problem of the radiative transfer in this geometry is exposed, and an inverse method for the retrieval of trace species concentration profiles from limb-scatter radiance is presented.

The second chapter is an attempt to show that one can also retrieve some quantitative information on the atmospheric state (here the pressure profile) by using image processing techniques only. In other words, we have shown that with an imager, but in absence of spectral information (or discrimination), it is possible to perform some kinds of measurements beyond the simple pattern detection. The application presented here is based on the analysis of the apparent flattening of the solar disk when observed through the atmosphere like an occultation instrument.

In the third chapter, we take the full potential of the ALTIUS concept by performing a detailed study of the expected retrieval performances for ozone and nitrogen dioxide. A formal description of the limb radiance simulations and of the inversion scheme is provided. The sensitivity of the retrieval to the major sources of uncertainty is tackled as well in a complete error budget for both species.

The fourth chapter is devoted to the description of the key element in the ATIUS instrumental concept: the AOTF. The fundamental equations of the acousto-optic interaction are derived and worked out to yield most of the required expressions for anyone interested in tayloring an AOTF to its needs (acceptance angle, passband, tuning range, diffraction efficiency, etc). Numbers are given for the well-known TeO_2 filter, but also for a KDP crystal that we could test for the UV channel.

Finally, the fifth chapter is dedicated to an outdoor campaign involving an optical breadboard of the ALTIUS visible channel. This real AOTF-based hyperspectral imager was used to detect the presence of NO_2 in the smoke exhausted by a waste incinerator smokestack. The description of the breadboard, the measurement strategy and all the steps involved in the data analysis to extract the NO_2 content from the spectral images are given. This ground-based experiment constitutes an important milestone on the way to demonstrating the capabilities of this new type of atmospheric remote sensing instrument.

Chapter 1

Atmospheric remote sensing

1.1 Milestones in atmospheric experimentation

When introducing a new instrumental concept, taking a look at the chronology of ideas and breakthroughs allows to distance oneself from the burden of the daily work. In atmospheric remote sensing like in many other fields, progresses were made by curious and talented people helped by always more advanced instruments. We propose to begin this dissertation by looking back at some remarkable milestones laid by previous pioneers. This short historical summary will differentiate between three main periods: the early developments of *in situ* measurements, the ground-based remote sensing and the advent of the spaceborne instrumentation.

The discussion will then move to specific aspects of our work. We will stress the importance of a class of data: those containing highly-resolved vertical information. With our instrumental concept precisely aiming at providing such kind of data, we will address two fundamental aspects of our baseline measurement strategy: the problem of light propagation through the atmosphere in limb observations, and the retrieval of trace gases concentration profiles from the limb-scattered light.

1.1.1 The *in situ* atmospheric measurements precursors

Back in the times when long-range instruments were lacking, atmospheric measurements consisted of *in situ* observations of basic meteorological variables, namely pressure (since the invention of the mercury-filled barometer by Torricelli in the 1640's), temperature (since Fahrenheit's fabrication of a mercury-filled glass tube in 1724), humidity (with reliable results obtained with Daniell's dew point hygrometer in 1820), and wind speed (with the relationship between the air speed and the ro-

tation speed of a cup anemometer first estimated by Robinson in 1847).

Aside from ground-based measurements and increasing use on trade ships, one had to wait for the 18th of July 1803 to see meteorological data being recorded at high altitude. This was on the occasion of the first balloon-borne scientific experiment: the flight of Robertson and Lhoest departing from Hamburg, and carrying a thermometer, a barometer, two birds and some equipment for an atmospheric electricity experiment[1]. Their flight lasted five hours and a half, they reached 7000 m, almost lost consciousness and suffered from eye bleeding. One bird died during the ascent, the other proved birds cannot fly at such altitudes...

Nowadays, *in situ* measurements extend well beyond the troposphere. It is with his first series of weather balloons launched in 1896 by Léon Teisserenc de Bort from the observatory of Trappes, France, that a pause in the temperature gradient was noticed at about 12 km. He named *stratosphere* the upper-lying region. Since then and up to the present day, balloon flights have been limited to the mid-stratosphere, at about 35 km. The development of liquid propelled rockets, successfully tested by Goddard in the U.S.A. in the late 1920's, allowed to break this boundary [97]. A second inversion of the temperature gradient was observed around 50 km, the lower limit of the *mesosphere*. Rockets pushed the *in situ* measurements up to the region where most of meteorites get destroyed by friction (80 km).

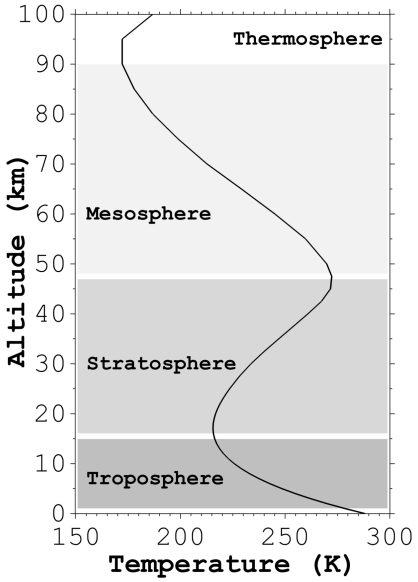

Figure 1.1 – Northern mid-latitude summer atmospheric temperature profile.

In 1957-58, the second International Geophysical Year (IGY) triggered a vast research effort, part of which devoted to studying proton flux energy spectrum and electron density in the ionosphere, the up-

[1]Details were reported by Arago and cited in the book *Wonderful balloon ascents* written by Fulgence Marion and published in 1870. Fulgence Marion is one of the recurrent pseudonyms of the famous French writer and editor Camille Flammarion.

per part of the atmosphere where ionization processes favours higher densities of ions and free electrons. This region spans the mesosphere (50-100 km) and the thermosphere (> 100 km) within which the temperature increases again and the air mixing becomes more heterogeneous. Under sponsorship of the IGY, rockets reached altitudes up to 200 km and provided insight into thermospheric composition. They opened a new field of increasing interest for radio communications: the space weather directly affected by solar activity.

Sputnik, the first artificial satellite launched on the 4th of October 1957, is the first of a long series of spaceborne instruments (a two-frequency transmitter in this particular case). Most of them have remote sensing capabilities, but some space missions like the four Cluster satellites are still performing *in situ* measurements. Launched two by two in summer 2000 by two Soyouz rockets, they are still probing the magnetosphere along their highly elliptical polar orbit.

1.1.2 The advent of ground-based atmospheric remote sensing

The attractiveness of studying the atmosphere from a remote location has numerous advantages: from the reduction of possible danger for the operators (volcano eruptions, high altitude air samples collection,...) to the broad coverage offered by a distant point of view. But as for the precursors of the meteorological measurements, intrigued scientists had to wait for technological breakthroughs before exploring out-of-reach air masses.

Before invention of any modern means of carrying information back and forth through high-altitude air parcels, the naked eye was often the only practical way for distant observations. Mysterious atmospheric processes were already observed for a long time, such as northern lights (named aurorae by French astronomer Pierre Gassendi (1592-1655)) or green flashes at sunset/sunrise. Another puzzling phenomenon is the appearance of noctilucent clouds (NLC) in late spring, beginning of summer. Scattering the Sun light not long after sunset, NLC's offer subtle glints (Fig. 1.2). But none of the contemporaries of the Krakatoa eruption in Indonesia in August 1883, which reportedly caused a significant increase of NLC occurrence long after ashes settled [67], could have guessed they were made of ice crystals formed on particles in suspension around 80 km.

Visible light is not the only information carrier, sound propagation can also be used to such ends. Worldwide reportings of anomalous sound

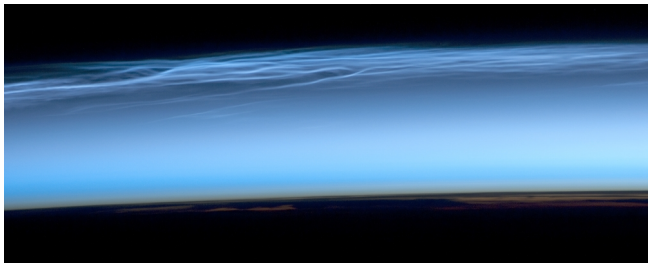

Figure 1.2 – Noctilucent cloud observed for the first time by a crew onboard ISS on July 2012.

propagation or sudden hearing of explosion are available. Perhaps one of the earliest mention of such phenomena lies in John Evelyn's diaries, recalling that on the 1st of June 1666, during an engagement between English and Dutch vessels on the Channel, firing cannons were heard in London, but not in the villages in between [17]. The physical explanation came with the pioneering works by Whipple [118] after World War One: low frequency acoustic waves can be reflected in the higher stratosphere where air temperature is back to ground levels.

The unexpectedly high temperature found at about 50 km had been postulated by Lindemann and Dobson using another type of remote measurements in the early 1920's [71]. Using a huge database of records of meteor appearance, brightness, velocity, path length and height, they inferred a number of conclusions in opposition to the accepted views of that time. The absence of a supposed hydrogen layer to the profit of a fairly constant nitrogen-oxygen mixing and much larger air density and temperature above 50 km than postulated constitute the main findings for what is our concern. Almost 30 years after the discovery of the tropopause, they had discovered the stratopause.

A few years later, Dobson made another major contribution to atmospheric science. In the second half of the 1920's, while at Oxford, he designed and built a photographic (and later, photoelectric) spectrophotometer with the objective of developing a network of instruments for the daily measurement of total ozone columns [39]. Lindemann thought there might be a correlation between pressure and ozone content, hence the need for a dissemination of identical instruments. At that time, the inconsistency between the short part of the solar spectrum and a black body equivalent was attributed to absorption in the atmosphere. Hartley had measured in 1880 the strong absorption of UV light by ozone and postulated its presence among the air trace species. Confirmation came with the experiments by Fabry and Buisson in 1913 [43]. Dob-

son's instrument followed their measurement principle: matching the difference between two UV lines of the scattered solar spectrum with a corresponding amount of ozone along the line of sight.

Since then, no better technique has been found for this purpose. Calibrated Dobson photometers provide daily total column ozone measurements at scattered locations all around the world. For example, the one situated at the Halley Bay Observatory in Antarctica saw the first evidence for ozone depletion in 1985 [44].

Ground-based atmospheric remote sensing has now evolved to encompass more species. Depending on their spectral signature, different measurement principles are used, mainly based on electromagnetic waves absorption, scattering or emission.

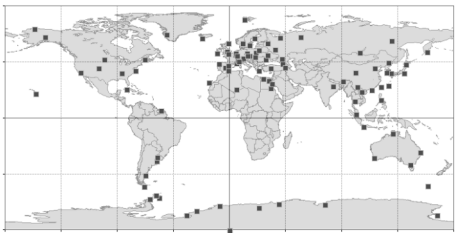

Figure 1.3 – Location of the GAW ground stations having delivered ozone data in 2013.

The need for trends assessment capabilities led to the creation of international organizations responsible for supervising quality, supporting teams and archiving data, such as the Global Atmosphere Watch (GAW) program from World Meteorological Organization (WMO) or the Network for the Detection of Atmospheric Composition Change (NDACC). The number of measurement stations has grown under their supervision, and the geographical coverage has spread. Fig. 1.3 shows the location of the GAW stations having delivered ozone data during year 2013. As expected, they are not evenly distributed, which complicates the assimilation of the measurements in large scale atmospheric models. The advent of satellite-based instruments allowed to fill these gaps.

1.1.3 Spaceborne atmospheric instruments

Spaceborne atmospheric remote sensing is almost as old as space exploration itself: about two years after the breakthrough of Sputnik 1 in October 1957, Explorer 7 (NASA) was performing global measurements of the radiative balance of the Sun-Earth system with bolometers. Besides intelligence applications and space environment studies, a lot of efforts were invested on developing reliable operational meteorological instruments. Their main initial contributions were the tracking of cloud patterns to derive wind and pressure fields.

For what is our main concern, namely the remote sensing of atmospheric constituents, one had to wait for the Nimbus IV satellite carrying the Backscatter Ultra-Violet radiometer (BUV) instrument to see total ozone columns being retrieved. But it survived for only one year and it is with Nimbus VII that long-term global measurements could be obtained. This platform hosted several instruments, among them the Solar Backscatter Ultra-Violet radiometer (SBUV) and the Total Ozone Mapping Spectrometer (TOMS). As the lifetime of the mission covered more than 15 years, it provided one of the longest series of records ever, and witnessed the formation and the spreading of the ozone hole year after year. Their measurement principle was similar to the method developed by Dobson: adjust the amount of ozone to the radiance levels observed at a number of selected UV wavelengths.

As the optical thickness varies strongly in the 260-330 nm region due to the steepness of the ozone absorption cross-section, measurements at different wavelengths see the backscattered radiance originating from different altitudes. Nadir-looking instruments like TOMS, SBUV and more recently GOME (Global Ozone Monitoring Experiment), OMI (Ozone Monitoring Instrument), SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) and others have therefore the capability to retrieve some information on the ozone vertical distribution. This is however generally limited to the stratosphere, with a coarse vertical resolution (several km).

The highest profile resolution and accuracy can be obtained by the occultation technique, i.e. the observation of a light source as it sets or rises behind the atmosphere. It is again onboard the Nimbus VII platform that the pioneering experiment was flown: SAM II (Stratospheric Aerosol Measurement). Due to the orbit, it provided vertical profiles of stratospheric aerosols above polar regions only. The measurements consisted of atmospheric transmittance at 1 micron. SAM II paved the way for, among others, the SAGE I, II and III instruments (Stratospheric Aerosol and Gas Experiment) which served extensively as references in ozone, nitrogen dioxide and aerosol validation exercises. All but one occultation instruments were designed to observe the Sun. GOMOS (Global Ozone Monitoring by Occultation of Stars) onboard ENVISAT looked at stars.

The excellent performance of solar occultation instruments is balanced by their limited coverage (one sunset and one sunrise per orbit). The limb scattering technique is an attempt to combine good vertical resolution with global geographical coverage. It relies on the measurement of the solar radiance scattered by the Earth atmospheric limb. The

drawback here is the uncertainty on the exact light path which complicates the numerical model, especially at lower altitudes where multiple scattering can occur. It was first applied by the SME instrument (Solar Mesosphere Explorer) in the 1980's for the upper atmosphere, but was proven to be a relevant technique (at least for ozone profiles) down to the lower stratosphere with the SOLSE (Shuttle Ozone Limb Sounding Experiment) and LORE (Limb Ozone Retrieval Experiment) instruments during the Shuttle flight STS-87. Since then, however, only three other experiments have used this method to deliver operational-grade products: OSIRIS (Optical Spectrograph and Infrared Imaging System) launched in 2001 and still operating, SCIAMACHY from 2002 to 2012, and OMPS (Ozone Mapping and Profiler Suite) in orbit since 2011. GOMOS and SAGE III processed some limb-scatter spectra as well, but more as a demonstration and for research purpose.

1.2 Importance of vertically resolved information

The number of sounders with accurate profiling capabilities is clearly on a negative trend compared to nadir-looking instruments which benefit from a sustained development (see Fig. 1.4 computed from the WMO database). While it does not reflect the seriously concerned recommendations of the scientific community [123], a future in which a total lack of high resolution information on the vertical distribution of key atmospheric species is becoming very likely, if not totally certain.

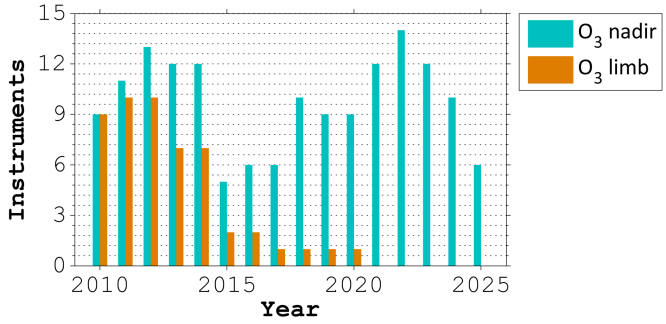

Figure 1.4 – Number of sounders capable of retrieving good ozone total columns (nadir) or profiles (limb). The data have been collected from the WMO database (`http://www.wmo-sat.info/oscar/`). Only instruments with at least *high* relevance have been considered, amongst the missions from NASA/NOAA (USA), JAXA (Japan) and ESA/third party (EU).

In this section, however, we address the importance of the knowledge of the trace gases concentration vertical distribution. The attention here will be focused on ozone. Despite being probably the most studied of all atmospheric species, it still attracts a significant part of efforts in the atmospheric science community. This is for mainly two reasons.

First, thanks to its strong absorption in the UV, stratospheric ozone is not only a natural barrier against harmful radiations affecting organic life at the surface, it is also the driver for the solar heating rates observed in the stratosphere and the lower mesosphere [16]. Indeed, photodissociation of ozone can occur at wavelengths shorter than 1180 nm [80], which roughly encompasses 78% of the total solar radiative energy at the top of the atmosphere. The dissociation yields two products:

$$O_3 + h\nu \longrightarrow O^* + {O_2}^*, \tag{1.1}$$

with the atomic and molecular oxygen excited states depending on the incident photon energy. It is essentially these two products which cause the heating as they react with other chemical species in exothermic processes. Depending on the altitude and the solar zenith angle, heating rates up to 10K/day can be attributed to the solar radiation absorption by ozone. Therefore, accurate knowledge of the ozone field (through measurements and modelling) is of utmost importance in climate research.

Second, in the pursuit of the response to the anthropogenic ozone depletion, there are currently immense efforts undertaken to detect ozone trends in reprocessed times series of many past and current sounders. After nearly two decades of continuous decrease and more than one decade of stagnation, a slight recovery is now expected. But it still lacks a statistically significant confirmation. In the following sections, we provide more insight into these entangled topics.

1.2.1 Stratospheric ozone trends

Questions remain open concerning the stratospheric ozone long-awaited recovery, and how well its altitude and latitude distribution is captured by models. In the 1970's it was known that the chlorofluorocarbons (CFC's) have a strong potential of ozone depletion. These molecules had become widely used as refrigerants, thanks to their non-reactivity. Once released in open air, their long lifetime allows them to diffuse slowly towards higher altitudes, eventually reaching the stratosphere where UV photons can photodissociate these molecules [86]. An example of reaction is $CFCl_3 + h\nu \longrightarrow CFCl_2 + Cl$. The released Cl atom triggers

a catalytic chain of reactions with O_3 and O:

$$Cl + O_3 \longrightarrow ClO + O_2$$
$$ClO + O \longrightarrow Cl + O_2.$$

The single chlorine atom can keep destroying ozone molecules for a long time if not converted into inactive species like $ClONO_2$ or HCl. With the general chemical understanding of that time, a worldwide decrease of O_3 by up to 10% around 40 km was expected, well above the ozone layer (25-30 km) [30].

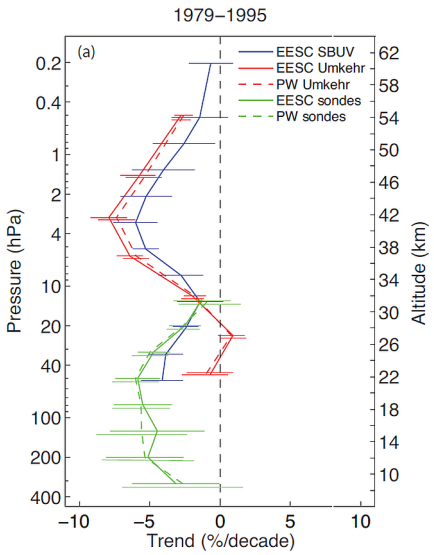

Figure 1.5 – Vertical profile of ozone trends over Northern midlatitudes estimated from ozonesondes, Umkehr, and SBUV(/2) measurements for the period 1979-1995. (Reproduced from [119])

When evidences of a much stronger ozone depletion over Antarctica during spring were published in 1985 [44], the general understanding got shaked. Observed in the time series of ozone total columns at Halley Bay (76° S) in 1984, the trend was already quite significant, such that early signs of the *ozone hole* could be traced back to the mid 1970's. A coordinated worldwide action against the halogen ozone depleting substances (ODS) was quickly initiated (the UN Vienna convention on the protection of the ozone layer in 1985 and the subsequent protocol on the chlorofluorocarbons in Montreal in 1987). In parallel, the ozone monitoring and the search for a better explanation for its depletion became a top priority for many groups. The observed loss over Antarctica in October was not only much more severe than expected, it was also occuring at a lower altitude. Until 1995, the negative trend was measured both in total columns and profiles (see Fig. 1.5), with significant differences in latitude and altitude [119].

After 1995, following the CFC's progressive phase out, the amounts of ODS in the atmosphere have levelled off and even started to decrease. Until now and besides its natural variability, stratospheric ozone is show-

ing only weak signs of recovery: total columns show worldwide positive trends, but the statistical significance is not yet reached everywhere [25]. The amplitude of the trend is also quite small (1-2% per decade) compared to pre-1995 figures. Moreover, zonal cross sections show a much more complex pattern: positive trends are found in the upper stratosphere, but with a strong hemispheric asymmetry, whereas below the ozone maximum, negative trends are still observed [14].

1.2.2 Dispersion of the measurements

Currently, the anticipated trend is smaller than 2%/decade. While this may be distinguished in the time-series of well-calibrated ground-based instruments, such a small slope is quite challenging for any kind of profiler, even for ozonesondes (accuracy and precision within 5-10%, [38]).

Among the satellite instruments, SAGE II was often depicted as a reference for intervalidation exercises. This is due to its unique lifetime — 20 years of continuous operations (1984-2005) — and also to its measurement technique, the solar occultation, which is self-calibrating, offers high SNR, and suffers less from pointing uncertainties. However, it was shown that despite the quality of its long dataset, no trend smaller than 2.9%/decade could be detected with statistical significance [62].

Recently, initiatives have been undertaken to reassess the data quality of many spaceborne limb sounders (the SPARC data initiative, the SI^{2}N initiative, the ozone climate change initiative (O3CCI),...) with the final objective of building a consistent dataset capturing the state of the ozone field in four dimensions (latitude, longitude, altitude and time). However, the early results of these activities shine a crude light on the capability of unearthing the expected trend from such datasets [100].

Fig. 1.6 illustrates the biases which exist between the ozone climatologies (monthly-averaged zonal means) derived from the measurements of 16 limb sounders. Clearly, the agreement is poor among the instruments, especially at the lowest altitudes. The figure, however, represents a worst case scenario, as the latitude bands illustrated here (high-latitudes and tropics) often contain the most perturbated ozone profiles. The agreement is better at the mid-latitudes for instance, such that the 1σ spread of all the instruments falls within $\pm 5\%$ at altitudes close to the ozone maximum.

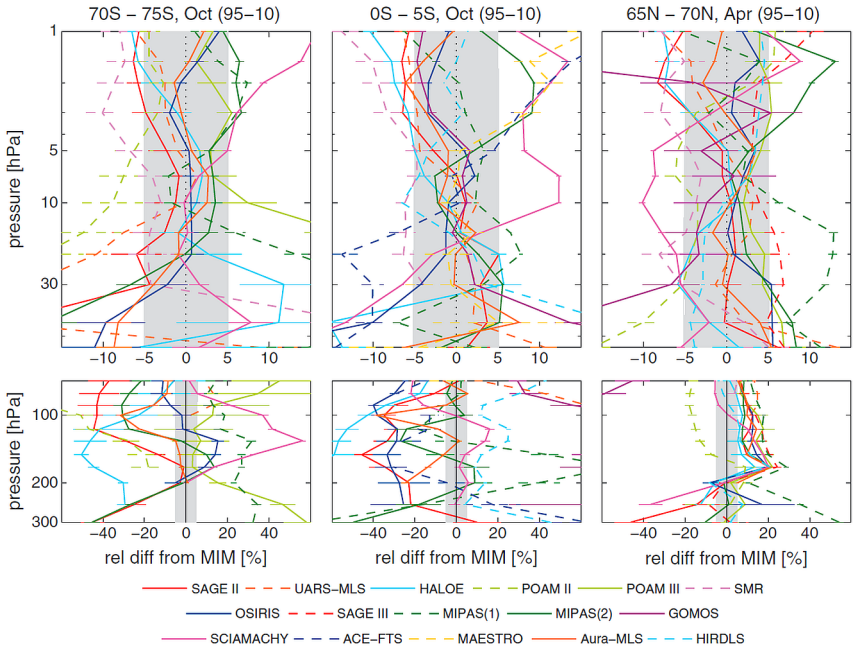

Figure 1.6 – The ozone monthly-averaged zonal mean measured by 16 limb sounders is given for 3 latitude regions. The data cover the years 1995-2010. The upper part shows the situation for the stratosphere, the lower part shows the UT/LS region. Each curve is associated to an instrument (the legend is in the bottom). The zero indicates a perfect agreement with the multi-instrument mean (MIM). The gray area always indicates the ±5% region of agreement. (Reproduced from [100])

1.2.3 Spatio-temporal sampling bias

Another important aspect has been recently emphasized by Toohey et al. [102]: the finite atmospheric spatio-temporal sampling capabilities of an instrument can potentially introduce large biases in its climatological products. This feature is directly related to the measurement technique. Limb emission instruments, for instance, have the widest geographical coverage and can measure at any time (day and night). The climatologies derived from their data deliver the highest accuracy, as monthly zonal means benefit from a dense sampling in space and time. On the other hand, solar occultation instruments acquire data only twice per orbit and are limited to small latitudinal bands. Their climatologies will suffer from the inhomogeneous sampling and the relatively small size of the dataset. In that case, biases will often be larger than 10% at high latitudes and in the UT/LS region, i.e. where the spatio-temporal

gradients in the ozone field are large.

1.2.4 Vertically-resolved data and atmospheric models

Despite the instrumental biases and the inherent limitations of each limb sounder measurement technique, profiles are essential to a better understanding of the atmosphere. More specifically, besides the monitoring of the trace species fields and the detection of trends, they provide the necessary inputs to the atmospheric modeling community. Climate models need the vertical distribution of absorbing species to simulate the radiative balance of the Sun-Earth system. Chemistry-climate models (CCM), which can potentially reproduce and forecast the state of the atmosphere, regularly assess their performance by comparing their output to actual measurements.

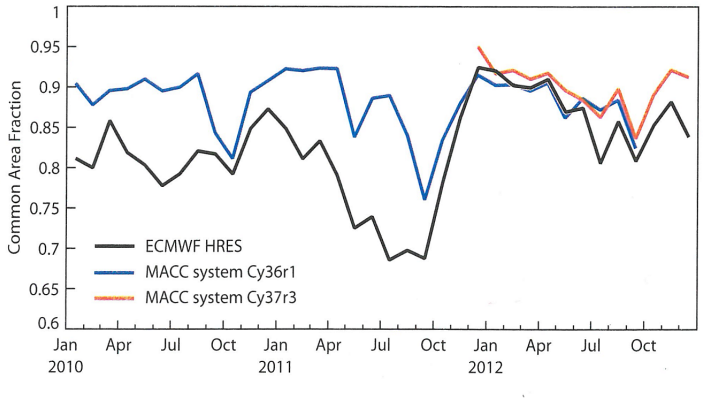

Figure 1.7 – Time series of the Common Area Fraction (CAF), an index representing the agreement between an actual measurement and its model counterpart (always between 0 and 1). Here, the CAF is for a polar ozone profile measured by an ozonesonde. From November 2011, ECMWF model assimilated limb data and filled the performance gap with the MACC systems which were already relying on such type of data for much longer. (Reproduced from [41])

A nice example arise from models which are constrained by assimilation of measurements. A recent paper [41] illustrates the jump in performance gained by the European Centre for Medium Range Weather Forecasts (ECMWF). Their High RESolution (HRES) forecasts assimilated the ozone total columns provided by SCIAMACHY during the entire ENVISAT lifetime. Comparison with MACC models (Monitoring Atmospheric Composition and Climate) revealed a gap in performance

when it came to model polar ozone profiles. The lower performance resulted from the absence of assimilation of vertically-resolved data, something which was done by MACC with the data from the Microwave Limb Sounder (MLS). After assimilating MIPAS data (Michelson Interferometer for Passive Atmospheric Sounding, another limb sounder onboard ENVISAT) from November 2011 onwards, the gap was filled until the loss of ENVISAT in Spring 2012 (Fig. 1.7).

Similarly, an intercomparison exercise between the four Chemistry Transport Models (CTM) taking part to the MACC project led to the same conclusion [69]: models assimilating total columns only are much less accurate than models using limb data, especially where large spatio-temporal gradients occur. At identical locations, the models lacking vertically-resolved measurements have biases sometimes larger than 50%, whereas the one assimilating MLS data was always below 10%.

From the various points discussed above, an ideal atmospheric instrument should have the following features. First, it should be a spaceborne limb sounder onto a polar orbit delivering the largest possible number of measurements to minimize the sampling error when building climatologies. Second, while the forthcoming gap in limb sounders should make any instrument welcome, careful attention should not only be paid on maximizing the accuracy and the precision of the retrieved products, but also on keeping the error budget constant over the instrument lifetime. The goal is to allow for easier trends detection, a capability which would benefit from a long-duration mission.

1.3 The radiative transfer problem for limb scattering instruments

The basic measurement method of the instrument being developed in this work is relying upon the spectral signature of atmospheric species (molecules or particles). As we are aiming at a passive remote sensing technique, we are left with a limited number of possibilites, namely the extinction of solar, stellar, lunar or planetary irradiance (mainly in the UV-VIS-NIR), or the emissions from the species themselves (thermal infrared up to mm wavelengths). Both methods exclude each other as their spectral ranges and inherent methodologies are completely different. While instruments looking at emissions offer the largest coverage in terms of geolocation and target species, the impossibility of a sub-kilometer vertical resolution and the lack of sensitivity to non-emitting particles (aerosols) make the first technique more attractive in view of

the challenges targeted in the previous section.

Instruments looking at the extinction of a celestial body irradiance can be grouped into two families: the occultation mode (direct transmittance) and the limb-scatter mode (solar light scattered in the direction of the instrument). The first one has clearly the simplest geometry with numerous examples of previous sounders (SAM II, SAGE I to III, HALOE, ORA, GOMOS, ACE-FTS, etc). Many papers have discussed the approach extensively, see for instance [28], [9] and [8]. On the other hand, the limb scattering technique is only well proven since the flying of SME in the 1980's, and OSIRIS and SCIAMACHY in the early 2000's.

The main difficulty with this observation mode is the determination of the light path. While in occultations, the instrument line of sight (LOS) points to the light source, here the radiance originates from a diffuse region around the LOS tangent point. With the neutral air density becoming exponentially larger as the tangent height decreases, the photon mean free-path gets shortened and multiple scattering becomes more likely. The retrieval of minor species is therefore fully dependent on the modelling accuracy of the basic physical processes: absorption, scattering and emission.

The key point for a limb scattering algorithm is to correctly solve the radiative transfer problem. The theoretical foundations have been raised by Chandrasekhar in the 1940's [23]. He stated the mathematical problem and solved it for a wide range of geometries and approximations. He also paid attention to the transfer of polarization. Since his work, several authors have gone deeper into the field, with the main motivation of proposing solving procedures for selected classes of geometries (plane parallel or spherical atmospheres), [85], [104], [98], [70], [72],...

The purpose of this work is not to make a significant contribution in the field of limb scattering modelling. However, as it will be mentionned on several occasions in the coming chapters, we have used a radiative transfer model for our simulations. Consequently, this section will describe the basic statements of the radiative transfer problem, and illustrate it with a few results.

1.3.1 The photon transport equation

The radiative transfer problem deals with the interaction between atmospheric particles and photons from the top of the atmosphere (TOA) to an observer, which in this case is the spaceborne instrument. Ideally, in limb scattering geometry, the only collected photons are those exiting the atmosphere along a tangent line. The point of shortest distance

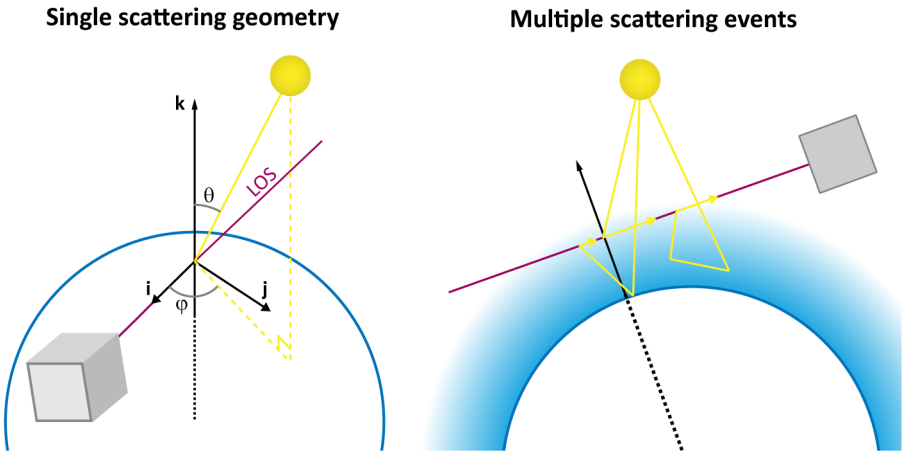

Figure 1.8 – Left: Single scattering event with the solar azimuth ϕ and zenith θ angles represented in a frame attached to the TP. Right: Side view illustrating three different processes participating to the total signal for a given LOS: single scattering, albedo and multiple scattering.

from the surface is called the tangent point (TP). There are different types of paths allowing a photon to be scattered into the LOS. The largest contribution comes from photons being scattered only once from an air mass close to the TP. With lower TP however, the probability of scattering increases due to the higher air density and photons are more likely to encounter multiple scattering, with a possible last kicking into the LOS. Albedo effects (reflection at a surface) from Earth or clouds are a significant source of upward flux too (Fig. 1.8).

The light source here is the Sun. The solar spectrum is well fitted by Planck's law of the black body radiation. The associated temperature is about 5800 K, a temperature found in a relatively thin layer of the solar sphere called the photosphere. The continuous spectrum is affected by sharp features called the Fraunhofer lines. They indicate absorption by atomic elements in the photosphere.

The radiation field can be described as a distribution function of photons, where

$$f(\boldsymbol{r}, \boldsymbol{n}, \nu, t)\, \mathrm{d}\Omega\, \mathrm{d}\nu$$

is the number of photons per unit volume centered at position $\boldsymbol{r}$ at time t, whose frequency belongs to the interval $[\nu, \nu + \mathrm{d}\nu]$ and propagating at speed c within a pencil of solid angle $\mathrm{d}\Omega$ in the direction of the unit vector $\boldsymbol{n}$. With each photon carrying an energy $h\nu$, the amount of energy in the volume $\mathrm{d}V = c\,\mathrm{d}t\,\boldsymbol{n}.\mathrm{d}\boldsymbol{S}$ (with $\mathrm{d}\boldsymbol{S}$ being an infinitesimal surface

vector) is given by

$$\mathrm{d}E = h\nu\, f\, c\, \mathrm{d}t\, \boldsymbol{n}.\mathrm{d}\boldsymbol{S}\, \mathrm{d}\Omega\, \mathrm{d}\nu. \tag{1.2}$$

The spectral radiance (or the specific intensity) I is defined as the amount of radiation energy of identical frequency crossing a surface $\mathrm{d}\boldsymbol{S}$ over time $\mathrm{d}t$, and propagating along $\boldsymbol{n}$ within the cone subtended by $\mathrm{d}\Omega$. From eq. (1.2), we have

$$I(\boldsymbol{r}, \boldsymbol{n}, \nu, t) = \frac{\mathrm{d}E}{\boldsymbol{n}.\mathrm{d}\boldsymbol{S}\, \mathrm{d}\Omega\, \mathrm{d}\nu\, \mathrm{d}t}, \tag{1.3}$$

$$= h\nu\, c\, f(\boldsymbol{r}, \boldsymbol{n}, \nu, t). \tag{1.4}$$

Typical units of I are W cm^{-2} str^{-1} μm^{-1}.

Considering the radiation field as a gas of photons, its distribution function f should satisfy the Boltzmann transport equation (BTE):

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}}{m} \cdot \boldsymbol{\nabla} f + \boldsymbol{F} \cdot \boldsymbol{\nabla}_{\boldsymbol{p}} f = S(\boldsymbol{r}, \boldsymbol{n}, \nu, t), \tag{1.5}$$

with $\boldsymbol{p}$ the momentum of the particles, $\boldsymbol{F}$ the force acting on them, $\boldsymbol{\nabla}$ and $\boldsymbol{\nabla}_{\boldsymbol{p}}$ the gradient operator for position and momentum respectively, and S, a source term usually accounting for the collisions which cause discontinuous changes in the distribution function (by adding or removing particles from their phase space). Now particularizing to the gas of photons, the BTE becomes

$$c\, \boldsymbol{n}.\boldsymbol{\nabla} f = S(\boldsymbol{r}, \boldsymbol{n}, \nu, t). \tag{1.6}$$

The force term in eq. (1.5) has been dropped in absence of general relativistic effects, while over sufficiently short intervals, time variations of the radiance field can be neglected as well. In the photon case however, the source term is not related to collisions. Instead, it describes the net change of photon density due to all the processes related to the interaction of light and atmospheric particles. Clearly, depletion *and* population of the radiation field should be considered. In the following sections, we describe the most important interactions competing within this source term.

There are essentially four classes of interactions which can take place in the radiometric budget: absorption, scattering, emission and reflection. In the frame of a UV-VIS-NIR instrument recording spectral radiance levels in a limb scattering geometry, we will not discuss the last two. The reason is that reflections at the Earth surface or at the cloud top are of utmost importance to nadir instruments but a second order

effect for limb sounders, while thermal emissions do not fall within our spectral range. Atomic emissions do increase the signal at very specific wavelengths, but they do not participate to the general shape of the limb-scatter radiance.

1.3.2 Rayleigh scattering phase function

What makes the limb-looking geometry feasible (and what gives the radiance its general shape above 350 nm) is the elastic scattering by air molecules, commonly referred to as Rayleigh scattering. This effect occurs when incident radiation couples with a dipole of negligible dimension compared to the optical wavelength. Part of the energy is absorbed and re-radiated by the dipole with an angular distribution function proportional to $\sin^2\theta$ (θ being the angle from the polarization direction). Rayleigh's theory predicts that for unpolarized light (which is the case for the solar light) the scattering is not homogeneous in all directions. Defining $\boldsymbol{n'}$ and $\boldsymbol{n}$ as the propagation vectors of the incident and scattered beams respectively, the scattering angle Θ is such that $\cos\Theta = \boldsymbol{n}.\boldsymbol{n'}$. The Rayleigh scattering phase function p for unpolarized light is given by

$$p(\boldsymbol{n}, \boldsymbol{n'}) = \frac{3}{4}\left(1 + \cos^2\Theta\right). \tag{1.7}$$

This function defines the probability for an incident beam propagating along $\boldsymbol{n'}$ to be scattered into $\boldsymbol{n}$. As it expresses a propability, p is normalized to unity:

$$\int_{4\pi} p(\boldsymbol{n}, \boldsymbol{n'})\mathrm{d}\boldsymbol{n} = 1. \tag{1.8}$$

The role of the dipole is played by the air most abundant diatomic molecules: N_2 and O_2. Knowing the Sun position, the probability of scattering in the instrument LOS can be found with eq. (1.7).

It is worth noting that the scattered field is not unpolarized anymore. Especially when $\Theta = 90°$, the polarization state should be completely linear. This is a consequence of the dipole re-radiation distribution. In practice, even at right angle, the polarization is never completely linear. This is due to a certain anisotropy of the scattering molecules. A better description should therefore account for the so-called depolarization ratio as described by Chandrasekhar [23]:

$$p(\cos\Theta) = \frac{3}{4(1+2\gamma)}\left((1+3\gamma) + (1-\gamma)\cos^2\Theta\right). \tag{1.9}$$

The parameter γ is a function of the depolarization factor ρ, a quantity which takes different values for each molecule and slightly varies with

wavelength [20]:

$$\gamma = \frac{\rho}{2-\rho}. \tag{1.10}$$

In air, the dispersion of ρ ranges from 0.045 at 200 nm down to 0.027 at 1 μm.

So far, the discussion on the Rayleigh phase function involved unpolarized incident light. As mentionned earlier, this is suitable up to the first scattering event, but it does not hold for successive orders. Oikarinen [88] investigated the error made in the total limb radiance computation when a scalar code (where only the radiance intensity is computed) is used instead of a vector code (which distinguishes between orthogonal polarizations). For an aerosol-free atmosphere, the worst case occurs when the Sun is exactly above the LOS (ZA = 0°, AA = 90°). In that case, Oikarinen estimated that, for instance, at a tangent altitude of 20 km and for $\lambda = 500$ nm, an error of about 3% will arise. On the other hand, the multiple scattering events decrease on average the degree of polarization of the radiance, i.e. the ratio of light intensity as measured in two orthogonal directions. Therefore, an instrument performing acquisitions in one linear polarization will be less sensitive to modelling errors (incorrect assumptions on the air density profile) at the lowest altitudes (where multiple scattering is more likely).

1.3.3 Rayleigh scattering cross-section

From a classical point of view, photons propagating through atmospheric layers have a finite probability of being scattered. Each particle exposes an effective cross-section σ_s to the incident radiation (usually given in cm^2). Considering an infinitesimal volume $\mathrm{d}V = A\,\mathrm{d}s$ filled with particles, the number of particles in the volume is $N = n\,\mathrm{d}V$, with n the number density in cm^{-3}. The probability for a photon to be scattered while propagating through the slab is given by the ratio of the total surface occupied by the particles in the volume to the slab surface: $P_s = \sigma_s N/A = \sigma_s n A \mathrm{d}s/A = \sigma_s n \mathrm{d}s$. For a radiation field of intensity I, the loss of intensity due to Rayleigh scattering in the slab is therefore

$$\mathrm{d}I = -P_s\, I = -\sigma_s n \mathrm{d}s\, I. \tag{1.11}$$

Solving the differential equation yields the Beer-Lambert law applied to Rayleigh scattering:

$$I(s) = I(0)\exp(-\sigma_s n s). \tag{1.12}$$

When $s = (\sigma_s n)^{-1}$, the intensity is reduced by a factor e, and this particular distance is called the mean free path of the photons. The parameter

$k_s = \sigma_s n$ is called the extinction coefficient for Rayleigh scattering. It depends on the wavelength.

The Rayleigh scattering cross-section derivation follows from the ratio of the power emitted by the excited dipole (electron–nucleus system) to the incident radiation power per unit area. In the approximation of the light frequency being much smaller than the dipole harmonic frequency, we have [20]

$$\sigma_s(\lambda) = \frac{24\pi^3}{\lambda^4 n^2} \frac{(n_r^2 - 1)^2}{(n_r^2 + 2)^2} \frac{6 + 3\rho}{6 - 7\rho}, \tag{1.13}$$

with n_r the refractive index of air, n the number density and ρ the depolarization factor. The first two fractions in eq. (1.13) account for the pure Rayleigh scattering for a mixture of air molecules, while the third one corrects for their anisotropy.

1.3.4 Transport equation in a purely Rayleigh scattering atmosphere

In a purely scattering atmosphere made of unabsorbing molecules, the photon transfer equation (eq. (1.6)) in which the distribution function f is replaced by the radiance I (using eq. (1.4) and the fact that the source term is the net balance of energy divided by the energy of one photon) takes the following form:

$$\boldsymbol{n}\boldsymbol{\nabla} I = h\nu \left(S^+ - S^- \right) \tag{1.14}$$

$$= k_s \int_{4\pi} I^*(\boldsymbol{n'})\, p(\boldsymbol{n}, \boldsymbol{n'})\, \mathrm{d}\boldsymbol{n'} - k_s\, I(\boldsymbol{n}). \tag{1.15}$$

By dividing the source term into a population (S^+) and an extinction (S^-) term, this equation expresses the fact that Rayleigh scattering is reponsible for some extinction of the radiance propagating along $\boldsymbol{n}$ (second term), while being also the cause of an increase of the intensity through the scattering of the incoming radiance I^* from all directions $\boldsymbol{n'}$ into direction $\boldsymbol{n}$ (first term). This is obviously a simple picture which requires further complications.

The first of them is probably to account for multiple scattering at the lower altitudes. It is a very demanding task from the computational point of view. The second scattering order for instance is obtained by estimation of the probability that a photon is first scattered towards the TP then again into the LOS. Triple scattering adds one more order of complexity. However, the fractional contribution of the higher orders to the detected signal is clearly minor. A figure from [11] illustrates the case

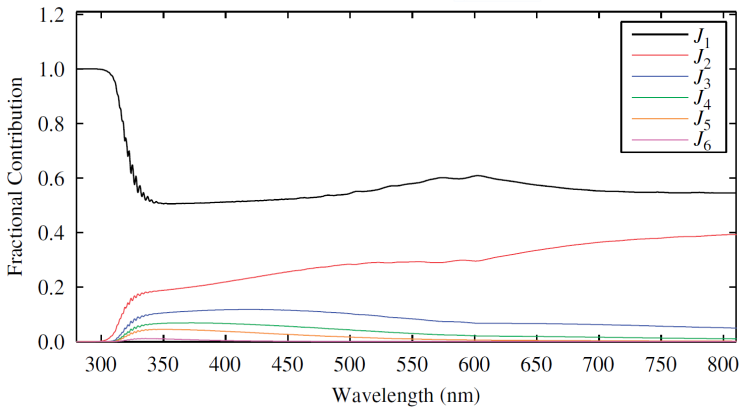

Figure 1.9 – Relative contribution of the scattering orders for a typical measurement geometry of OSIRIS and a tangent point at 25 km. (Reproduced from [11])

for the measurement geometry of the OSIRIS instrument and a TP at 25 km (Fig. 1.9). It shows how important the successive scattering orders are to the total radiance as a function of the wavelength. Below 300 nm, the absorption by ozone is so strong that only singly scattered photons contribute. At longer wavelengths, the first two orders are dominant, with the sum of all higher orders amounting to 30% at 350 nm, but to only 5% at 800 nm.

Up to this point, we have only discussed Rayleigh scattering, i.e. the scattering of solar light by atmospheric molecules acting as tiny dipoles. However, volcanoes, violent storms, or simply global atmospheric circulation are responsible for the loading of the stratosphere with much larger particles referred to as *aerosols*. The general understanding of the formation of aerosols is that H_2SO_4 and H_2O molecules are nucleating on Aitken particles (radius smaller than 100 nm). Once initiated, the growth can yield sizes up to 1 micron. The particle size distribution is an important parameter when aerosol extinction measurements must be converted to number densities. It changes with altitude and it strongly depends on the volcanic activity which appreciably changes the mean composition (for up to several years).

The scattering of light by spherical particles of non-negligible dimension compared to the wavelength is described by the Mie theory of scattering. The phase function is very different from Rayleigh scattering, as the forward and backward directions are dominant. Also, the dispersion of the scattering cross-section as a function of wavelength is different. The extinction and the scattering by aerosols is an important aspect of

radiative transfer modelling, especially in the lower stratosphere (below 40 km). Details on aerosol scattering theory can be found in [105] for instance.

1.3.5 Molecular absorption

Finally, a relatively small effect in terms of absolute magnitude of the limb-scatter signal, but a fundamental aspect for the remote sensing principle: the absorption by molecular species. In the UV-VIS range, photons carry sufficient energy to induce eletronic state transition or even photodissociation. This often translates into smooth absorption cross-sections or broad absorption peaks which are determined experimentally (e.g. O_3, NO_2, BrO, OClO). In the VIS-NIR domain, the transitions occur between vibrational and rotational states. In that case, extremely good spectral resolution (obtained by interference techniques) is needed to resolve these lines (H_2O, CH_4, CO_2,...). If not, the spectrum convolved with the instrument transfer function will show a broad structure mixing lines from saturated and non-saturated states. The retrieval can therefore become quite tricky as the radiance will vary in a nonlinear way with respect to the absorber concentration.

From the radiative transfer point of view, the molecular absorption is well represented by an extinction coefficient k_a, such that $\mathrm{d}I = -k_a I \mathrm{d}s = -\sigma_a(\nu) n I \mathrm{d}s$. However, at a given wavelength, all the photosensitive species (denoted below by the upperscript (i)) must be included, such that

$$\mathrm{d}I(\nu) = -\sum_i \sigma_a^{(i)}(\nu)\, n^{(i)}\, I(\nu)\, \mathrm{d}s. \tag{1.16}$$

1.3.6 Radiative transfer equation for a single scattering and absorbing atmosphere

The integral form of the radiative transfer equation for a single scattering and absorbing atmosphere can now be written. Fig. 1.10 illustrates the problem of computing the total radiance propagating along the LOS and caused by several single scattering events. The respective positions of the Sun, the instrument and its LOS define a range of scattering angles $\Theta \in [\Theta_{\min}, \Theta_{\max}]$. For convenience, we will use two variables for the distance travelled by the photons. For the slant path connecting the TOA to the scattering point, s measures the distance. Extreme values are $s = 0$ for the TOA and $s = s_s$ for the scattering point. If we consider the LOS as a reference axis, then we can use the variable l_s to refer to

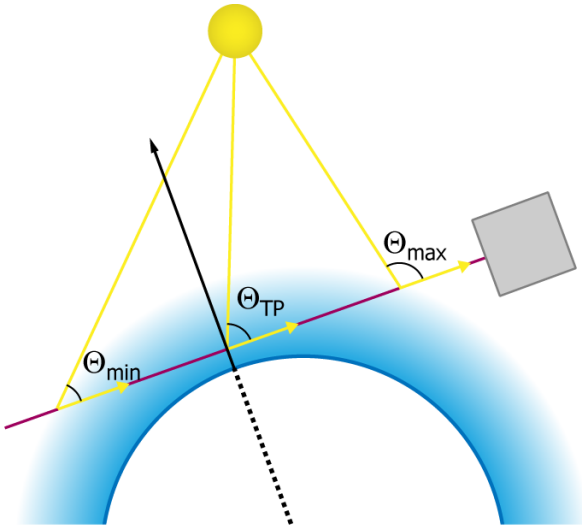

Figure 1.10 – Single scattering geometry.

the coordinate of the scattering point along this axis. Moreover, the geometry of the problem makes the scattering point uniquely associated to Θ, such that we can write an explicit dependence $l_s(\cos\Theta)$. One can impose $l_s(\cos\Theta_{\min}) = 0$.

In our approximation, the radiance will first experience extinction (absorption and scattering) from the TOA ($I_0(\nu)$) down to a given scattering point somewhere along the LOS. Integrating eq. (1.11) and (1.16) along that path, the incident radiation at $s_s = l_s(\cos\Theta)$ follows:

$$I\left(l_s(\cos\Theta),\nu\right) = \int_0^{s_s} I_0(\nu)\,\exp\left(-k_s(s,\nu)\,s - \textstyle\sum_i k_a^{(i)}(s,\nu)\,s\right)\mathrm{d}s. \quad (1.17)$$

There, a fraction of the incident beam will be scattered into the LOS. This forms the single scattering radiance (I_{ss}) at a given point:

$$I_{\mathrm{ss}}\left(l_s(\cos\Theta),\nu\right) = k_s\left(l_s(\cos\Theta),\nu\right)\,I\left(l_s(\cos\Theta),\nu\right)\,p(\cos\Theta), \quad (1.18)$$

where the arguments $\boldsymbol{n}$ and $\boldsymbol{n'}$ have been replaced by $\cos\Theta$ according to the scattering angle definition: $\boldsymbol{n}.\boldsymbol{n'} = \cos\Theta$. It is worth noting that as no other light source than the Sun is considered here, there is no need for a full sphere integration as in eq. (1.15).

After scattering, extinction is again experienced along the rest of the path until escaping the atmosphere. This last part must be repeated for each scattering event such that all the source terms along the LOS are

added (integration over $[\cos\Theta_{\min}, \cos\Theta_{\max}]$):

$$I_{\text{out}}(\nu) = \int_{\cos\Theta_{\min}}^{\cos\Theta_{\max}} \left[\int_{l_s(\cos\Theta)}^{l_s(\cos\Theta_{\max})} I_{\text{ss}}\big(l_s(\cos\Theta), \nu\big) \times \right.$$
$$\left. \exp\left(-k_s(l_s,\nu)\, l_s - \textstyle\sum_i k_a^{(i)}(l_s,\nu)\, l_s\right) \mathrm{d}l_s \right] \mathrm{d}\cos\Theta. \quad (1.19)$$

1.3.7 Typical limb-scatter radiance profile

To conclude this section on radiative transfer, we would like to illustrate some key aspects by a numerical example obtained with the radiative transfer code MODTRAN. This model and the parameters used for the simulations will be further discussed in the next chapters.

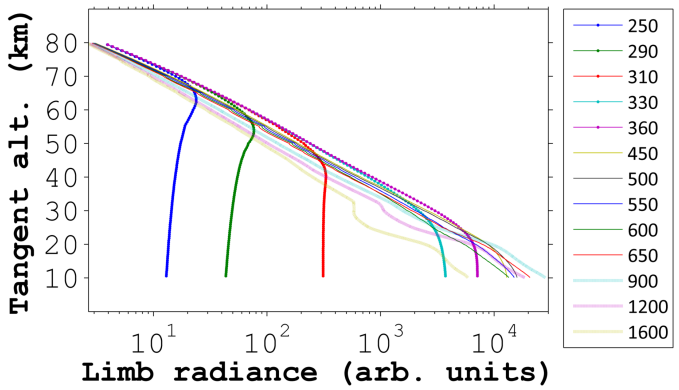

Figure 1.11 – Limb-scatter radiance computed by MODTRAN for selected wavelengths in function of the tangent altitude. The wavelengths are in nm.

This working case shows the limb radiance at selected wavelengths as a function of the tangent height (Fig. 1.11). Pressure, temperature and species concentration profiles follow a standard mid-latitude summer atmosphere. No clouds are present in the scene and a surface albedo of 0.5 is used. Background atmospheric aerosol loading is assumed.

As can be seen, all radiance curves follow a similar pattern. As explained above, the limb-scatter signal is shaped after a competition between absorption and scattering, but the weight of each process varies with the wavelength. The first curve we should have a look at is the purple one at 360 nm. At this wavelength, ozone absorption is negligible such that Rayleigh scattering is dominant. From the TOA down to about 30 km, the signal increases exponentially with lower altitudes. This simply follows from the increase of air density, and hence, of the

number of scatterers. Below 30 km, the extinction along the LOS gets so large that it starts cancelling any additional scattering. Lower, the radiance levels off under a balanced competition between extinction and scattering. This behaviour almost vanishes for longer wavelengths (certainly above 450 nm). Rayleigh scattering is still responsible for the exponential shape, but as its cross-section varies as λ^{-4} (see eq. (1.13)), the extinction is reduced and no *knee* appears in the profile.

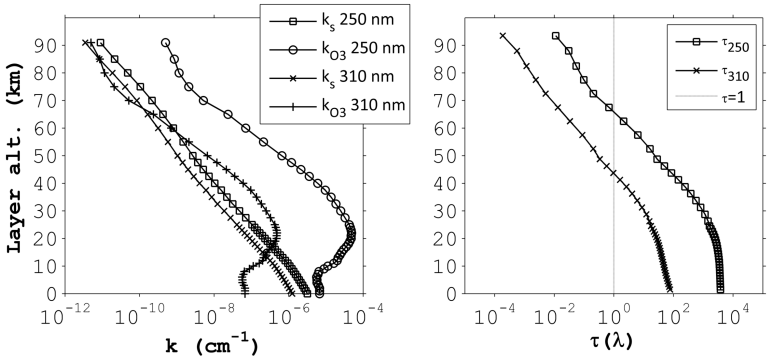

Figure 1.12 – Left: Extinction profile at 250 and 310 nm for Rayleigh scattering (k_s) and O_3 absorption (k_{O_3}) in a standard atmosphere. Right: Optical thickness for a limb geometry (single scattering approximation) based on extinction by Rayleigh scattering and ozone absorption only.

If we now look at the shortest wavelengths, we see that the inflection point is lifted up. Clearly, the strong absorption by ozone at these wavelengths is playing an important role. To see which process is dominant, one has to go back to the standard pressure, temperature and ozone profiles. We will focus on the curves $\lambda_1 = 250$ nm and $\lambda_3 = 310$ nm. Taking values from [22] and [20], we have: $\sigma_{O_3}(\lambda_1) = 1.1 \times 10^{-17}$, $\sigma_{O_3}(\lambda_3) = 1.0 \times 10^{-19}$, $\sigma_s(\lambda_1) = 1.3 \times 10^{-25}$ and $\sigma_s(\lambda_3) = 4.9 \times 10^{-26}$. All cross-sections are in cm^2 molecule^{-1}. From there, the respective extinction profiles can be obtained with $k(\lambda, z) = \sigma(\lambda)\, n(z)$, where z is the air mass altitude (left panel in Fig. 1.12). Clearly, at 250 nm, ozone absorption is the driver. At 310 nm however, this is already not the case anymore.

A valuable indication of the opacity of the medium, is obtained with the optical thickness:

$$\tau(\lambda) = \int \Big[\sigma_{O_3}(\lambda)\, n_{O_3}(s) + \sigma_s(\lambda)\, n_s(s)\Big]\, \mathrm{d}s. \tag{1.20}$$

The integration goes over the total light path, i.e. from the TOA to the scattering point and then along the LOS. Basic geometrical considera-

tions show that in single scattering approximation, a large part of the photon path is covered in the scattering layer (because of the tangent viewing geometry). A rough estimation yields a ratio of 40:1 between the tangent path length and the incident path length. Taking this factor into account for the integration limits in eq. (1.20), we end up with the optical thickness experienced by a limb-scattered photon of a given wavelength (right panel of Fig. 1.12).

It is often accepted that $\tau = 1$ indicates the boundary between an optically thin ($\tau < 1$) and an optically thick medium ($\tau > 1$). Comparing the right panel of Fig. 1.12 with Fig. 1.11, we see that the altitude of the *knee* of the radiance profiles at 250 and 310 nm (blue and red thick curves respectively) corresponds to the intersection of the curves τ_{250} and τ_{310} with the vertical line at $\tau = 1$.

1.4 From radiance to concentration profiles: the inverse problem

Independently of the limb-scattered radiance measurement technique (scanning grating spectrometer or spectral imager), the retrieval of atmospheric species concentration profiles is based on the fitting of a numerical model to the measured signals. Often, the data are first treated in order to discriminate interfering species (DOAS method, [90]) or to simply increase the signal-to-noise ratio (SNR) by associating distinct measurements, although direct radiance fitting can also be used. In all cases, we are facing an inverse problem for which many solving strategies may apply.

In atmospheric remote sensing however, one method is widely adopted for its simple formalism and the diagnostic tools it comes with. The comprehensive theory has been described by Rodgers [94] and the following paragraphs mainly derive from his work. This method is often referred to as *optimal estimation*, which actually encompasses a range of techniques based on the maximizing or minimizing of some parameter (probability, information content, ...). Here, we will focus on one of these techniques which provides the *maximum a posteriori solution* (MAP).

Let $\boldsymbol{x}$ be the unknown of the inverse problem: the concentration profile of a target atmospheric species, often referred to as the *state vector*. It is a $m \times 1$ vector. Each element of this vector corresponds to the concentration of the species within an air mass layer of the atmosphere. The concentration is expressed in molecules cm^{-3} and is assumed constant

within a layer.

Let $\boldsymbol{y}$ be the *measurement vector*. It contains n elements corresponding to the radiance emitted from a number of tangent altitudes and measured at one wavelength. Measurements at more than one wavelength are appended one after the other. If ratios of radiance or other types of combination are used instead of the direct measurements, they can also enter the vector. We will always remain in the over-constrained case by forcing $n > m$. This can be ensured by either performing more acquisitions, either decreasing the number of layers in the atmospheric model.

With a perfect radiative transfer model (simply denoted hereafter as the *forward model*, by opposition to the inversion algorithm), that is, a code which accurately models all the physical processes and receives as input a correct representation of the atmospheric state (pressure, temperature, spatial distribution of other absorbers and scatterers, Sun position, viewing directions, albedo, ...), then $\boldsymbol{x}$ and $\boldsymbol{y}$ are related by a function F:

$$\boldsymbol{y} = F(\boldsymbol{x}, \boldsymbol{b}) + \boldsymbol{\epsilon}, \tag{1.21}$$

where the vector $\boldsymbol{b}$ contains all the other parameters characterizing the atmospheric state which are not targeted by the inverse problem, and $\boldsymbol{\epsilon}$ is the measurement error vector. Retrieving the concentration profile of the particle which is partly the cause of the detected signal is precisely finding the $\boldsymbol{x}$ corresponding to $\boldsymbol{y}$ through the function F. Due to the complexity of the problem, F^{-1} is not available, such that a simple step consists of approximating the forward model by a linear expression about a linearization point $\boldsymbol{x}_0$:

$$F(\boldsymbol{x}, \boldsymbol{b}) = F(\boldsymbol{x_0}, \boldsymbol{b}) + \boldsymbol{K}(\boldsymbol{x} - \boldsymbol{x_0}) + \boldsymbol{\epsilon}, \tag{1.22}$$

with $\boldsymbol{K}$ the Jacobian of the forward model, a $n \times m$ matrix whose elements are given by $K_{ij} = \partial y_i / \partial x_j$ $(i = 1, \ldots, n,\ j = 1, \ldots, m)$. In some measurement geometries, it is possible to derive analytical Jacobians, but in limb-scattering mode, the only choice is to obtain $\boldsymbol{K}$ from successive perturbations of the elements of $\boldsymbol{x}_0$ in the forward model and estimate the resulting perturbation on the measurement vector:

$$K_{ij} = \frac{F_i(x_{0_j} + \delta x) - F_i(x_{0_j})}{\delta x}. \tag{1.23}$$

In the method described by Rodgers, the statement of the inverse problem follows a Bayesian approach with the probability density function (pdf) of $\boldsymbol{x}$ and $\boldsymbol{y}$ given by normal distributions. This assumption,

certainly valid for well-behaved measurements (signals made of a large number of counts such that the central limit theorem applies), allows for a convenient description of the Bayes theorem:

$$P(\boldsymbol{x}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{x})\,P(\boldsymbol{x})}{P(\boldsymbol{y})}, \tag{1.24}$$

where $P(\boldsymbol{x})$ is the pdf of the state vector before taking into account any further knowledge provided by the measurements (this is the *a priori* pdf); $P(\boldsymbol{x}|\boldsymbol{y})$ is the pdf of the state vector including the knowledge provided by the measurements (this is the *a posteriori* pdf); $P(\boldsymbol{y}|\boldsymbol{x})$ is the pdf of the measurements when $\boldsymbol{x}$ is effectively a good description of the atmospheric state; and $P(\boldsymbol{y})$ is the pdf of the measurements before being performed (which is often so broad that it is uninformative and can be seen as a constant normalization factor).

In vector notation, the normal pdf of a vector $\boldsymbol{z}$ of mean $\bar{\boldsymbol{z}}$ and variance $\boldsymbol{S_z}$ takes the form[2]

$$\mathcal{N}(\bar{\boldsymbol{z}}, \boldsymbol{S_z}) = \frac{1}{\sqrt[n]{2\pi}\,|\boldsymbol{S_z}|^{1/2}}\,\exp\left(-\frac{1}{2}(\boldsymbol{z}-\bar{\boldsymbol{z}})^T\,\boldsymbol{S_z}^{-1}\,(\boldsymbol{z}-\bar{\boldsymbol{z}})\right). \tag{1.25}$$

Propagating the measurement errors into the covariance matrix $\boldsymbol{S_\epsilon}$, and assuming that $P(\boldsymbol{x})$ is correctly represented by a normal pdf $\mathcal{N}(\boldsymbol{x_a}, \boldsymbol{S_a})$, then the natural logarithm of eq. (1.24) gives

$$\ln P(\boldsymbol{x}|\boldsymbol{y}) = -\frac{1}{2}(\boldsymbol{y}-\bar{\boldsymbol{y}})^T\,\boldsymbol{S_\epsilon}^{-1}\,(\boldsymbol{y}-\bar{\boldsymbol{y}}) - \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x_a})^T\,\boldsymbol{S_a}^{-1}\,(\boldsymbol{x}-\boldsymbol{x_a}) + c. \tag{1.26}$$

By definition, the expectation value $\bar{\boldsymbol{y}}$ should correspond to the error-free measurement vector, which can be replaced by the linearized output of the forward model: $\bar{\boldsymbol{y}} = \boldsymbol{K}\boldsymbol{x}$. Therefore, we can re-write eq. (1.26) as

$$\ln P(\boldsymbol{x}|\boldsymbol{y}) = -\frac{1}{2}(\boldsymbol{y}-\boldsymbol{K}\boldsymbol{x})^T\,\boldsymbol{S_\epsilon}^{-1}\,(\boldsymbol{y}-\boldsymbol{K}\boldsymbol{x}) - \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x_a})^T\,\boldsymbol{S_a}^{-1}\,(\boldsymbol{x}-\boldsymbol{x_a}) + c. \tag{1.27}$$

From the *normality* of $P(\boldsymbol{y}|\boldsymbol{x})$ and $P(\boldsymbol{x})$, it follows that the posterior pdf $P(\boldsymbol{x}|\boldsymbol{y})$ is also a normal distribution, such that

$$\ln P(\boldsymbol{x}|\boldsymbol{y}) = -\frac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^T\,\boldsymbol{S_x}^{-1}\,(\boldsymbol{x}-\bar{\boldsymbol{x}}) + c'. \tag{1.28}$$

[2]It is common to denote a normal distribution by $\mathcal{N}(\mu, \sigma^2)$, i.e. a function which does not declare its variable, but only its parameters (mean μ and variance σ^2). We will follow this convention throughout the manuscript, though some readers would prefer a notation like $\mathcal{N}(x; \mu, \sigma^2)$ or $\mathcal{N}(x|\mu, \sigma^2)$ which explicitly identifies the variable.

Consistently, the right hand side of eq. (1.27) and (1.28) should be equal. By equating like terms, we obtain the expectation value of $P(\boldsymbol{x}|\boldsymbol{y})$, which indeed corresponds to the maximum of the atmospheric state pdf knowing the information provided by the measurement (MAP solution):

$$\bar{\boldsymbol{x}} = \boldsymbol{x_a} + \left(\boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \boldsymbol{K} + {\boldsymbol{S_a}}^{-1}\right)^{-1} \boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \left(\boldsymbol{y} - \boldsymbol{K}\boldsymbol{x_a}\right). \tag{1.29}$$

The uncertainty on the retrieved concentration profile $\bar{\boldsymbol{x}}$ is described by the covariance matrix of the pdf, which is obtained in a similar manner:

$${\boldsymbol{S}_{\bar{\boldsymbol{x}}}}^{-1} = \boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \boldsymbol{K} + {\boldsymbol{S_a}}^{-1}. \tag{1.30}$$

If the forward model is too much nonlinear and/or $\boldsymbol{x_0}$ is too distant from the true state vector, several iterations involving re-computation of $\boldsymbol{K}$ will be necessary.

The similarity of expression (1.29) with a least-square fit weighted by the measurement error is striking[3]. The advantage here is that some weight is given to the solution by the *a priori* knowledge of the state vector, providing some *stability* to the algorithm when the information added by the measurements is not significant for some part of the state vector.

The analysis of the obtained results is delicate. When working with simulated quantities ($\boldsymbol{y}$, $\boldsymbol{\epsilon}$), it is easy to see which part of the solution vector is driven by the *a priori*, because the true state is known. With real remote sensing measurements, it is not the case, and a convenient way of assessing the added value of the measurements is by looking at the sensitivity of the solution to the atmospheric state. This quantity is usually referred to as the *averaging kernel matrix*:

$$\boldsymbol{A} = \frac{\partial \bar{\boldsymbol{x}}}{\partial \boldsymbol{x}} = \left(\boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \boldsymbol{K} + {\boldsymbol{S_a}}^{-1}\right)^{-1} \boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \boldsymbol{K}. \tag{1.31}$$

The rows of $\boldsymbol{A}$ are normalized to unity and should ideally have the shape of a peaked function associating one element of the retrieved vector to one or more elements of the state vector. If the function is zero, then the measurements do not perform better than noise and they do not allow to distinguish the *a posteriori* pdf from the *a priori* pdf. In other words, and for the particular case of limb-scattering geometry, peaked averaging kernels indicate that the solution is strongly influenced by the measurements over a limited range of tangent altitudes. The width of the averaging kernels therefore emphasizes the actual spatial

[3]Solution to the weighted least-square problem: $\boldsymbol{x}_{\text{ls}} = \left(\boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \boldsymbol{K}\right)^{-1} \boldsymbol{K}^T {\boldsymbol{S_\epsilon}}^{-1} \boldsymbol{y}$.

resolution of the retrieved profiles. However, this width is limited by the forward model which divides the atmosphere into a given number of homogeneous layers. Practical examples will be given and discussed in Chapter 3.

Finally, when there are too many measurements for the number of unknowns ($n \gg m$), which can happen for instance when the spatial sampling is high compared to the model grid, numerical instabilities can arise from the largely over-constrained problem. In that case, adjacent rows of $\boldsymbol{K}$ can be quite redundant and possibly decrease the rank of the matrix. To overcome possible numerical problems, a singular value decomposition (SVD) is performed on $\boldsymbol{K}$ such that

$$\boldsymbol{K} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T. \tag{1.32}$$

$\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, $\boldsymbol{\Lambda}$ is diagonal. If all singular values of $\boldsymbol{\Lambda}$ are non-zero, one can reduce the over-constrained problem to a well-determined system of equations by subsituting $\boldsymbol{y}^* = \boldsymbol{U}^T\boldsymbol{y}$ and $\boldsymbol{x}^* = \boldsymbol{V}^T\boldsymbol{x}$ in the MAP equation, with $\boldsymbol{\Lambda}$ taking the place of $\boldsymbol{K}$. The covariance matrices ($\boldsymbol{S}_\epsilon$ and $\boldsymbol{S}_a$) are adapted accordingly.

Chapter 2

Imaging and atmospheric remote sensing — The refracted Sun case

2.1 Quantitative atmospheric information based on imaging techniques

If someone was asked to think of possible applications of a spaceborne imaging instrument, very likely the answers would be *geography*, *land cover*, *vegetation*, *intelligence*, or *astronomy*. Clearly, the term *imaging* suggests that the information seeked lies in the contrasts, the shapes, or the patterns, something which is readily captured by our eyes, although requiring some algorithmic development for machines. On the other side, as recalled in the previous chapter, atmospheric remote sensing overwhelmingly makes use of spectral information to identify the presence of absorbing or emitting species in the line of sight. The imaging aspect is often a by-product of scanning systems which call themselves hyperspectral imagers. Therefore it seems there is a gap between both approaches: one cannot derive atmospheric properties from a method based on image processing techniques, whereas scanning spectrometers claim they build images but their quality is poor.

While this statement is certainly true for trace gases which reveal themselves under spectroscopic techniques, we will show in this chapter that atmospheric pressure (or temperature) profiles can be unveiled using images of a setting or rising Sun. We believe that this application is one of the few examples of quantitative atmospheric information retrieved from an image processing-based technique. A recent similar effort ([49]) has been conducted by the team of the SOFIE instrument (Solar Occultation For Ice Experiment, [50]) onboard the AIM platform (Aeronomy of Ice in the Mesosphere). Other examples include automatic PSC flagging ([40]), or PMC shape and occurence trends analysis

([101]), but these are more qualitative approaches.

It is true that the number of atmospheric processes leaving their signature in the form of imaging features is quite limited. Refractive effects form the majority of this family (flattening, scintillation, scene shift). We are not aware of any other imaging-based quantitative atmospheric remote sensing technique, but maybe, in the future, a concept translating spectral signatures into 2-D patterns will exist and turn out to be useful...

The meaning of this chapter at this point of the dissertation is to investigate a possible atmospheric remote sensing method *based on imaging aspects only*. The next chapter will introduce the spectral dimension and from there on, *imaging and spectroscopy* will be coupled into the hyperspectral imager concept. As imaging is clearly a sub-domain of hyperspectral imaging, the method which will be presented below does not consitute a departure from the general topic of the present work. It is simply an application of the same instrument which does not use its full potential (the spectral dimension).

This work has been published in Atmospheric Measurement Techniques under the title *Zernike polynomials applied to apparent solar disk flattening for pressure profile retrievals*, [37].

2.2 Scope of the theoretical work: pressure profile retrieval from the apparent flattening of the solar disk

The measurement of refraction angles was one of the first remote sensing techniques in the field of spaceborne atmospheric science experiments. From the first radio occultation experiment during the Mariner IV mission that successfully retrieved pressure and temperature information in Mars atmosphere to the pictures of the Sun as seen through the Earth's atmosphere by SOFIE onboard AIM, this technique has proved to be a valuable alternative solution with respect to other remote passive methods like the detection of CO_2 or O_2 thermal emissions.

Gordley et al. ([49]) have shown that one can achieve temperature retrievals with an accuracy of ± 5 K up to 60 km provided that high spatial sampling of the solar disk is available. Their method is based on

accurate solar upper and lower edge detection during sunset or sunrise in order to measure refraction angles and to retrieve temperature profiles. Here, we follow a different strategy. By taking advantage of the asymmetrical flattening of the apparent solar disk caused by the exponential gradient in the air density profile, we retrieve atmospheric pressure from a quantified description of the whole solar disk deformation. For this purpose, we used Zernike polynomials to guarantee a unique orthogonal decomposition of the images. By simulating a wide variety of test cases, we built a look up table to map the link between the Zernike moments of the images and the pressure state of the atmosphere.

This theoretical study is intended to show the capabilities of a simple technique based on Sun image analysis that could be applied to any solar imager performing occultations. It does not require highly efficient pointing stabilization, and it has the advantage of using all the pixels constituting the solar disk image, allowing for measurements with very high signal-to-noise ratios.

In the following sections, we first describe the steps leading to the simulation of solar images including the solar limb darkening and the apparent solar disk flattening. Then we present the Zernike polynomials and we apply them to simulated sunset pictures for a large and non-redundant set of atmospheric pressure states, generated by a principal component analysis performed on a representative pressure dataset. Finally, we assess the limitations of the method by testing it on two different datasets.

2.3 Sun image simulation

2.3.1 Virtual instrument definition

The goal of this study is to exploit the relation between Sun images and atmospheric pressure profiles. To do so, we developed a virtual but realistic imaging instrument whose essential characteristics match those of ACE, the Atmospheric Chemistry Experiment on-board SCISAT-1 [8] which possesses two imagers that are used for a twofold purpose [48]. They serve primarily as a pointing support to the payload by keeping the full Sun in their field of view. Second, as they observe solar occultations through the Earth's atmosphere, they provide information about the spectral transmittance related to the atmospheric extinction profiles in two different spectral regions. The filter's central wavelengths are identical to two of the seven channels of the SAGE II instrument [28]: a visible channel at 525 nm weakly sensitive to aerosols and ozone, and

a near infrared (NIR) channel at 1020 nm sensitive to aerosols as well.

The second channel was selected as the reference for our simulations in order to minimize the effect of trace gas absorption (mainly NO_2 and O_3) on the transmitted irradiance in case of actual application of the method to real data. For the sake of realism, we also kept ACE's pixel size and field of view: the image is sampled by an array of 128×128 squared pixels dividing a total field of view of 30 mrad, which is about three times larger than the full Sun apparent angular size. Finally, we also preserved ACE's orbital parameters, and placed our virtual sensor on a 650 km circular orbit. By choosing the wavelength, the field of view and the orbit altitude, all the parameters required to simulate Sun images are defined.

2.3.2 Solar limb darkening

The Sun image simulation requires a description of the solar limb darkening, i.e. the apparent radial decrease of intensity across the solar disk. This darkening is also wavelength dependent and has been empirically parameterized by [87]. Defining $\mu = \cos\beta$ where β is the angle of emission of a Sun ray with respect to the local vertical at the Sun surface, the ratio between the light intensity emitted at this angle $I(\mu)$ and the intensity of rays emitted with $\beta = 0$ (i.e. from the center of the apparent disk) can be fitted by a 5th-order polynomial

$$\frac{I(\mu,\lambda)}{I(\mu=1,\lambda)} = \sum_{i=0}^{5} A_i(\lambda)\mu^i, \tag{2.1}$$

where the A_i coefficients are functions of the wavelength λ. Keeping the notation of Neckel, the six coefficients are defined by the following relations, where the values of the a_{ij} are given in Table 2.1:

$$\begin{aligned}
A_0 &= a_{00} + a_{01}\lambda^{-1} \\
A_1 &= a_{10} + a_{11}\lambda^{-1} + a_{15}\lambda^{-5} \\
A_2 &= a_{20} + a_{25}\lambda^{-5} \\
A_3 &= a_{30} + a_{35}\lambda^{-5} \\
A_4 &= a_{40} + a_{45}\lambda^{-5} \\
A_5 &= a_{50} + a_{55}\lambda^{-5}
\end{aligned}$$

`Table 2.1` – Values attributed to the limb darkening coefficients as given in [87] for wavelengths belonging to the spectral range 422 nm–1100 nm.

Neckel coefficients					
a_{00}	=	0.75267	a_{30}	=	2.42234
a_{01}	=	−0.265577	a_{35}	=	−0.017117
a_{10}	=	0.93874	a_{40}	=	−1.71150
a_{11}	=	0.265577	a_{45}	=	0.011977
a_{15}	=	−0.004095	a_{50}	=	0.49062
a_{20}	=	−1.89287	a_{55}	=	−0.003347
a_{25}	=	0.012582			

2.3.3 Atmospheric refractive index

Atmospheric refraction is the key physical process involved in the apparent flattening of the solar disk when observed through Earth atmosphere. The solar rays that penetrate the Earth's atmosphere experience an air concentration gradient giving rise to refraction. This process depends on the refractive index n of each particular layer. In this study, we assumed hydrostatic equilibrium, making n dependent on pressure (P) and temperature (T). CO_2 and water vapor were also considered according to the Ciddor formula [29] which is recommended by the International Association of Geodesy (IAG) for refractive index calculations in the visible and near infrared range in air.

In our model, the P, T and H_2O profiles are based on the US Standard 1976 atmosphere. Only CO_2 has been changed to a more up to date value of 400 ppmv. The division of the atmosphere into layers also follows the US Standard atmosphere: the layers are 1 km thick below 25 km, 2.5 km thick between 25 km and 50 km, and 5 km thick from 50 km to 100 km. A spline interpolation was performed to avoid discontinuities over multiple layers.

2.3.4 1-D ray tracing problem

The problem of simulating solar pictures as acquired after propagation of the light in the atmosphere mostly relies on the knowledge of three parameters (highlighted in the following sentence): there is only one path for light rays leaving the solar photosphere with a given **emission angle** β to graze the Earth's surface at a minimal **tangent height** h and hit the detector with a given **viewing angle** θ. The geometry is illustrated in Fig. 2.1 (left panel) for rays lying in the plane containing the centers of the Sun, the Earth and the spacecraft. The formalism

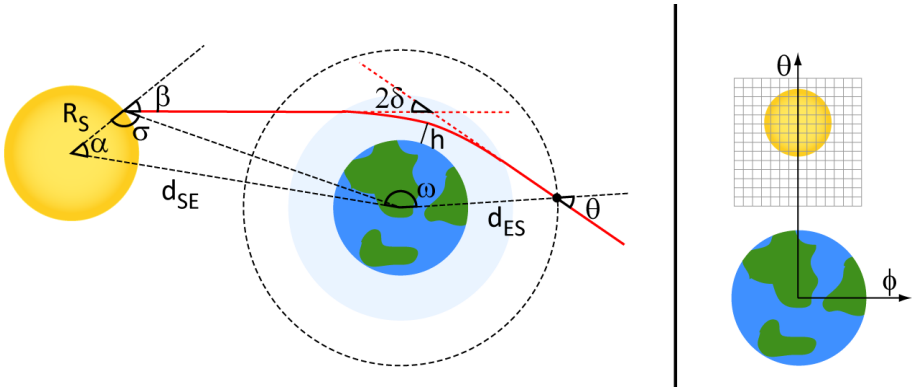

Figure 2.1 – Left: Illustration of the 1-dimensional ray tracing problem for photon trajectories in the Sun-Earth-Spacecraft plane. Right: Illustration of the reference frame for vertical and horizontal viewing directions as seen from the instrument. The origin of the coordinate system is at the intersection of the axes and the grid illustrates the pixel division of a possible instrument field of view. Relative sizes of the Earth, the Sun and the pixels are not representative of the orbit and instrument parameters used in this study.

used in the present section and the next one is an adapted version of a previous study from [109].

The trajectory of a light ray in a refractive medium is governed by Fermat's principle, stating that light always follows the fastest path. The associated distance L traveled between time t_A and time t_B is

$$L = \int_{s(t_A)}^{s(t_B)} n(s)\mathrm{d}s,$$

where $n(s)$ is the refractive index at the coordinate s along the path. Furthermore, assuming spherical homogeneous atmospheric layers, the following conservation formula holds for any value of s:

$$n(r(s)) \cdot r(s) \cdot \sin\psi(s) = c^{\mathrm{st}} \tag{2.2}$$

where c^{st} is a constant value, r denotes the geocentric distance of a point located at a coordinate s along the trajectory, and ψ is the angle between the local vertical and the ray trajectory. Applying eq. (2.2) to the point of the light trajectory situated at the smallest distance from the ground, at the altitude h, one can find the relation between the tangent height h and the viewing angle θ (angle between the light ray hitting the detector and the geocentric position vector of the spacecraft):

$$n(h(\theta)) \cdot (R_{\mathrm{E}} + h(\theta)) = d_{\mathrm{ES}} \cdot \sin\theta \tag{2.3}$$

where R_{E} is the Earth radius and d_{ES} the geocentric spacecraft distance.

In order to find the relation between the viewing angle θ, the tangent height h and the emission angle β, one can apply some trigonometric relations to the triangles of Fig. 2.1:

$$\beta(\theta) = \alpha - \xi(\theta), \tag{2.4}$$

where

$$\xi(\theta) = \pi + 2\delta(h(\theta)) - \omega - \theta, \tag{2.5}$$

2δ being the total refraction angle, ω, the Sun-Earth-Satellite angle, and α, the angle between the Sun geocentric position vector and the heliocentric position vector of the emission spot. The refraction angle 2δ depends on all the atmospheric layers crossed by the light ray, from the top of the atmosphere (TOA) to the lowest altitude, i.e. the tangent height. The tangent height-dependent refraction angle can be obtained by using the following expression (see for instance [33] for more details on the computation of refraction angles in the atmosphere):

$$2\delta(h) = -\int_{R_{\mathrm{E}}+h}^{\infty} \frac{2n(h)\cdot(R_{\mathrm{E}}+h)}{n(r)\cdot\sqrt{(n(r)\cdot r)^2 - (n(h)\cdot(R_{\mathrm{E}}+h))^2}} \left(\frac{\mathrm{d}n}{\mathrm{d}r}\right) \mathrm{d}r \tag{2.6}$$

The last step to solve eq. (2.4) with respect to θ is to remove the dependence on α by applying trigonometric relations, leading to the emission angle associated to a light ray having reached the tangent height h and being seen by the instrument at the angle θ (more details can be found in [109]):

$$\beta(\theta) = \left|\sin^{-1}\left(\frac{1}{\rho_1}(\rho_2 \sin\theta - \sin\xi(\theta))\right)\right| \tag{2.7}$$

where $\rho_1 = R_{\mathrm{S}}/d_{\mathrm{SE}}$ with R_{S} being the solar radius and d_{SE} being the Sun-Earth distance, and $\rho_2 = d_{\mathrm{ES}}/d_{\mathrm{SE}}$.

2.3.5 2-D ray tracing problem

Up to now, only emission points lying in the Sun-Earth-spacecraft plane (SES plane) have been considered, as illustrated by the left panel of Fig. 2.1. It was possible to address them with a single viewing angle (θ). For points located outside the SES plane, one can split their associated viewing angle into the vertical viewing angle θ and a horizontal viewing angle ϕ, allowing to univocally associate the image pixels with the couple of coordinates (ϕ, θ). The origin of the frame is chosen such that $(0, 0)$

points towards the Earth center and the θ-axis belongs to the SES plane. The solar disk center receives coordinates $(0, \theta_c)$. This is illustrated in the right panel of Fig. 2.1.

When no refraction occurs along the light path, the solar limb darkening is a radial effect only, i.e. the emission angle β of the detected photons is the same for points located at an equal distance of the solar disk center ($\phi^2+(\theta-\theta_c)^2 = c^{\text{ste}}$), allowing for a 1-dimensional treatment. But when observed through an inhomogeneous refractive medium, the radial symmetry does not hold anymore and the relation between the light ray emission angle and the viewing angle has to be computed for any combination of θ and ϕ.

Eq. (2.7) relates β to θ in the SES plane, whose intersection with the Sun forms a disk. It allows one to compute the limb darkening along the vertical direction for this particular meridian. We will first extend the validity of eq. (2.7) to the other adjacent slices of the Sun, then correct it for the limb darkening along the horizontal direction (which does not suffer from any atmospheric effect as no horizontal atmospheric gradient is considered).

This is more easily done by dividing the Sun into vertical slices of infinitesimal width. Every point belonging to the surface of a given slice is seen with the same ϕ angle and each slice is actually a disk whose radius $R_S(\phi)$ is approximately computed by the following relation:

$$R_S(\phi) \approx \sqrt{R_S^2 - (d_{SE} \cdot \phi)^2}. \tag{2.8}$$

Obviously, $R_S(\phi)$ decreases with ϕ and vanishes when ϕ is equal to half of the angular size of the Sun.

The limb darkening caused by a non-zero β angle within a ϕ-slice is obtained by using $R_S(\phi)$ instead of R_S in the calculation of ρ_1 in Eq. (2.7). One obtains an updated expression for β taking into account the reduction of the vertical angular size of the ϕ-slices:

$$\beta(\phi, \theta) = \left| \sin^{-1} \left(\frac{1}{\rho_1(\phi)} (\rho_2 \sin\theta - \sin\xi(\theta)) \right) \right|. \tag{2.9}$$

There remains to update the parameter μ involved in eq. (2.1) with respect to $\beta(\phi, \theta)$, and account for the limb darkening along the horizontal direction:

$$\mu(\phi, \theta) = \cos\left(\sqrt{\beta^2(\phi, \theta) + \arcsin^2\left(\frac{d_{SE} \cdot \phi}{R_S} \right)} \right) \tag{2.10}$$

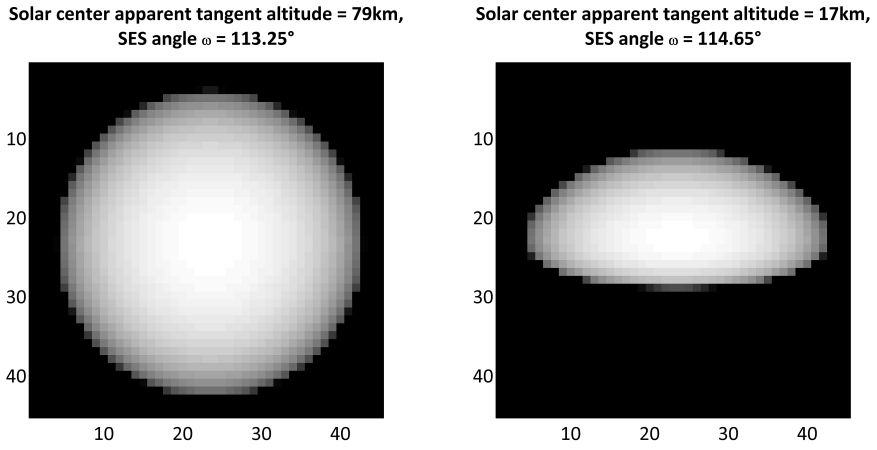

Figure 2.2 – Simulated pictures of the Sun. Left picture shows a high Sun, right picture shows a low Sun already experiencing strong flattening. Notice the asymmetrical flattening.

To illustrate the effect of the atmospheric refraction on the apparent solar disk, we simulated images of the Sun for different Sun-Earth-spacecraft arguments ω. Two of them are shown in Fig. 2.2. When ω increases during the occultation, the light ray bending increases also due to higher air density resulting in an asymmetrical flattening of the Sun image as light rays from the lower part of the Sun are more bent than those from the higher part.

The Sun images were computed according to the following procedure:

- For a given state of the atmosphere (pressure, temperature, humidity, ...) the total refraction angle $2\delta(h)$ is computed for all tangent altitudes using eq. (2.6) (the computation grid size is 10 m, with interpolation of the geophysical quantities on this finer grid).

- The intensity received by a pixel depends on the solar limb darkening of the region of the solar disk it is looking at ($\mu(\phi, \theta)$), which in turn depends on the light ray emission angle ($\beta(\phi, \theta)$). Both quantities are computed on a much finer grid than the final digital image resolution (900 grid nodes per pixel typically). This is particularly important for pixels capturing the solar disk edge. Eq. (2.3) is solved iteratively to match θ with h using the precomputed $2\delta(h)$.

- The final pixel intensity is obtained by averaging the radiance com-

puted at the 900 grid nodes and using the assumed optical parameters of the virtual imager.

2.4 Zernike moments

As the goal of this method is to retrieve atmospheric pressure information from the apparent flattening of the solar disk, a unique quantitative description of the Sun shape is preferable. It is also necessary to have a method insensitive to spacecraft pointing uncertainties responsible for rotation and translation of the Sun in the FOV. For these reasons, we selected the Zernike polynomials.

Zernike polynomials [124] constitute an orthogonal basis of continuous functions defined on the unit disk. They are widely used in adaptive optics to correct for image deformation due to atmospheric effects in high precision telescopes for instance. The definition of the Zernike polynomial of order n and repetition m (n and m being always integers with $n \geq 0$ and $|m| \leq n$ with $n - |m|$ even) in polar coordinates (ρ, α) is

$$Z_n^m(\rho, \alpha) = R_n^m(\rho) \cdot e^{im\alpha}. \tag{2.11}$$

The radial function is given by the following polynomial:

$$R_n^m(\rho, \alpha) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s}. \tag{2.12}$$

One immediately notices that $Z_n^{m^*} = Z_n^{-m}$. Fig. 2.3 shows the Zernike polynomials computed on the unit disk up to $n = 4$.

For our application, Cartesian coordinates are preferred over polar, i.e. $Z_n^m(\rho, \alpha)$ is equivalent to $Z_n^m(x, y)$ with $x^2 + y^2 \leq 1$. By orthogonality of the Zernike functions, we have

$$\iint\limits_{x^2+y^2 \leq 1} Z_n^{m^*}(x, y) Z_j^k(x, y) \mathrm{d}x\mathrm{d}y = \frac{\pi}{n+1} \delta_{nj} \delta_{mk}. \tag{2.13}$$

When applying Zernike polynomials to images, one needs to discretize the functions to account for the discrete nature of digital pictures. This may lead to a loss of orthogonality. This problem was already discussed by [116] but an extensive study of the orthogonality requirement of the discrete Zernike polynomials has been performed by [2] where different grid geometries (Cartesian, polar, random, ...) and grid resolutions (number of grid nodes) were tested. The goal was to find the best discretization strategy, close to the perfect orthogonality of the continuous

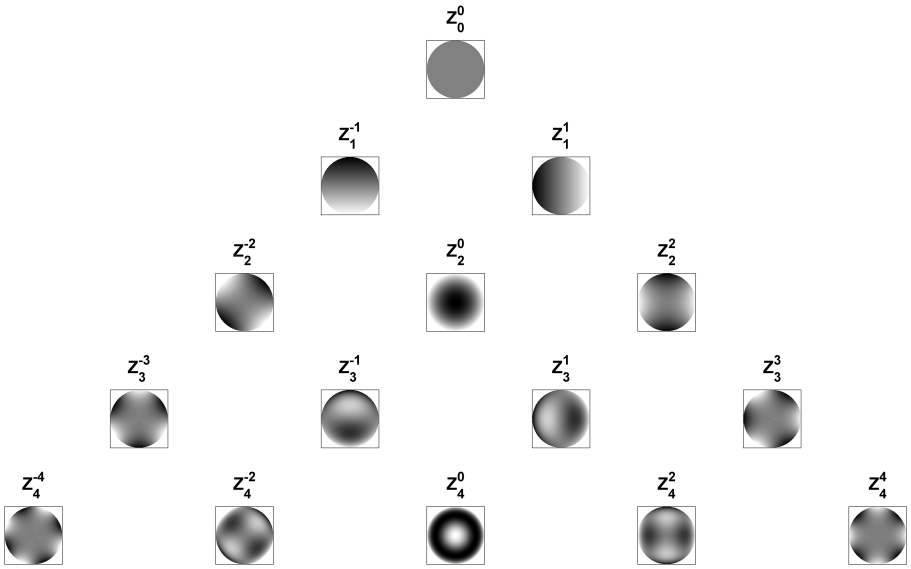

Figure 2.3 – Usual representation of the Zernike polynomials evaluated on the unit disk up to order $n = 4$. If $m > 0$, the real part is represented. If $m < 0$, the imaginary part is represented.

case. Allen showed, as expected, that a finer grid allows for better orthogonality. In the case of a Cartesian grid like in digital images, a rule of thumb for estimating the minimal number of grid points n_p ensuring *decent orthogonality* between the n_z first Zernike polynomials is to satisfy $\log_{10} n_\mathrm{p} > 1.73 \log_{10} n_\mathrm{z} + 0.27$. According to this rule, for applications where the 10 first Zernike polynomials are needed, the picture should be made of 100 pixels at least. In our case, images are cropped such that they closely fit the left and right edges of the solar disk, which makes the final array 45×45 pixels large. Not all of them belong to the unit circle, the pixels close to the corners are excluded from the Zernike domain, which leaves us with a bit more than 1500 pixels concerned. As will be shown later, we will not use many Zernike polynomials, so that the uniqueness of the decomposition will be guaranteed.

The moment of order n and repetition m of the image whose pixel intensities are represented by $f(x, y)$ is given by the following expression:

$$A_n^m = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) \cdot Z_n^{m^*}(x, y), \qquad x^2 + y^2 \leq 1. \tag{2.14}$$

There is no guarantee that the solar disk will always remain exactly at the same position in the field of view. It can be shown that Zernike moments are invariant under translation and rotation [65]: the Zernike

moment A_n^m of a given image becomes $A_n^{m^{\mathrm{r}}}$ after a rotation by an angle ϵ, and clearly

$$A_n^{m^{\mathrm{r}}} = A_n^m \cdot e^{-im\epsilon},$$

meaning that the moment modulus is preserved under rotation. A consequence of this property is that, as $|A_n^m| = |A_n^{-m}|$, it is sufficient to rely on moments with $m \geq 0$ only.

The translation invariance can be achieved by pre-processing the Sun image in order to have the centroid of the solar disk co-located with the center of the unit disk. It can be done by using the regular moments [59]. In their discrete form, they are defined by the following expression:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y). \tag{2.15}$$

In the Cartesian frame of the initial image, the centroid is located at $(\overline{x}, \overline{y})$ where

$$\overline{x} = \frac{m_{10}}{m_{00}}, \qquad \overline{y} = \frac{m_{01}}{m_{00}}.$$

In the translated coordinates frame (x', y'), the image is centered in such a way that $f(x', y') = f(x - \overline{x}, y - \overline{y})$.

It can be shown [65] that Zernike moments of order $n = 1$, being A_1^{-1} and A_1^1, are proportional to the regular moments m_{10} and m_{01}. Both of them vanish in the translated frame, resulting in the disappearance of the first order Zernike moments as well.

In Fig. 2.4, Zernike moments up to order 4 are computed for the high and low Sun cases. The first moment A_0^0 is by definition the total integrated intensity on the whole solar disk and is always dominant.

2.5 Pressure profiles retrieval

The central problem addressed by this method is retrieving pressure profiles from images of the Sun during occultation. The measurements are represented by a number of Zernike moments $A_n^m(\omega)$ computed for each image $I(\omega)$ acquired successively during sunset or sunrise, ω being the Sun-Earth-Spacecraft argument. From a given set of Zernike moments, the measurement vector $\boldsymbol{a}$ is written as:

$$\boldsymbol{a} = \begin{pmatrix} \boldsymbol{A}(\omega_1) \\ \vdots \\ \boldsymbol{A}(\omega_s) \end{pmatrix} \quad \text{with} \quad \boldsymbol{A}(\omega_i) = \begin{pmatrix} A_{n_1}^{m_1}(\omega_i) \\ \vdots \\ A_{n_k}^{m_k}(\omega_i) \end{pmatrix}, \tag{2.16}$$

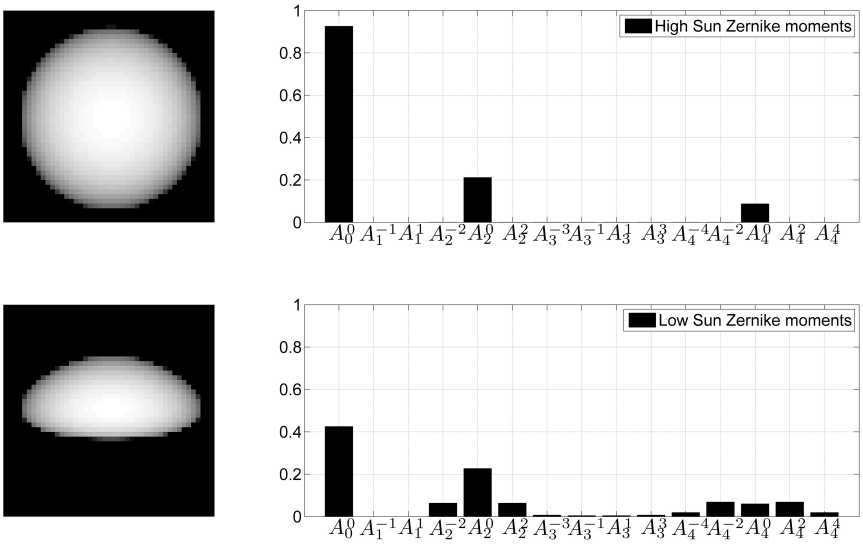

Figure 2.4 – First 15 Zernike moments computed on a high Sun image and a low Sun image (same as Fig. 2.2, arbitrary units). In the case of the high Sun image, moments with radial features ($m = 0$) are the only ones to play a role in characterizing this unrefracted solar disk. For the low Sun, more moments are needed to account for the asymmetrical flattening.

$A_{n_j}^{m_j}(\omega_i)$ being the j-th selected moment computed on the i-th image. The atmospheric pressure profile will be referred to as the state vector $\boldsymbol{p}$, made of 46 layers of constant pressure (US Standard atmosphere). The forward model can be simply written as:

$$\boldsymbol{a} = F(\boldsymbol{p}) \tag{2.17}$$

where F describes the ray propagation, the image acquisition and the moments computation.

Eq. (2.17) is not solved directly due to the non-linearity of operator F through the refraction process itself. Ill-conditioning might also be an issue: even if the number of image acquisitions is larger than the number of retrieved parameters, the problem might be ill-conditioned due to the redundancy between successive measurements. Furthermore, iterative approaches are not suitable as all the pictures of the setting or rising Sun need to be simulated at each iteration, which is not efficient regarding the processing time. To circumvent this difficulty, we preferred a retrieval strategy based on a lookup table. By simulating sunsets for a large number of plausible pressure profiles, we built a large set of direct relations between pressure profiles and associated Zernike moments. The pressure profiles used to create the lookup table constitute the training

dataset. If the training dataset is sufficiently exhaustive with respect to the natural variation of atmospheric pressure, then any pressure profile should be well captured by the lookup table.

2.5.1 Training dataset

The first step is to have a clear representation of the variety of realistic pressure states. We used the COSPAR International Reference Atmosphere model [64] referred in the following to CIRA-86. It consists of monthly averaged zonal pressure data on an altitude grid spacing of 5 km from 20 km to 120 km between $-80°$ and $80°$ of latitude by steps of $10°$, forming a dataset of 204 profiles. To comply with the US Standard atmosphere, the pressure profiles were extrapolated down to 0 km by setting a common value of 101 300 Pa at ground for all profiles. They were then interpolated onto the US Standard atmosphere grid using a spline method. The dataset consists of a matrix of 204 realizations (the modified CIRA-86 profiles) of 46 variables (the number of atmospheric layers). Let $\boldsymbol{p} = (p_1 \cdots p_q)$ be a pressure vector where $q = 46$ in this case. The whole dataset is represented by the matrix

$$\mathbf{P} = \begin{pmatrix} p_{11} & \cdots & p_{1q} \\ \vdots & & \vdots \\ p_{r1} & \cdots & p_{rq} \end{pmatrix},$$

with p_{ij} being the i-th realization of the j-th variable (i.e. the i-th pressure value of the j-th layer), where $i = 1, \ldots, r$ ($r = 204$ for CIRA-86) and $j = 1, \ldots, q$. The CIRA-86 climatology does reflect the annual and worldwide variability of atmospheric pressure states. Though, the averaging made the profiles probably smoother and removed the extrema. This makes the dataset a good starting point for testing the method, but not yet suitable for real cases.

2.5.1.1 Principal component analysis

The profiles may contain redundant structures between each other and it could be useful to describe the pressure profiles with a reduced number of parameters. We used Empirical Orthogonal Functions (EOF) [75], often referred to as Principal Component Analysis (PCA) [61] to solve these issues. The technique is based on finding orthogonal vectors (the principal axes) along the directions in which most of the dispersion of repeated realizations of the same phenomenon occurs. In the following paragraphs, we focus on finding the principal axes explaining most of the

variation of the atmospheric pressure profile dataset and on estimating the error of such a reduced representation.

The dataset is first centered about its mean. By averaging the realizations of each variable, we define

$$\overline{\mathbf{P}} = \begin{pmatrix} \overline{p_1} & \cdots & \overline{p_q} \\ \vdots & & \vdots \\ \overline{p_1} & \cdots & \overline{p_q} \end{pmatrix} \quad \text{where} \quad \overline{p_j} = \frac{1}{r}\sum_{i=1}^{r} p_{ij}.$$

The centered dataset is denoted by $\tilde{\mathbf{P}} = \mathbf{P} - \overline{\mathbf{P}}$. Then the data are normalized by their standard deviation. The covariance matrix of $\tilde{\mathbf{P}}$ is $\mathbf{\Sigma} = \tilde{\mathbf{P}}^{\mathrm{T}} \cdot \tilde{\mathbf{P}}$ with its elements being the σ_{ij}, $i, j = 1, \ldots, q$, and $\mathbf{\Sigma}_0$ is the same matrix but with all off-diagonal elements removed. The factor $1/(q-1)$ has been neglected. The normalized dataset is then

$$\mathbf{P}^* = \tilde{\mathbf{P}} \cdot \mathbf{\Sigma}_0^{-1/2} \tag{2.18}$$

whose elements are the $p^*_{ij} = (p_{ij} - \overline{p_j}) / \sqrt{\sigma_{jj}}$. By performing a singular value decomposition (SVD) on the normalized dataset, the eigenvectors (the principal axes) are identified and their associated singular values are sorted in a decreasing order. Defining the matrices $\mathbf{U}$, $\mathbf{L}$ and $\mathbf{V}$ such that $\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$, and $\mathbf{L}$ is diagonal, we have $\mathbf{P}^* = \mathbf{U}\mathbf{L}\mathbf{V}^{\mathrm{T}}$. The covariance matrix of $\mathbf{P}^*$ follows: $\mathbf{R} = {\mathbf{P}^*}^{\mathrm{T}}\mathbf{P}^* = \mathbf{V}\mathbf{L}\mathbf{L}\mathbf{V}^{\mathrm{T}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}$, where the columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{R}$, and the λ_{ii} of the diagonal matrix $\mathbf{\Lambda}$ are its squared singular values. From now on, the columns of $\mathbf{V}$ will be called the principal axes (PA) of the normalized dataset. The principal components (PC) of a given pressure profile are its coordinates in the orthogonal basis formed by the PA. Computing the PCs of the entire normalized dataset, we find the matrix of PCs $\mathbf{C}$ of size $r \times q$:

$$\mathbf{C} = \mathbf{P}^*\mathbf{V}.$$

Exploring the eigenvalues of the SVD, we find that the first ones already explain a very large portion of the overall dataset variability. Table 2.2 shows the eigenvalues and their cumulative importance for the 10 first PA. One notices that the first four PA are responsible for more than 99 % of the variance.

As the goal is to reduce the number of dimensions of the pressure profiles, we are looking for the minimum number of PA needed to allow for a reconstruction of the dataset profiles within acceptable errors. The satisfaction criterium is based on the global mean quadratic relative error (MQRE) that we required to be lower than 1 % for the full dataset. The

Table 2.2 – PCA applied to the CIRA-86 pressure dataset. The eigenvalues of the 10 first principal axes and their cumulative importance for capturing the overall dataset variability are given in the first two columns. The third column gives the total relative error made when the dataset is reconstructed from the first PAs only.

	Eigenvalue amplitude	**Cumulative importance (%)**	**MQRE (%)**
λ_1	37.1415	82.537	6.08
λ_2	5.7789	95.379	2.85
λ_3	1.5160	98.748	1.95
λ_4	0.3634	99.555	1.29
λ_5	0.1372	99.860	0.66
λ_6	0.0399	99.949	0.42
λ_7	0.0139	99.980	0.26
λ_8	0.0054	99.992	0.17
λ_9	0.0020	99.996	0.13
λ_{10}	0.0011	99.998	0.09

global MQRE is computed this way: from a number of m PA ($1 \leq m \leq q$), the profiles are reconstructed using the associated m PCs:

$$\begin{aligned} & \mathbf{P}^*_{(m)} = \mathbf{C}_{(m)}\mathbf{V}^{\mathrm{T}}_m \\ \Leftrightarrow & \left(\mathbf{P}_{(m)} - \overline{\mathbf{P}}\right) \cdot \mathbf{\Sigma}_0^{-1/2} = \mathbf{C}_{(m)}\mathbf{V}^{\mathrm{T}}_m \\ \Leftrightarrow & \mathbf{P}_{(m)} = \mathbf{C}_{(m)}\mathbf{V}^{\mathrm{T}}_m\mathbf{\Sigma}_0^{1/2} + \overline{\mathbf{P}} \end{aligned} \tag{2.19}$$

where $\mathbf{V}_m$ is a $q \times m$ matrix containing the m first PA, and the subscript (m) denotes quantities obtained from the truncated basis of these m first PA. The MQRE ($\overline{\epsilon}$) is computed by averaging the relative deviation between the initial data (p_{ij}) and the reconstructed ones ($p_{ij}^{(m)}$):

$$\overline{\epsilon}_{(m)} = \sqrt{\frac{1}{rq}\sum_{j=1}^{q}\sum_{i=1}^{r}\left(\frac{p_{ij} - p_{ij}^{(m)}}{p_{ij}}\right)^2} \cdot 100\,[\%] \tag{2.20}$$

The last column of Table 2.2 provides the MQRE computed for up to 10 PA. As can be seen, five PA are needed to go below the 1 % threshold.

When the 204 profiles of the CIRA-86 database are projected onto the five first principal axes, their PCs can be seen as a spectrum. Fig. 2.5 shows the spectrum of PCs for each principal axis. In order to build a training database that would span a wider range of atmospheric states and to avoid redundancies between profiles (as it is the case in CIRA-86), we used the principal component spectra to identify a small number

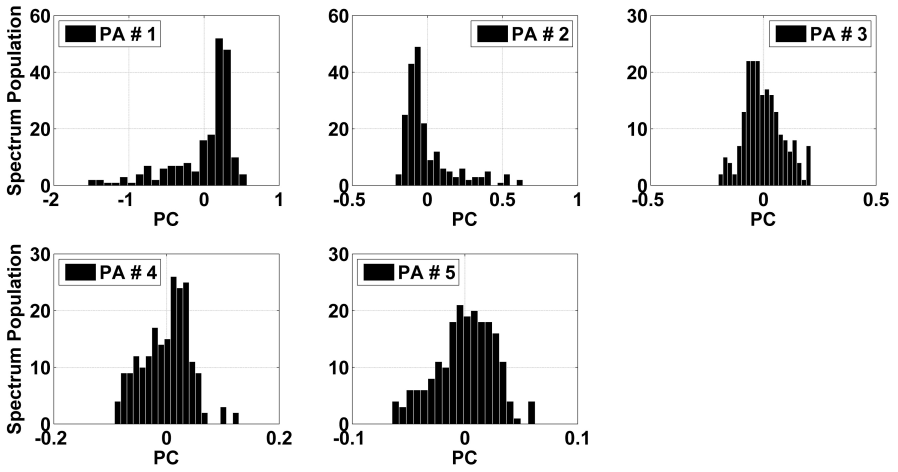

Figure 2.5 – Distribution of the PCs of the CIRA-86 dataset along the five first principal axes. From these distributions, pivots can be defined in order to generate a training dataset of non-redundant elements.

of pivot values from which an exhaustive but non-redundant dataset can be generated. Table 2.3 shows the coordinates of the pivots along the five principal axes and how they have been computed. By combining the pivots, one can create a new database of $4^2 \cdot 3^3 = 432$ different pressure profiles that constitutes the training dataset that will be used in the retrieval algorithm.

Table 2.3 – Coordinates of the pivots along the five first principal axes. The median $\tilde{m}_i$ and the standard deviation σ_i of each distribution ($i = 1, \ldots, 5$) were used to compute the pivots. They were chosen such as to span the spectrum with a reduced number of points. See Fig. 2.5 for a representation of the PC spectrum.

	Pivots			
	#1	**#2**	**#3**	**#4**
PA#1	$\tilde{m}_1 - 3\sigma_1 = -1.08$	$\tilde{m}_1 - 1\sigma_1 = -0.23$	$\tilde{m}_1 = 0.20$	$\tilde{m}_1 + 1\sigma_1 = 0.63$
PA#2	$\tilde{m}_2 - 1\sigma_2 = -0.23$	$\tilde{m}_2 = -0.066$	$\tilde{m}_2 + 1\sigma_2 = 0.10$	$\tilde{m}_2 + 3\sigma_2 = 0.44$
PA#3	$\tilde{m}_3 - 1\sigma_3 = -0.27$	$\tilde{m}_3 = -0.01$	$\tilde{m}_3 + 1\sigma_3 = 0.25$	
PA#4	$\tilde{m}_4 - 1\sigma_4 = -0.12$	$\tilde{m}_4 = 0.01$	$\tilde{m}_4 + 1\sigma_4 = 0.13$	
PA#5	$\tilde{m}_5 - 1\sigma_5 = -0.08$	$\tilde{m}_5 = 0.00$	$\tilde{m}_5 + 1\sigma_5 = 0.08$	

2.5.1.2 The transfer matrix

The measurement simulations consist of 23 pictures of the setting Sun. Each picture corresponds to a given value of ω, the Sun-Earth-Spacecraft argument: from 113.25° (high Sun) up to 115.45° (low Sun) by steps of 0.1°. The images are then projected onto the Zernike polynomial basis and the Zernike moments are computed. Only a small number k of them are kept for the training. If there are l image pixels falling within the circular domain of the Zernike functions, ordered in a column vector $\boldsymbol{f}(\omega_i, \boldsymbol{p}_j)$ (ω_i is the i-th angle during the sunset and $\boldsymbol{p}_j$ is the j-th pressure profile in the dataset), then in matrix formalism, the vector containing the Zernike moments is computed by

$$\boldsymbol{a}_{ij} = \underbrace{\boldsymbol{a}(\omega_i, \boldsymbol{p}_j)}_{k\times 1} = \underbrace{\mathbf{Z}}_{k\times l} \cdot \underbrace{\boldsymbol{f}(\omega_i, \boldsymbol{p}_j)}_{l\times 1}$$

where the rows of $\mathbf{Z}$ are the k selected Zernike polynomials evaluated at the l pixels lying in the circular domain of the image. The Zernike moments of each image and each pressure profile are then concatenated into a matrix $\mathbf{A}$ of size $[23 \cdot k \times 432]$ to form the measurement matrix of the training dataset.

The link between the measurements of matrix $\mathbf{A}$ and the 5 principal components describing the pressure profiles arranged in the matrix $\mathbf{C}$ of size $[5 \times 432]$ is expressed by the transfer matrix $\mathbf{X}$ of size $[5 \times 23 \cdot k]$:

$$\mathbf{C}_{\text{train}} = \mathbf{X} \cdot \mathbf{A}_{\text{train}} \tag{2.21}$$

Finally, the elements of $\mathbf{X}$ can be easily computed through the resolution of this overconstrained problem using a least-square regression method:

$$\mathbf{X} = \mathbf{C}_{\text{train}} \cdot \mathbf{A}_{\text{train}}^{\text{T}} \cdot \left(\mathbf{A}_{\text{train}} \cdot \mathbf{A}_{\text{train}}^{\text{T}}\right)^{-1} \tag{2.22}$$

2.5.2 Test cases

The training dataset was obtained from non-redundant combinations of key values (the pivots) taken in the spectrum of principal components of the CIRA-86 dataset. In that sense, not a single pressure profile of the training dataset is identical to those contained in CIRA-86. Though, it is correct to write that they are similar: the principal axes of both datasets are almost parallel.

As such, the pressure profiles of the CIRA-86 climatology constitute an optimal, though non-trivial, test case. The retrieved accuracy will

unveil the upper limits of the method. Results obtained with independent pressure states should exhibit larger biases, particularly when they are not accurately represented by a linear combination of the principal axes of the training dataset. In other words, the adopted retrieval strategy performs better when the sounded atmospheric state lies in the multi-dimensional space defined by the principal axes of the training dataset. Poor results mainly indicate that the training dataset is not comprehensive enough.

We illustrated this discussion with two test cases. First, we applied the method to the 204 pressure profiles of the CIRA-86 climatology. This is actually the optimal situation as they are almost perfectly represented in the basis of the training dataset principal axes. Second, we picked up 30 pressure profiles from the ACE-FTS v3.0 products of the occultations observed in 2011 [10], constituting a truly independent test sample. The results are presented and discussed below.

2.5.2.1 CIRA-86

Fig. 2.6 **(a)** shows the 204 pressure profiles of the CIRA-86 climatology (dark gray) above the 432 profiles of the training dataset (pale gray). A solar occultation has been simulated for each of the CIRA-86 atmospheric states. Zernike moments have been computed for each of the 23 pictures constituting a solar occultation measurement in order to form the measurement vector. Finally the transfer matrix allowed to retrieve the pressure profile from the Zernike moments. Figure 2.6 **(b)** shows the relative difference between the retrieved profiles and the original ones when only the first Zernike moment is used (A_0^0, the solar disk total intensity), while the first and the fifth moments are used together in Fig. 2.6 **(c)** (A_2^0 better captures the solar limb darkening).

The excellent results obtained below 20 km simply follow from the extrapolation of the CIRA-86 profiles from 20 km down to a common value at ground (101300 Pa). This extrapolation was then inherently reflected in the training dataset, so that the least-square fit cannot be mistaken in this altitude range and it should be overlooked. Between 20 and 60 km however, the average bias is almost 0 with increasing spread though. Figure 2.6 **(d)** shows the mean relative difference $\pm 1\sigma$ for the retrieval performed with two Zernike moments. In this altitude range, the reasons of the discrepancy are twofold. First, not all the profiles are well described by the training dataset principal axes (though very good for most of them). Second, the solar disk flattening is decreasingly measurable with higher altitudes given the spatial resolution of the virtual

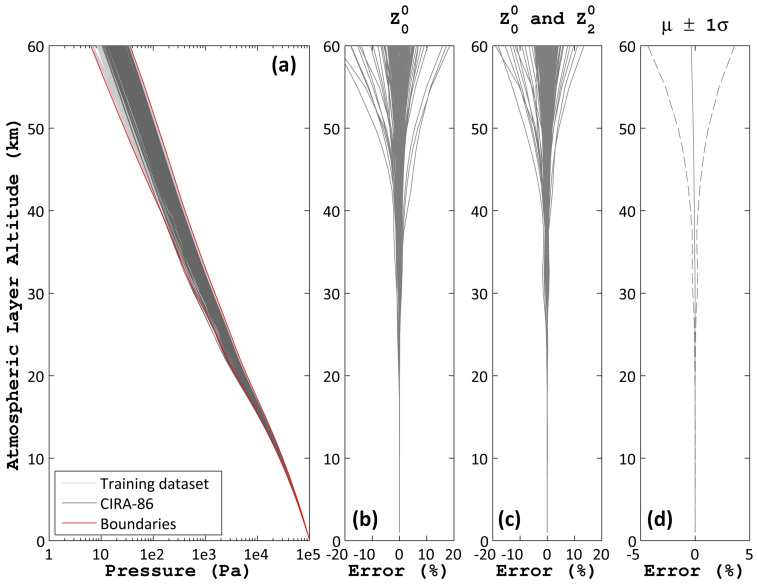

Figure 2.6 – Retrieval algorithm applied to the simulated sunsets for each of the 204 pressure profiles of the CIRA-86 climatology. The first panel shows the CIRA-86 profiles (dark gray) above the training dataset profiles (pale gray) on a logarithmic scale. The red lines emphasize the boundaries of the training dataset. The second panel shows the relative error of the retrieved pressure in function of altitude by using information from the first Zernike polynomial only. Third panel shows the relative error obtained with the first and fifth Zernike polynomials. The last panel shows the mean relative error $\mu \pm 1\sigma$ for the two Zernike polynomials retrieval case.

imager. This is reflected in the Zernike moments as their amplitude becomes less and less dependent on the atmospheric state as the air density decreases. Fig. 2.7 **(a)** shows the evolution of the first Zernike moment during the sunset for all 204 pressure profiles of CIRA-86. Above 60 km, all pictures give the same moment amplitude, whatever the pressure in the upper layers. Below 60 km, different pressure profiles are characterized by distinct moment values. A measure of this distinction can be done by comparing the standard deviation of the moment population with the mean amplitude (Fig. 2.7 **(b)**). As expected, this ratio falls down to zero above 60 km: the spatial resolution of the virtual imager is too small and the refraction too weak.

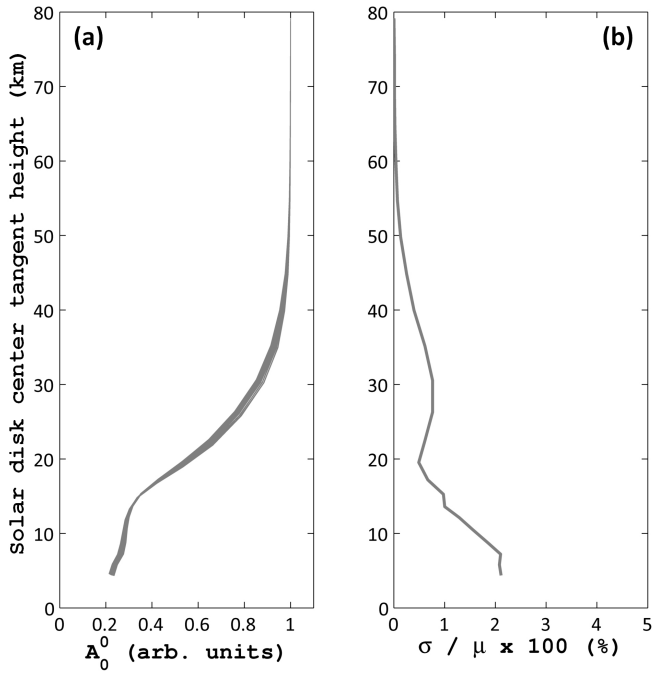

Figure 2.7 – Left panel shows the moment amplitude of the first Zernike polynomial computed for every simulated solar disk image and every profile of the CIRA-86 climatology. At the beginning of the occultation, the weakness of the refraction makes the Zernike moment insensitive to the particular pressure of the upper atmospheric layers. At lower altitudes, the moment starts taking different values for different pressure profiles. This is emphasized in the second panel where the standard deviation of the moment amplitude is compared to the mean value.

2.5.2.2 ACE-FTS pressure profiles sample

We picked 30 pressure profiles among the solar occultations performed by ACE-FTS in 2011. They originate from the version 3.0 of the level 2 products, sampled on an altitude grid from 5.5 km to 89.5 km by steps of 1 km. In order to fit with our atmospheric model, they were extrapolated towards 0 km (101300 Pa) and 100 km (0.03 Pa) and interpolated on the US Standard grid. The ACE-FTS profiles were selected at random among those lying between the boundaries of the training dataset (Fig. 2.8 **(a)**). Unlike the CIRA-86 profiles, they are truly independent from the training dataset. Their coherence with the training dataset can be assessed by computing their principal components and reconstructing the profiles from the 5 first principal axes (as it was done to check

the number of independent variables needed to describe the CIRA-86 profiles). Fig. 2.8 **(b)** shows the relative error paid by doing so. Clearly, the basis extracted from the CIRA-86 climatology is not suitable for any kind of profile. As expected, by simulating sunsets for these 30 profiles, and applying the retrieval algorithm, we obtained poor results (Fig. 2.8 **(c)**).

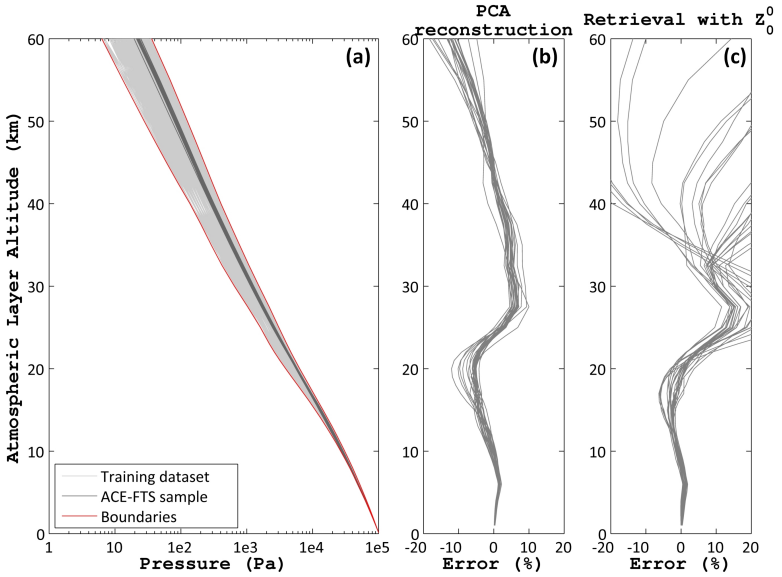

Figure 2.8 – First panel shows the ACE-FTS sample of 30 pressure profiles (dark gray) above the training dataset (pale gray). The red lines emphasize the boundaries of the training dataset. Second panel shows the relative difference between the original profiles and those reconstructed from a projection onto the principal axes of the training dataset. The third panel shows the relative difference between the original profiles and the retrieved ones using the first Zernike moment.

2.5.2.3 Discussion of the results obtained with the two test cases

Two test cases have been presented to assess the limitations of this method. The first one gives a flavor of what can be achieved at best with the assumed instrument specifications. For most of these atmospheric states, the pressure profile is retrieved up to the lower mesosphere with less than 5% relative error (less than 1% below 30 km). Better spatial resolution should probably lift the detection limit a few kilometers upward.

The second test case revealed a critical aspect of the retrieval strategy: the suitability of the training dataset (through its basis of PA) with the considered pressure profiles must be optimized. Clearly, the smoothed profiles of the CIRA-86 climatology did not turn out to be very representative of the ACE-FTS pressure profiles. Despite this lack of representativeness, the algorithm managed to deliver acceptable results on average for the lowest layers. For instance, at 20 km, the mean relative error is close to $1 \pm 1.8\%$ whereas the 3σ-variability of CIRA-86 is close to 25%.

2.5.3 Impact of error sources

In real conditions, many sources of error should be included in the development with a major contribution from shot noise and detector dark current. The error budget can be computed by starting from the covariance matrix of each image: a diagonal matrix $\mathbf{S}_{\boldsymbol{f}_i}$ of size $[l \times l]$ containing the shot noises of each pixel belonging to the circle domain. From there, $\mathbf{S}_{\boldsymbol{a}_i}$, the covariance matrix of $\boldsymbol{a}_i$, of size $[k \times k]$ follows:

$$\mathbf{S}_{a_i} = \mathbf{Z} \cdot \mathbf{S}_{\boldsymbol{f}_i} \cdot \mathbf{Z}^{\mathrm{T}}. \tag{2.23}$$

The covariance matrix of the measurement vector (the Zernike moments of the 23 images) is formed by concatenating the $\mathbf{S}_{\boldsymbol{a}_i}$ to form a block-diagonal matrix $\mathbf{S}_{\boldsymbol{a}}$. The covariance matrix of the retrieved PCs is then given by:

$$\mathbf{S}_{\mathrm{C}} = \mathbf{X} \cdot \mathbf{S}_{\boldsymbol{a}} \cdot \mathbf{X}^{\mathrm{T}}. \tag{2.24}$$

Finally, the covariance matrix of the error on the pressure profile is found using the same procedure on eq. (2.19):

$$\mathbf{S}_{\mathrm{P}} = \boldsymbol{\Sigma}_0^{1/2} \mathbf{V} \mathbf{S}_{\mathrm{C}} \mathbf{V}^{\mathrm{T}} \boldsymbol{\Sigma}_0^{1/2}. \tag{2.25}$$

In order to estimate the expected retrieval uncertainty, we simulated a test case where the brightest pixel of the solar pictures receives 10 000 counts and suffers from a dark current of 500 counts. Taking into account the associated shot noise, we used the above formula to derive the retrieval error by taking the square root of the diagonal elements of $\mathbf{S}_{\mathrm{P}}$. By examining the relative importance of this uncertainty to the profile, we find values ranging typically from 0.001 % for the lower layers, to 10 % for the highest ones. These typical numbers have to be compared with the natural variability of the pressure profiles as can be computed using the CIRA-86 dataset: a few percent at lower altitudes to a hundred percent at the highest ones. This exercise, though very simple,

allows to conclude that thanks to the high signal to noise ratio of the measurement vector (the Zernike moments use all the pixels of the solar disk resulting in a huge amount of signal), the retrieval uncertainty is expected to be 10 to 500 times smaller than the natural variability of the retrieved quantity.

Obviously, shot noise is the easiest and more predictable source of uncertainty. Other typical imager issues are related to the optics performance (blur, ghost images, distortion,...) or detector features (dead or hot pixels, pixel response non-uniformity,...). Each of them will require careful consideration in future developments but the inherent properties of the method lead to some general comments.

- Fixed patterns in the array of pixels or systematic distortions are image features which are captured by the Zernike moments. In some situations, only specific moments are affected so that a workaround is simply not to include them in the retrieval.
- If all moments are corrupted (like in the case of large distortion), comparing the distorted moment spectrum with the clean simulated case will allow to isolate what depends on the refraction (flattening), and what is due to the optics. The training dataset can then be adapted to take the instrumental effects into account.
- Non-systematic distortions like the presence of a cloud (PSC or PMC) in the field of view cannot be corrected and make the image unusable for the proposed method. However, when considering the question of reliable automatic flagging of solar pictures corrupted by clouds [40], the Zernike moments turn out to be extremely efficient as their amplitude will be considerably affected in an unusual way with respect to the cloud-free situation.

2.6 Adequacy of the proposed method and future work

This theoretical study was performed in order to assess the feasibility of retrieving atmospheric pressure profiles up to the lower mesosphere by means of solar disk image analysis. It relies on an accurate simulation of the apparent solar disk (including solar limb darkening and refractive effects), the use of the Zernike polynomials as orthogonal functions to measure the apparent flattening quantitatively, the reduction of the number of variables describing the pressure profiles by using a principal component analysis, and the construction of a comprehensive dataset to

link the Zernike moments of the sunset pictures to the pressure profile. Its intrinsic practical advantages are: the absence of strong pointing requirements (as the Sun is a very intense light source, one only needs to take pictures with sufficiently small integration times to have sharp solar disk images and to keep it in the instrument field of view for the complete sunset/sunrise), high signal to noise ratios (all the pixels sampling the solar disk are used) and potentially applicable to any solar occultation imager dataset provided that the images have been sent to ground in their entirety. The Zernike moment computation could even be performed onboard, requiring the download of a few numbers only (the Zernike moments themselves) instead of the full pixel map intensity. This could be an interesting feature for systems having small downlink capacity.

The limitations of the method have been assessed by two test cases. Pressure profiles particularly well represented by the training dataset show retrieval relative error smaller than $1 \pm 5\%$ up to 60 km. It constitutes what can be best achieved with this method, except if an instrument with higher spatial sampling is used. In that case, results could be slightly improved. Profiles that are not well represented by the training dataset are less well retrieved, though acceptable errors are found up to 20 km. The main point here is the need for a more comprehensive dataset with its principal axes allowing to correctly reconstruct most of the real pressure profiles.

Future investigations should address the issue of a training dataset showing better representativeness of the variability of natural pressure profiles. Other developments should also focus on the vertical resolution of the retrieved profiles, consider inhomogeneous atmospheric layers and account for possible instrumental effects.

Chapter 3

Spectral imaging and atmospheric remote sensing — The ALTIUS mission

3.1 The pointing problem of atmospheric profiling instruments

Atmospheric remote sensing refers to a large set of measurement techniques addressing different products. Those able to resolve the vertical distribution of a species concentration can be divided into two families: those tracking atomic and molecular absorptions, and those looking at emissions. While both processes occur over the entire spectral range (from UV to millimeter), instruments relying on these processes tend to work in different spectral domains (absorption mainly in the UV-VIS-SWIR, emissions mainly in the MWIR-mm). The source of the signal is completely different too. The first type of instruments collects light originating from a celestial body (Sun, star, Moon, planets) that was directly transmitted through or scattered by the atmosphere, but the spectrum of the body itself needs to be removed. The second family of instruments captures emissions from molecules within the atmosphere and must subtract possible background signals.

Each approach has its own advantages and drawbacks. Table 3.1 compares their respective performance for high-level aspects such as the coverage, the number of measurements or the geophysical targets. Of course, besides their filiation to one of these general categories, the number of measurable species and the performance are linked to the instrumental concept and considerably vary from one mission to the other.

Independently of the kind of signal detected (absorption or emission),

Table 3.1 – High level comparison between several types of atmospheric instruments with vertical resolving capabilities. Abbreviations used: solar occultation (So. Oc.), Stellar and planetary occultation (St. Oc.), scattering (Sc.), emission (Em.), terminator (term.). Crosses indicate potential measurement capability.

	Absorption			**Emission**
	So. Oc.	St. Oc.	Limb Sc.	Limb Em.
Local time coverage	term.	night	day	day/night
Spatial coverage	limited	scattered	global	global
# profiles/day	30	100	1000	2000
Altitude range (km)	15-100	15-100	15-50	5-100
Vertical sampling (km)	< 1	< 2	< 3	< 5
O_3 chemistry species	x	x	x	x
Greenhouse gases	x	x	x	x
Aerosols	x	x	x	

the capability of resolving the species concentration profile is often provided by a tracking (for occultation instruments) or a scanning system (for limb scattering and emissions).

Most of the time, it involves the use of a movable mirror, but it can also happen that the entire spacecraft performs the manoeuvre (OSIRIS onboard ODIN f.i., [73]). Clearly, occultation measurements offer the highest accuracy in determining the tangent point of the light path (the point of shortest distance between the light trajectory and the Earth surface). On the other side, instruments looking at the limb do not benefit from a bright and sharp irradiance source. They look at a diffuse scene whose brightness is globally exponentially decreasing with altitude. Typical pointing performance for those kinds of instruments is given in Table 3.2, where *accuracy* refers to the systematic bias

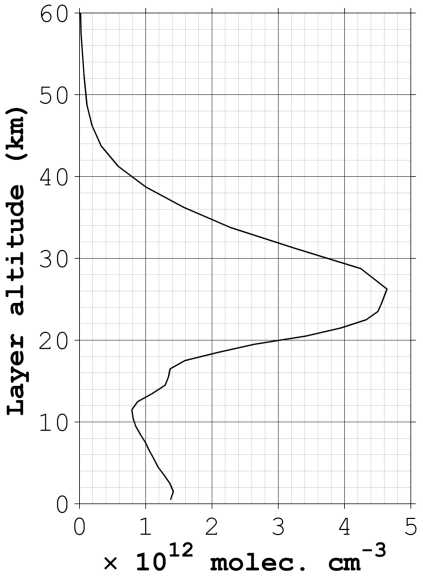

Figure 3.1 – Typical mid-latitude ozone profile.

`Table 3.2` – Pointing performance of limb sounding instruments. The first two record limb-scattered radiances (absorption), the third one measures emissions in the mm range.

Instrument	**OSIRIS**	**SCIAMACHY**	**MLS**
Spacecraft	ODIN	ENVISAT	EOS AURA
accuracy (m)	100	150	500
precision (m)	400	200	15
reference	[81]	[92]	[117]

and *precision* refers to the random error (jitter). Note that these values are not frozen in time: continuous efforts are made to reduce them, sometimes even after the instrument loss, since both hardware and software solutions can correct pointing offsets. In the case of SCIAMACHY for instance, which suffered from strong biases in the early years, the performance improved regularly at the cost of several complete dataset re-analyses [51].

Pointing is often considered as the major source of error in retrieved profiles. It is known that an offset of 500 m immediately leads to an error of 10% in ozone concentration above and below the peaking altitude of the ozone layer [114]. This is due to the strong gradient in ozone concentration, as can be seen in Fig. 3.1. In the context of ozone recovery monitoring (see section 1.2.1), an instrument aiming at detecting stratospheric trends of 2%/decade should therefore have a pointing accuracy drift of less than about 100 m per decade. To date, this seems still quite challenging.

3.1.1 Imaging as an answer to pointing uncertainty

Though many other atmospheric instruments call themselves *imager* because they sequentially record slices of the atmosphere (either by mechanical scan (whisk broom technique), either by orbital movement (push broom technique)), they do not solve the issue of the pointing uncertainty between successive acquisitions. By design, instantaneous full 2D spectral snapshots offer four direct advantages:

- The inter-pixel pointing uncertainty reduces to thermoelastic deformation and optical imperfections, as a 2D solid state detector is used with both dimensions corresponding to spatial coordinates (in many other instruments, where a grating is used, one dimension is for space, the other for the spectrum).
- Modern pixel arrays allow for a much higher vertical sampling

than usually achieved with scanning mechanisms. It can easily cope with ten times more sub-divisions of the atmosphere. Fine structures then become observable.

- The geographical sampling integrity is preserved as all the tangent points of one image are located above the same geolocation (the scanning of the atmosphere can often take 20 to 60 seconds at a typical orbital speed of 7 km/s).
- Several pointing calibration techniques arise from the instant imaging capability, such as background stars or geographical details detection.

3.2 The ALTIUS concept

Understanding the fundamental asset of the imaging capability in terms of pointing knowledge and geophysical performance, we can now elaborate on the overall ALTIUS concept. This concept is based on three axes: an **imager** coupled to a **tunable spectral filter** (AOTF) onboard an **agile platform** (PROBA). Table 3.3 lists the main points which are further elaborated in the following paragraphs.

Table 3.3 – The three axes making the ALTIUS concept and their key points summarized.

Imager	**AOTF**	**PROBA**
-minimization of pointing errors	-freedom of wavelength selection over broad spectral range	-manœuvrability
-image-based pointing calibration schemes	-moderate spectral resolution	-large telemetry and data processing capabilities
-direct interpretation of signals	-robust and compact	-good attitude knowledge and control
-occultations more easily performed	-independent straylight measurement	-allow multimode scenarii

We have discussed above several advantages of an imager. Here we repeat them and add a few more. An imager considerably *reduces the inter-pixel pointing uncertainty*, as the distance between the detector rows is fixed and each row looks at a different tangent altitude. It offers

simple but very efficient independent *means of pointing calibration*, by capturing background patterns such as stars through the highest atmospheric layers, or geographical details in cloud-free regions. It makes the *occultation measurements much easier* to perform, as the large field of view (FOV) alleviates the need for a dynamical tracking system. Finally, for post-processing aspects, the *flagging of bad data* is facilitated when it comes to identifying stratospheric or mesospheric clouds whose signature is not so evident to detect looking only at radiance profiles.

The second axis consists of the use of a tunable spectral filter. Obviously, the remote sensing of atmospheric species relies crucially on the identification of their spectral signature. Depending on the molecules, relatively high spectral resolution can be required. A few technologies can be used in combination with an imaging system. They share one feature: the instantaneous filtering of a large number of resolvable spots. Here are the most important types of instrument: basic interference filters of course, but also actuated Fabry-Perot interferometers (F-P) [3], liquid crystal tunable filters (LCTF) [52], Lyot filters [1], Fourier-transform spectrometers (FTS) [96] or acousto-optical tunable filters (AOTF). The latter is the technology selected for ALTIUS and it will be further described in Chapter 4. At the exception of the interference filters whose passband and central wavelengths are frozen very early in the design phase, the other technologies are truly tunable (F-P, LCTF, Lyot, AOTF) in the sense that the position of their passband can be shifted continuously over a broad spectral range, or even provide the full spectrum along with the image (imaging FTS).

The imaging FTS is probably the most powerful solution, but it is also the most expensive and complicated instrument. Tunable filters were therefore the good compromise. Having the freedom of tuning the wavelength in-flight is an interesting feature as it allows to *adapt the measurement strategy* in case it is found later that other combinations of wavelengths perform better. The AOTF is a small robust device made of a piezo-electric transducer bonded to a birefringent crystal. Compared to F-P, LCTF or Lyot filters, it is able to deliver narrower bandwidths while preserving a non-negligible angular aperture (up to 10°). The tuning time is of the order of a few μs. On top of that, the AOTF principle has a great advantage: the transmitted light has its polarization rotated by 90°, and is spatially separated from the incident beam. An instrument based on an AOTF has therefore the unique capability of being *able to measure its own straylight* independently: by switching the AOTF off, the acquired image only contains the parasitic light which can be later subtracted from the measurements (see Chapter

5 for a specific example).

The third axis is about the platform. In the past, multi-payload satellites of ever increasing size have been launched. The climax was reached with ENVISAT, the largest civil satellite ever deployed in space carrying ten instruments (developed by ESA, launched in 2002 and weighing 8.2 tons). The recent trend is nowadays not to take such risks anymore (risks at launch and of major failure during operations in orbit). A good example is the replacement of ENVISAT by a suite of five smaller platforms (the Sentinel program). Things are evolving even further as micro-, nano- and pico-satellites (< 100 kg, < 10 kg and < 1 kg respectively) are becoming platform standards.

One driver of the recent interest for these small platforms is of course the launch costs, as many such small volumes can fit into a single rocket. In the case of ALTIUS, the driver was the good volume-performance ratio found with the PROBA platform. Another (economical and political) advantage was that it is designed and manufactured in Belgium by QinetiQ Space.

PROBA stands for PRoject for OnBoard Autonomy, a family of experimental microsatellites counting already three members flying: PROBA-1 (launched in 2001, taking high resolution pictures of the Earth), PROBA-2 (launched in 2009, carrying four instruments for solar physics and space weather), and PROBA-V (launched in 2013, hosting one main payload dedicated to mapping land cover and vegetation, and several other smaller experiments).

One of the key features of this platform is its attitude and orbit control system (AOCS). Based on a set of star trackers, magnetotorquers and reaction wheels, it is designed to ensure an absolute pointing accuracy of 75 arcsec and a precision of 10 arcsec (2σ confidence level). Among the three flying missions, PROBA-2 is the closest to ALTIUS in terms of pointing requirements. Analysis of in-flight AOCS performance [31] reveals that the actual figures are better: an accuracy smaller than 50 arcsec and a jitter smaller than 5 arcsec are obtained most of the time.

Another key feature of the PROBA platform is its ability to rotate about three orthogonal axes at an average angular velocity of 1° per second. Having ALTIUS as only payload onboard, it offers the opportunity to perform measurements in different geometries. Within the limits of the electrical power resources, bright limb observations followed by solar and stellar occultations become feasible in one orbit, extending the global coverage of the mission and the list of target species.

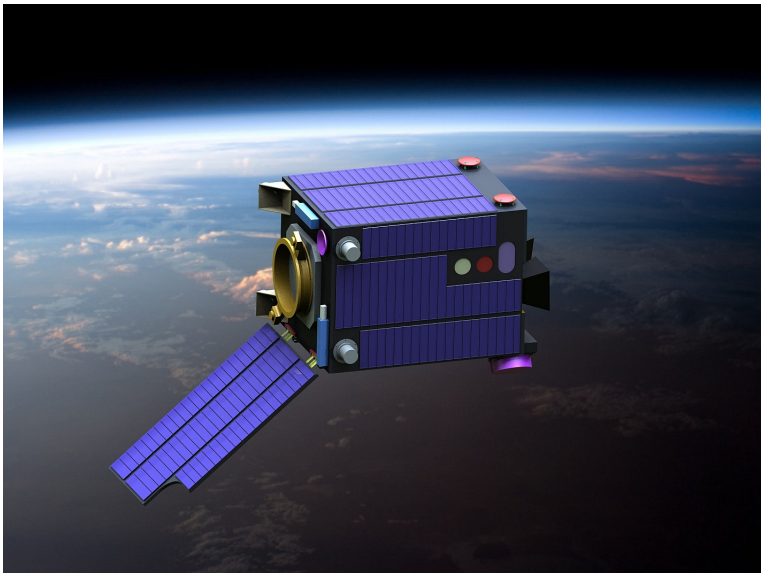

Figure 3.2 – Initial view of a PROBA platform hosting ALTIUS.

In summary, ALTIUS is a global atmospheric remote sensing concept. It consists of an agile platform capable of changing its attitude to allow for several measurement geometries, and maintaining an accurate pointing. Onboard this platform lies an imager which greatly helps in the burden of the tangent height registration of bright limb data. At the heart of this imager a tunable filter enables the identification of various molecular spectral signatures. The capabilities of this concept are discussed in the following sections.

We will first describe a nominal orbit. Then we will focus on the measurement mode which will provide most of the scientific data: the limb scattering geometry. Some simulations will be performed to illustrate the typical signal levels that are to be expected in flight. Then, starting from the ALTIUS mission requirements, we will simulate some measurements from which O_3 and NO_2 concentration profiles can be retrieved. Finally, we will test the mission requirements by performing a sensitivity study on the retrieved profiles by varying the most important parameters: the pointing and spectral accuracy, and the measurement noise.

3.3 Nominal orbit and measurement geometries

ALTIUS aims at providing information on the vertical distribution of key atmospheric species concentration. To do so, it relies on a set of spectral imagers and on a spacecraft showing good attitude control capabilities to improve on the pointing uncertainty aspect. Its primary measurement geometry is the observation of limb scattered radiance as it allows for global coverage. But the agility of the PROBA platform combined to the imaging feature make solar, stellar and planetary occultations easily achievable. The large field of view relaxes the pointing constraints and celestial objects are captured in inertial pointing (not compensating for the orbital movement). ALTIUS is therefore the only mission dedicated to atmospheric profiling, working in the UV-VIS-NIR, capable of measurements in the bright side, at the terminator and in the dark side of the orbit. Added values are: a wider coverage, access to photosensitive species in different illumination conditions (NO_2 for instance is strongly depleted during day but recovers at night), the opportunity of self-validation campaigns (probing the same air masses with different techniques in a short timescale),...

Over one nominal orbit, ALTIUS will perform measurements in three modes: atmospheric limb scattering (LS), solar occultations (SoO) and stellar occultations (StO), the latter encompassing planetary and lunar occultations as well. An orbit propagator was developed at IASB-BIRA. Coupled to a star catalogue, it allows for predicting the occurence and the tangent location of star-set and star-rise at any time. For illustrative purpose, we show below a coverage map computed for three consecutive days of nominal operations (Fig. 3.3). Each spot corresponds to the tangent point of the central pixel LOS of one image taken by the VIS channel. The star/planet selection algorithm takes into account the object magnitude and its angular distance from the current pointing direction. No stars of magnitude > 3 are considered.

In limb scattering mode, the number of atmospheric species profiles which can be retrieved over one orbit depends in a first place on the optical throughput of the instrument. If large enough, the exposure time needed for an image to reach its signal-to-noise ratio requirements is reduced, clearing some time for more acquisitions. In practice however, the electrical power budget is restricting the use of the three channels and a balanced work load must be found. In occultation modes, the question is less relevant as it completely depends on the availability of celestial bodies.

As it will be described later, each target atmospheric species retrieval

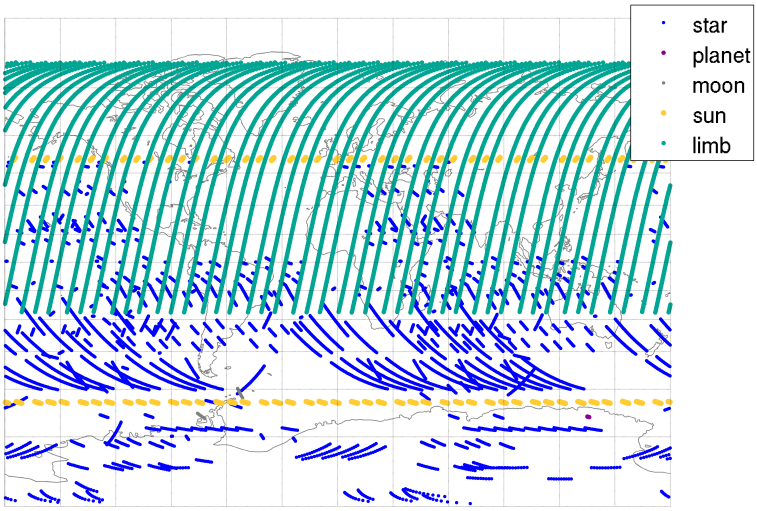

Figure 3.3 – Illustration of ALTIUS coverage after 3 days of nominal operations started on the 21st of June 2014.

requires two or more spectral images. A nominal orbit aims at meeting the highest priority scientific mission objectives: ozone of course, but also nitrogen dioxide, aerosols, water vapor and methane. The acquisition of the set of spectral images required to fulfill these objectives is referred to as performing one *observation*. The total number of images it will contain is not frozen yet, but five snapshots per channel is a reasonable estimation. The typical exposure times depend on the measurement geometry: one to two seconds in bright limb, ten times less during sunset/sunrise, and half a second in stellar occultation. Table 3.4 shows a tentative observational budget for ALTIUS.

Table 3.4 – ALTIUS preliminary observational budget for one nominal orbit.

	LS	**SoO**	**StO**[a]
# occurence/orbit	1	2	10
# obs./occurence	50	50	25
# image/ch./obs.	5	5	5
typical exp. time	1s	0.1s	0.5s
total # obs.	50	100	250
total # images	750	1500	3750
total # profiles[b]	250	10	50

[a] Also for planetary and lunar occultations.
[b] Assuming that one LS observation or one occultation allows to retrieve the five nominal species: O_3, NO_2, aerosols, H_2O and CH_4.

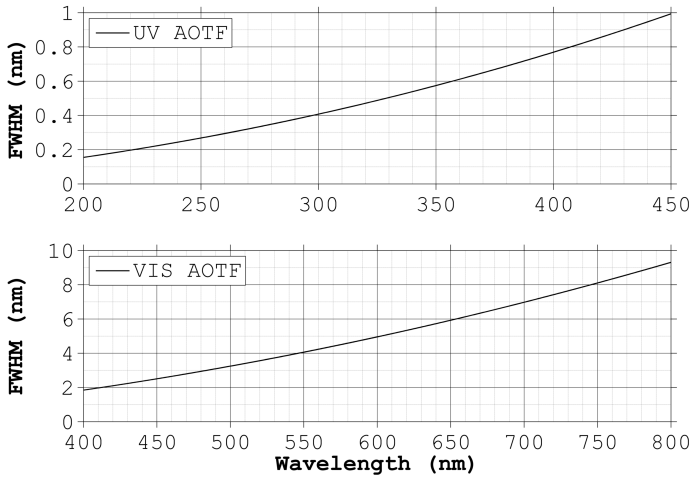

Figure 3.4 – Dispersion of the FWHM of the AOTF spectral transfer function for the UV and VIS channels.

3.4 Synthetic limb-scatter radiance

An end-to-end simulator is needed in order to assess the implications of the design choices on the quality of the final scientific products. The final simulator will accurately model all the steps connecting the celestial body radiance (Sun, stars, Moon, planets) out of atmosphere to the final retrieved profiles. It will involve a vectorial multiple scattering radiative transfer code, an accurate representation of the instrument transfer function (spectral and spatial), a model of the detectors counting efficiency and their read out electronics, a quantification of all the error sources and an inverse problem solver.

Our part of this huge task has been to estimate the quality of the inverted profiles regarding the payload and platform requirements. To do so, a downsized simulator has been used: the radiance levels have been computed by a state-of-the art scalar radiative transfer code, the signals (in terms of photon counts/second/pixel) have been derived from the knowledge of the spectral filter bandwidth, the total field of view and the number of pixels sampling the scene. The bandwidth and the spatial sampling are given in Fig. 3.4 and Table 3.5 respectively. The NIR channel specifications are not shown as it was not necessary for the subsequent simulations (owing to the spectral features of the tackled species).

The radiative transfer model (RTM) used is MODTRAN 5 [7]. It

Table 3.5 – Detector spatial specifications for two ALTIUS channels. The number of pixels and the corresponding sampling at the tangent point are indicated.

Channel	Detector	Spatial sampling
UV	170×170	590×590 m
VIS	512×512	195×195 m

offers sufficiently good accuracy in terms of scattering and absorption modelling, especially in our case where normalized radiances will be used (i.e. multiple scattering and albedo effects are reduced by normalization with high-altitude measurements). The molecular absorption spectral database is HITRAN [95], which is retrieved at a resolution of 1 cm^{-1}.

The RTM was set to reproduce ALTIUS viewing conditions in bright limb. The atmosphere was modelled as a stack of layers of homogeneous concentrations. The vertical division followed the U.S. 1976 standard atmosphere (US1976), which consists of layers of 1 km thickness from the surface up to 25 km, then 2.5 km up to 50 km, and 5 km above. Otherwise specified, the species concentration profiles have been interpolated onto that grid from the COSPAR International Reference Atmosphere 1986 database (CIRA86) [64].

All the retrieval simulations presented in the following sections rely on a common set of parameters constraining the RTM calculations. They essentially define the general scattering properties of the surface (albedo) and of the atmosphere (aerosol loading), and the orbit and LOS geometric aspects (day of the year, local time at tangent point, tangent point coordinates and LOS azimuth angle). They are summarized in Table 3.6.

Fig. 3.5 illustrates typical photon counting rates in the three channels when observing limb scattered radiance. They were obtained for tangent altitudes of 10 to 60 km by steps of 5 km. The radiances were convolved by Gaussian AOTF transfer functions.

3.5 Major payload and platform requirements

At the early stages of the ALTIUS project, the scientific payload performance and the mission requirements have been defined. They drive the overall instrument design and the mission profile. The achievement of the ALTIUS scientific objectives is directly related to the fulfillment

Table 3.6 – Relevant static MODTRAN parameters for the radiative transfer simulations.

Surface albedo	0.5
Surface temperature	15°C
Cloud top height	3 km
Tropospheric aerosol model	rural, 23 km visibility
Stratospheric aerosol model	background content
Atmospheric CO_2 content	400 ppmv
Day of the year	# 180
Tangent point coord.	45°N, 20°W
Tangent point local time	10:30 AM
LOS azimuth angle	180° (backward looking)

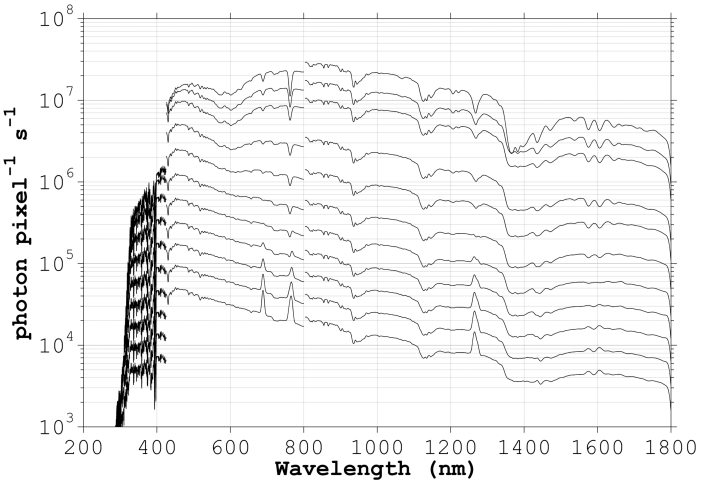

Figure 3.5 – Simulated flux of photons falling on one pixel in limb scattering geometry over mid-latitudes. Eleven tangent altitudes are represented: from 10 to 60 km by steps of 5 km. Curves can be labelled in this order starting from the highest one. The discontinuity occuring at the boundaries of the channels spectral domain is due to their difference of throughput. Optical transmission losses have been neglected.

of these requirements. The most important ones are the acceptable radiometric noise levels, the pointing uncertainty and the wavelength registration uncertainty. The first one is responsible for the size of the final error bars of the retrieved profiles, whereas the two others can induce systematic biases. Below, we give a little more insight into the

three error sources. Their quantitative requirements are summarized in Table 3.7-3.8.

3.5.1 Signal-to-noise ratio

The ALTIUS measurement principle is based on digital images acquisition. Photons of a given energy will hit the pixel photodetector, releasing electrons with a wavelength-dependent efficiency. The counting rate of electrons follows a Poisson distribution. In nominal operations (i.e. with a number of photons $N \gg 1$), it is well approximated by a normal distribution of mean $\mu = N$ and standard deviation $\sigma = \sqrt{N}$ (see Appendix A). This is called the *shot noise*, and it corresponds to the fundamental limitation of the signal-to-noise ratio: $\text{SNR} = \mu/\sigma = \sqrt{N}$. Of course, other sources of noise can add up and further reduce the SNR: read-out noise, dark current shot noise, thermal noise,... As low SNR means higher uncertainty on the measurements, threshold values have been defined independently of all the sources. Moreover, as the images will exhibit a high dynamic range (see Fig. 3.5), and as not all the wavelengths have the same importance for the retrieval, the SNR requirements vary with the wavelength and the tangent height (Table 3.8). The ways to meeting these values are an optimization of the optical throughput, the binning of pixels belonging to the same row (assuming rows are parallel to the horizon and therefore their pixels do not see different atmospheric composition), and larger integration times.

3.5.2 Pointing error

As already emphasized in the beginning of this chapter, pointing errors can easily become the largest source of bias in ozone retrieval. Following ESA standards, a number of performance indices have been used to separate the uncertainty contributors: errors on the spacecraft pointing performance itself, but also on the knowledge of the actual pointing, the drift over different timescales, the reproducibility of identical pointing commands,... However, referring to what is generally found in the litterature, concepts of *accuracy* and *precision* are still more meaningful. The pointing accuracy measures how far is the actual pointing from the target, whereas the pointing precision measures the jitter around the actual mean pointing direction. While the accuracy of the pointing can always be mitigated by a better calibration (and the imaging capabilities of ALTIUS are a clear asset here), the fundamental uncertainty lies in the precision. Obviously, for an atmospheric profiler such as ALTIUS, the most stringent requirements are assigned to the vertical dimension.

3.5.3 Wavelength misregistration

Atmospheric species often exhibit highly structured spectra, although these patterns are smoothed by the passband of the instrument. Grating spectrometers capture the radiance spectrum over broad wavelength ranges. A spectral imager, however, usually does not acquire continuous spectra (except if equipped with an imaging FTS): it takes snapshots at a limited number of wavelengths maximizing the sensitivity to the target species. A standard approach is to build ratios of weakly and strongly absorbing wavelengths. An error on the wavelength selection can induce a loss of sensitivity which in turn reduces the apparent optical thickness and leads to a smaller gas concentration. With the AOTF technology, such error will mainly arise from the uncertainty on the crystal temperature: as it will be shown in Chapter 4, a drift of 1°C shifts the wavelength scale by about 0.1 nm, which is precisely the threshold uncertainty for ALTIUS.

Table 3.7 – Requirements on three major error sources for the ALTIUS mission. The pointing requirement is for uncertainties along the vertical dimension. All quantities are defined as equivalent to two standard deviation of a Gaussian random perturbation.

Error source	Requirement	Confidence level
Meas. noise (SNR)	Table 3.8	2σ
Pointing	< 0.1 mrad	2σ
Wavelength	< 0.1 nm	2σ

3.6 O_3 retrieval in bright limb

We will now address the measurement strategy of O_3 profiles from limb scattered radiances observed by a spectral imager fulfilling the ALTIUS SNR requirements and possessing its spectral bandwidth. In this first analysis, no pointing or wavelength mis-registration is considered. The first step is to identify the measurement wavelengths and their vertical sensitivity range. Then we will apply the retrieval algorithm presented in section 1.4 to match the simulated measurements to the concentration profiles of the synthetic atmospheric model.

3.6.1 O_3 measurements

The selection of the wavelengths is based on previous limb sounder heritage. Flittner [45] described a simple method applied to one of the first

`Table 3.8` – ALTIUS signal-to-noise ratio requirements expressed as a function of tangent altitude (column-wise, km) and wavelength (row-wise, nm). The values in boldface are the firm requirements to be met by binning the pixels of the same row and adjusting the integration time during the image acquisition if necessary.

	250	**300**	**350**	**400**	**450**	**500**	**550**	**600**	**650**	**700**	**800**
5	1	55	1120	1160	5660	3940	2210	1400	1170	908	831
10	1	55	1130	1160	4790	**3290**	**1800**	**1090**	**913**	707	631
15	1	56	1130	**1150**	**4060**	**2760**	**1470**	**850**	**711**	551	479
20	1	58	1100	**1070**	**3430**	**2310**	**1200**	**662**	**554**	429	365
25	1	59	971	**862**	**2910**	**1940**	**978**	**516**	**431**	334	278
30	1	61	**765**	**640**	**2460**	**1630**	**801**	**402**	**336**	260	212
35	2	**64**	**568**	**462**	**2080**	**1370**	**657**	**313**	**262**	203	162
40	2	**67**	**407**	**325**	**1760**	**1150**	**540**	**244**	**204**	158	124
45	2	**73**	**291**	229	1490	**968**	**445**	**190**	**159**	123	95
50	2	**78**	**208**	161	1260	**814**	**366**	**148**	**124**	96	73
55	2	**65**	**147**	111	1070	**685**	**302**	**115**	**96**	75	56
60	2	**44**	99	72	904	**577**	**250**	**90**	**75**	58	43
65	3	26	62	44	765	486	207	70	58	45	33
70	2	14	36	24	648	410	172	54	45	35	25
75	1	7	18	12	548	345	142	42	35	27	20

limb scattering experiment during the SOLSE/LORE mission [82]. He showed that a few pairs of UV wavelengths (Hartley-Huggins bands) and a triplet of visible wavelengths (Chappuis band) can already enable accurate stratospheric ozone remote sensing. Each pair or triplet combines one strongly absorbing wavelength λ_s with one or two weakly absorbing wavelength(s) λ_w. The advantage is that such combination preserves a high sensitivity to ozone, while removing many interfering features like Rayleigh and aerosol scattering, albedo effects,... This method was also used by von Savigny [113] with OSIRIS spectra (with only a slightly adapted version of the triplet) and further developped by Rault [93] on SAGE III limb measurements using multiple triplets in the visible range. More recently, the same procedure was tested on OSIRIS spectra with six pairs in the UV and a single triplet in the Chappuis band [35].

For the ozone case, we tested the expected performance of the instrument by using only five wavelengths : the first of the six doublets presented in the last-cited paper of the OSIRIS team, and the triplet of Flittner. We also followed the recommendations for the respective altitude ranges and restricted the UV doublet to the higher part of the stratosphere and the VIS triplet to the lower part. Finally, in order to avoid the need for absolute radiometric calibration and remove albedo effects, one pixel row in each image is used to normalize the others below.

A summary of the O_3 measurement parameters is provided in Table 3.9 and the ozone absorption cross-section at room temperature is shown in Fig. 3.6 for illustrative purpose.

Table 3.9 – Parameters used for the O_3 measurements simulations.

	λ (nm)	**Altitude range (km)**	**Norm. altitude (km)**	**Heritage**
Doublet	$\lambda_s = 292$ $\lambda_w = 351$	30–60	65	OSIRIS, [35]
Triplet	$\lambda_{w1} = 525$ $\lambda_s = 600$ $\lambda_{w2} = 675$	10–35	40	SOLSE-LORE, [45]

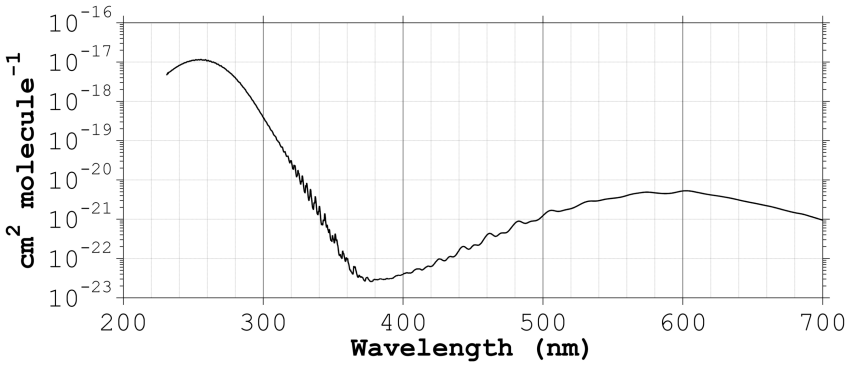

Figure 3.6 – Absorption cross section of O_3 at 293 K [22].

Let us define h_i, the tangent altitude of the ith pixel row, and h_n, the tangent altitude of the normalization row. The signal registered by the pixels of the ith row at a given wavelength is denoted by $I(h_i, \lambda)$. The doublet takes therefore the following form:

$$D(h_i) = \frac{I(h_i, \lambda_s)}{I(h_n, \lambda_s)} \frac{I(h_n, \lambda_w)}{I(h_i, \lambda_w)}, \tag{3.1}$$

while the triplet is given by

$$T(h_i) = \frac{I(h_i, \lambda_s)}{I(h_n, \lambda_s)} . \sqrt{\frac{I(h_n, \lambda_{w1}).I(h_n, \lambda_{w2})}{I(h_i, \lambda_{w1}).I(h_i, \lambda_{w2})}}. \tag{3.2}$$

Taking the geometric mean of the *weak* measurements in the triplet might look risky: if one of the two quantities is close to zero, the ratio will take unreasonably large values. This situation, however, is very

unlikely given the expected typical photon counting rates (see Fig. 3.5). Moreover, a proper data pre-processing should flag unexpectedly small signals.

The justification for using a weighted product of signals instead of a simple sum is found in the spectrum of the limb-scattered radiance [76] (see Fig. 3.5). In absence of molecular absorption, and for an optically thin atmosphere, the limb radiance is exponentially decreasing with the wavelength (this statement is only valid for $\lambda > 500$ nm, i.e. after the maximum of the solar irradiance spectrum): $I(h,\lambda) = A(h).\exp(\alpha\lambda)$. This is best seen in the lowest curves of Fig. 3.5. In an attempt to determining the ozone content, the ideal measurement strategy would consist of comparing signals recorded with and without O_3 for the same light path (i.e. with all the other atmospheric parameters frozen). While this is a standard approach in the laboratory, this cannot be done with the atmosphere... unless the following trick is used.

If one agrees that both *weak* measurements correspond to radiance without ozone interference, then their weighted geometric mean gives access to the pseudo-measurement at the *strong* wavelength in absence of ozone. Using the approximate representation by the exponential, we have

$$I(h,\lambda_s)\,\big|_{\text{no } O_3} = I(h,\lambda_{w1})^a.I(h,\lambda_{w2})^b = A(h)^{a+b}e^{\alpha(a\lambda_{w1}+b\lambda_{w2})}. \quad (3.3)$$

If a and b are chosen such that

$$\begin{aligned} a &= 1 - \frac{|\lambda_{w1}-\lambda_s|}{|\lambda_{w1}-\lambda_s|+|\lambda_{w2}-\lambda_s|}, \\ b &= 1 - \frac{|\lambda_{w2}-\lambda_s|}{|\lambda_{w1}-\lambda_s|+|\lambda_{w2}-\lambda_s|}, \end{aligned}$$

then the right term of eq. (3.3) reduces to $A(h).\exp(\alpha\lambda_s)$, which is precisely what should be the radiance at λ_s without ozone. In our case $a = b = 0.5$ which explains the square root in eq. (3.2). It is worth noting that this reasoning does not hold for weak wavelengths situated on the same side of λ_s. Similarly, the definition of the weighting roots depends on the number of weak wavelengths used.

Assuming that the SNR requirements for the mission are met (Table 3.8), the noise levels $\epsilon(h_i,\lambda)$ at any row of a spectral image acquired at any wavelength can be obtained from interpolation in Table 3.8. The uncertainty on D and T follows from propagating the ϵ values through eq. (3.1) and (3.2).

3.6.2 Measurement error

As the doublet and triplet quantities are nonlinear functions of the measurements, their statistical distribution does not strictly follow any classical function, although each measurement has its statistical occurence well described by a normal distribution. Moreover, the normalization introduces correlation among the vectors elements. Their respective covariance matrices should therefore not be diagonal. The doublet case is tackled below with the classical error propagation method. An equivalent, but more statistically robust method is described in Appendix A.

For the sake of clarity, we use the following notations: $I(h_i, \lambda_s) = I_i^s$, $I(h_i, \lambda_w) = I_i^w$, $I(h_n, \lambda_s) = I_n^s$ and $I(h_n, \lambda_w) = I_n^w$. Eq. (3.1) becomes:

$$D_i = \frac{I_i^s}{I_i^w} . \frac{I_n^w}{I_n^s}. \tag{3.4}$$

Let $\boldsymbol{x}$ be a vector containing all the measurements used for the computation of the doublet. With $i = 1, \ldots, m$, $\boldsymbol{x}$ is a vector of size $[z \times 1]$, with $z = 2m + 2$ (two spectral images within which m rows are used, plus one more row in each image for the normalization). We chose to organize its elements as follows:

$$\boldsymbol{x} = \left(I_1^s \cdots I_m^s \, I_1^w \cdots I_m^w \, I_n^s \, I_n^w\right)^T . \tag{3.5}$$

By linearization of the problem (1st order Taylor expansion), we can write in general:

$$D_i(\boldsymbol{x}) \approx D_i(E(\boldsymbol{x})) + \sum_{j=1}^{z} (x_j - E(x_j)) \left. \frac{\partial D_i}{\partial x_j} \right|_{E(\boldsymbol{x})}, \tag{3.6}$$

with $E(.)$ denoting the mathematical expectation of the parameter. By definition, the covariance between D_k and D_l is given by

$$\mathrm{cov}(D_k, D_l) \equiv E\Big[\big(D_k(\boldsymbol{x}) - D_k(E(\boldsymbol{x}))\big)\big(D_l(\boldsymbol{x}) - D_l(E(\boldsymbol{x}))\big)\Big]. \tag{3.7}$$

Using eq. (3.6), this formula is approximated by

$$\mathrm{cov}(D_k, D_l) \approx \sum_{p=1}^{z} \sum_{q=1}^{z} \left. \frac{\partial D_k}{\partial x_p} \right|_{E(\boldsymbol{x})} \left. \frac{\partial D_l}{\partial x_q} \right|_{E(\boldsymbol{x})} E\Big[\big(x_p - E(x_p)\big) . \big(x_q - E(x_q)\big)\Big], \tag{3.8}$$

where the expectation term is the pq element of the covariance matrix of $\boldsymbol{x}$: $C_{\boldsymbol{x}}$. This is a diagonal matrix as all the measurements are supposed to be independent ($\mathrm{cov}(x_p, x_q) = 0$ for $p \neq q$).

As usual, there is an advantage of using the matrix formalism. If we denote the Jacobian of $\boldsymbol{D(x)}$ by $J_{\boldsymbol{D}}$, such that

$$J_{\boldsymbol{D}} \equiv \begin{pmatrix} \frac{\partial D_1}{\partial x_1} & \cdots & \frac{\partial D_1}{\partial x_z} \\ \vdots & \ddots & \vdots \\ \frac{\partial D_m}{\partial x_1} & \cdots & \frac{\partial D_m}{\partial x_z} \end{pmatrix}, \tag{3.9}$$

then this $[m \times z]$ matrix is upper triangular and sparse as

$$\frac{\partial D_i}{\partial x_j} \neq 0 \qquad \text{only for} \qquad \begin{cases} j = i \\ j = m + i \\ j = 2m + 1 \\ j = 2m + 2 \end{cases} \tag{3.10}$$

In matrix formalism, eq. (3.8) takes the following form:

$$C_{\boldsymbol{D}} = J_{\boldsymbol{D}}.C_{\boldsymbol{x}}.J_{\boldsymbol{D}}^T \tag{3.11}$$

The same reasoning applies to the triplet.

3.6.3 Statement of the inverse problem

We can now address the retrieval method. MODTRAN 5 was used to compute the radiances on its finest spectral grid (1 cm^{-1}) and over a spectral range covering the desired wavelengths. The spatial sampling of the measurements followed the detector pixel densities as presented in Table 3.5. The true ozone profile was picked up from a recent ozone climatology combining MLS and ozonesonde data [83]. It corresponds to the zonal mean taken between 40°N and 50°N in June, which is quite close to the chosen MODTRAN geometrical parameters. MODTRAN was run and produced the radiances $I(h, \lambda)$. Their high spectral resolution structures were smoothed by convolution with gaussian functions whose FWHM are taken from Fig. 3.4 at the measurement wavelengths. Finally, eq. (3.1) and (3.2) were used to compute the doublet and triplet vectors.

Fig. 3.7 shows typical doublet and triplet values as a function of the tangent altitude for the chosen ozone profile. The uncertainty on these data points follows from their covariance matrix: the diagonal elements of $C_{\boldsymbol{D}}$ ($C_{\boldsymbol{T}}$) provide the variance of each doublet (triplet) element: $\sigma^2_{D_i} = \mathrm{cov}(D_i, D_i)$ ($\sigma^2_{T_i} = \mathrm{cov}(T_i, T_i)$). By taking the square root to obtain the

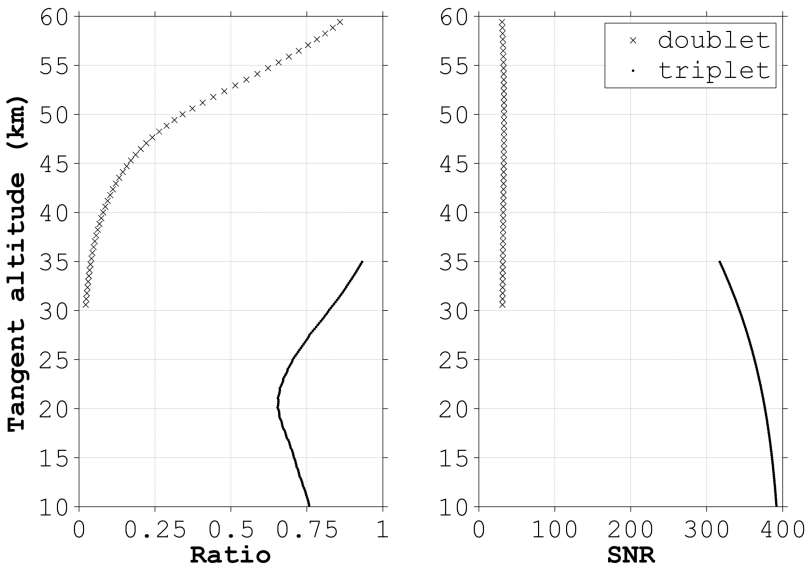

Figure 3.7 – Left: Doublet and triplet values computed from eq. (3.1) and (3.2) in agreement with the O_3 measurement parameters (Table 3.9) for a standard atmosphere. Right: Doublet and triplet SNR obtained from spectral images fulfulling the ALTIUS radiometric noise requirements (Table 3.8).

standard deviation, we can compute the expected SNR: D_i/σ_{D_i} (T_i/σ_{T_i}) (right panel).

Assembled into one vector,

$$\boldsymbol{y} = (D_1 \cdots D_m \, T_1 \cdots T_n)^T \tag{3.12}$$

the data points from the left panel of Fig. 3.7 constitute the effective measurements which will be used by the inversion algorithm to retrieve the ozone profile. Our retrieval scheme is based on finding a solution $\boldsymbol{x}$ which maximizes the conditional probability $P(\boldsymbol{x}|\boldsymbol{y})$ (the MAP solution). The general algebra has been described in section 1.4 and we will only discuss some particular aspect with a quantitative approach now.

As mentioned earlier, the *true* ozone profile comes from a recent climatology built from ozonesondes and satellite data. The retrieval scheme makes use of an assumption on the profile: the *a priori* ozone profile $\boldsymbol{x_a}$. It was picked up in the COSPAR International Reference Atmosphere (CIRA-86, [64]) as the mean mid-latitude ozone profile. In order not to skew the inversion, the uncertainty on $\boldsymbol{x_a}$ is chosen very large: 100% for each point. As we can presumably assume that the

ozone concentration does not vary too wildly between adjacent points, we populated the off-diagonal elements of the a priori vector covariance matrix ($C_{\boldsymbol{x_a}}$) by assuming a correlation length (L) of 1.25 km. In this work, we define the correlation length as the half-width at half-maximum (HWHM) of a Gaussian function connecting the uncertainty of neighbouring layers. This is maybe best understood with the following familiar expression:

$$\mathrm{cov}(x_{ai}, x_{aj}) = \sigma_{ai}.\sigma_{aj} \, . \, e^{-\left(\frac{z_i - z_j}{L}\right)^2 . \ln 2}, \tag{3.13}$$

where σ_{ai} is the uncertainty on the *a priori* knowledge of the ozone concentration in the ith layer, and z_i is the mean altitude of this layer. The measurement vector covariance matrix ($C_{\boldsymbol{y}}$) is built by assembling the blocks $C_{\boldsymbol{D}}$ and $C_{\boldsymbol{T}}$ along the diagonal:

$$C_{\boldsymbol{y}} = \begin{pmatrix} C_{\boldsymbol{D}} & 0 \\ 0 & C_{\boldsymbol{T}} \end{pmatrix}. \tag{3.14}$$

As the solver needs a linearization of the physical model, the Jacobian of the forward model (K) must be evaluated. To reduce numerical instabilities, we will work with the natural logarithm of the O_3 concentration:

$$K_{ij} = \frac{\partial y_i}{\partial \ln x_j}, \tag{3.15}$$

where y_i is the ith element of the measurement vector (eq. (3.12)) with $i = 1, \ldots, m+n$, and x_j is the ozone concentration in the jth atmospheric layer, $j = a, \ldots, b$, a and b denoting the lowest and highest layers belonging to the target altitude range (Table 3.9). The K_{ij} were computed by discrete differentiation: the ozone content in each layer was reduced by 5% by turns, and the change in the measurement vector elements was estimated. The rows of K are peaked functions of altitude whenever there exists sensitivity in the measurements to the concentration of a species within a given layer. They are usually called the weighting functions.

Fig. 3.8 is a plot of these weighting functions. Their density is directly related to the spatial sampling of the instrument: each homogeneous layer of the atmospheric model grid is probed by several detector rows (the sampling is three times larger in the VIS channel than in the UV channel). The limb scattering geometry is responsible for the shape of the weighting functions: light collected by a given detector row has propagated through the upper layers of the atmosphere before being scattered towards the instrument. Neglecting multiple scattering, a

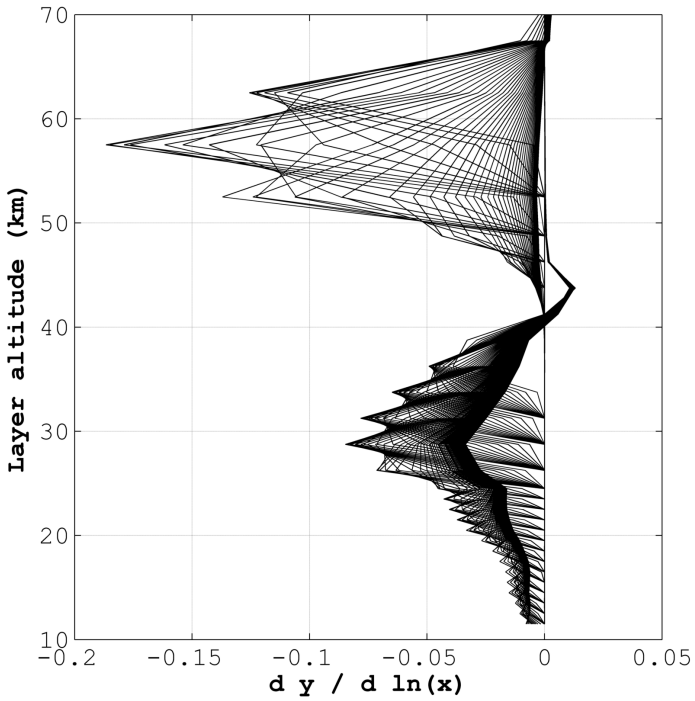

Figure 3.8 – Weighting functions representing the response of the measurement vector to a 5% perturbation of the ozone profile.

measurement performed at a given tangent altitude is mainly sensitive to the scattering layer, and its closest upper neighbours. This approximation is valid down to about 20 km, depending on the wavelength (multiple scattering is more likely to happen at shorter wavelengths). The scale of the horizontal axis is not really important. What matters is how sharp are the peaks, to which altitude do they point and are there regions of lower sensitivity?

Clearly, there are two regions in this figure. In the UT/LS, it is mainly the triplet which is at work: the strongly absorbed UV wavelength (292 nm) is completely removed from the light spectrum in this region because of a two-orders of magnitude larger ozone cross-section than in the visible domain. Only the triplet wavelengths are allowed to travel down to the stratosphere as they get depleted by the Chappuis band.

It seems however that the choice of wavelength is not optimal. There is a gap of sensitivity around 45 km. This gap could be filled by adding one more wavelength (or shifting up the one at 292 nm) around 300-310

nm, a part of the spectrum which is allowed to go a little deeper into the stratosphere. While this will be foreseen for the real operations, we chose not to change the wavelengths as it allows to better emphasize the respective roles of the UV and visible wavelengths.

3.6.4 Retrieved profile

Having described all the quantities playing a role in the inversion algorithm, we can address an example of retrieved profile and discuss its error bars. Usually, one retrieves ozone in a few iterative steps: after the first inversion, most of the profile is already well captured but the bias is still large in the lowest layers. The first solution is then used as a new linearization point and the weighting functions are recomputed around this new atmospheric state. After three inversions, any other step does not improve the solution any further.

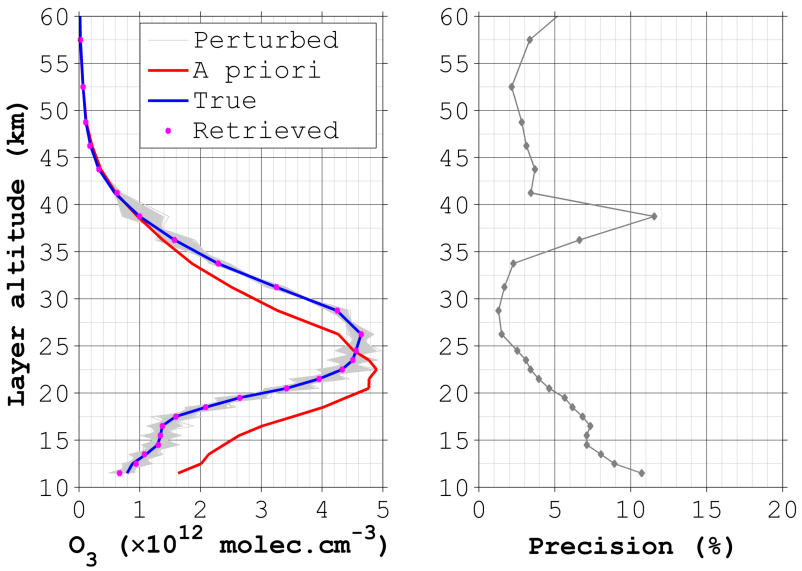

Figure 3.9 – Left: Example of retrieved O_3 profile on the US standard grid using 5 spectral images taken at the wavelengths defined in Table 3.9 (dots). The initial guess (red curve) and the true ozone profile (blue curve) are shown. The gray curves are the profiles retrieved from measurements contaminated by noise. Right: The gray curve indicates the relative precision of the retrieved profile as measured by the standard deviation of the noisy retrievals (1σ confidence level).

In Fig. 3.9, the true ozone profile, the *a priori* and the retrieved solu-

tion are plotted (left panel). As can be seen, the amount of information lying in the spectral images is sufficient to accurately converge towards the true atmospheric state (blue curve), starting from a seriously biased initial guess (red curve).

The uncertainty on the retrieved solution has been assessed by adding noise on the measurements before repeating the inversion. One thousand measurement vectors were built by perturbing each element of eq. (3.12) with an additive or subtractive contribution taken in the limits of a normal distribution of zero-mean and of variance given by the $C_{\boldsymbol{y}_{ii}}$ elements ($C_{\boldsymbol{y}_{ii}} = \sigma_{\boldsymbol{y}_i}$). The explicit perturbation formula is the following:

$$\boldsymbol{y}^* = \boldsymbol{y}.\left(1 + \sigma_{\boldsymbol{y}}\,.\mathcal{N}(0, \sigma_{\boldsymbol{y}}^2)\right). \tag{3.16}$$

The * indicates a perturbed measurement vector and the standard notation convention is adopted for specifying the normal distribution of given mean and variance ($\mathcal{N}(\mu, \sigma^2)$).

The random noise in the measurements introduces local biases in the retrieved profile. By simulating a large number of repeated measurements, one ends up with a statistically significant group of inverted profiles. Computing the standard deviation in each layer and comparing it to the noise-free case, we obtain an estimate of the retrieval relative uncertainty (precision). This is illustrated in the right panel of Fig. 3.9. As can be seen, the SNR requirements combined with a relevant choice of wavelengths ensure uncertainties lower than 10% from the high troposphere to the lower mesosphere. Smaller values than 5% can even be expected above 20 km (assuming a slightly larger value of λ_s in the UV doublet is used to remove the loss of sensitivity around 40 km which was already noticed in the weighting functions).

3.6.5 Vertical resolution

The sensitivity of the retrieved profile to the atmospheric state is captured in the rows of the averaging kernel matrix (eq. 1.31). As for the weighting functions, a good sensitivity of the solution to a particular atmospheric layer corresponds to a sharp peak pointing at this layer. Here, values are important: the closer to 1, the more information comes from the actual atmospheric state. Conversely, values nearing zero indicate that most of the information comes from the a priori, or that the layer simply does not contribute to the inverse problem.

The averaging kernels for our ozone retrieval have been plotted in function of the layer altitude in Fig. 3.10 (left panel). The good sensitivity is confirmed by the sharp peaks, except around 40 km. Again,

this is due to the choice of wavelengths and can easily be fine tuned. Each averaging kernel is identified by a unique color and associated to a layer altitude in the adjacent colorbar.

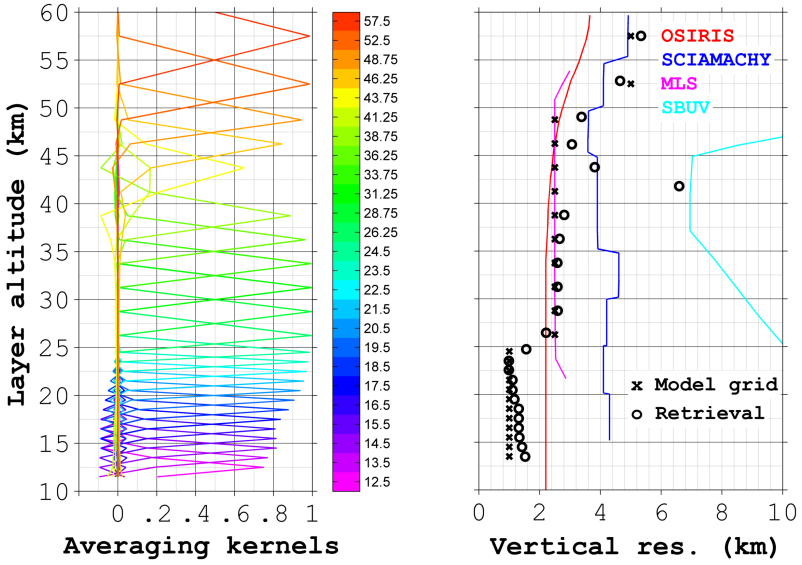

Figure 3.10 – Left: Rows of the averaging kernel matrix. Each curve represents the sensitivity of the retrieved O_3 concentration within a given layer (denoted by its mean altitude and associated to a unique color in the adjacent colorbar) to the entire atmosphere. Right: Comparison between the atmospheric model vertical sampling (crosses) and the retrieval spatial resolution (circles). For comparison, the vertical resolution of four other instruments is indicated.

Perhaps an example allows to better understand this figure. Let us consider the cyan curve which is related to the atmospheric layer located at 22.5 km (air found between 22 and 23 km). This curve is peaking sharply at 22.5 km and is flat everywhere else. This means that the measurement strategy combined to the retrieval algorithm are such that the obtained value for the O_3 concentration in this layer is in fact very sensitive to the actual amount of O_3 at that altitude, which is of course ideal for a profiling instrument. On the other hand, the green curve associated to the retrieval in the layer located at an altitude of 41.25 km shows a broader peak at a slightly higher altitude. It indicates that most of the information used to derive the amount of O_3 in the layer spanning 40-42.5 km is not only coming from that layer, but also from the adjacent ones, and presumably slightly more from the upper ones.

The right panel of Fig. 3.10 compares the width of the layers used in the radiative transfer model (US standard grid) to the FWHM of the averaging kernels. This is a widely adopted means for estimating the spatial resolution of the retrieved profile. As illustrated by the spread occuring around 40 km, a loss of resolution is translated into a broad function of reduced amplitude. Fine structures (< 5 km) in this region are therefore out of reach with the current measurement parameters. For the rest of the profile, the achieved resolution is comparable to the model grid. Higher resolution may even be reached between 20 and 40 km if a finer atmospheric grid was used.

The vertical resolution of other instruments has been added in the right panel of Fig. 3.10: two UV-VIS limb-scatter instruments (OSIRIS and SCIAMACHY), one limb emission sounder (Aura-MLS) and one nadir-looking instrument (SBUV) (data source: [57]). For most of the stratosphere, ALTIUS will show better or equivalent performance than the state-of-the-art. With further developments involving a finer model grid resolution, it is very likely that the ALTIUS stratospheric ozone product will match the vertical resolution of occultation instruments (1 km), but with a much more extended geographical coverage.

3.6.6 Pointing uncertainty

It has been emphasized in the introduction of this chapter that the mis-registration of the LOS tangent point altitude can be the driver of the total error budget. This was known since the early times of the limb scattering instruments and is very well documented from the point of view of the post-processing data analysis [114], [81]. Workarounds rely on the fitting of measured radiances with either ozone features in the deep UV or Rayleigh scattering in the near UV [60] (see section 1.3.7). However, pre-launch error budgets have often proven to be too optimistic by considering only small pointing uncertainties and often absolutely no jitter [76].

ALTIUS being an imager, one of its advantages is that it removes completely the uncertainty on the relative distance between the tangent points. Contrary to instruments equipped with scanning systems, the assumption that each pixel LOS is affected by the same pointing bias is a sound hypothesis. We will therefore assess the impact of the platform attitude uncertainty on the instrument performance by shifting the entire images upwards and downwards.

The pointing uncertainty is specified in Table 3.7 as an error smaller than 0.1 mrad along the vertical direction at a 2σ confidence level.

Translated into the more standard 1σ level, and at the expected spacecraft altitude (650 km), this corresponds to drifts of the tangent point location by $\Delta z = \pm 150$ m. In order to simulate such a pointing mis-registration, we computed new measurement vectors from spectral images showing a shifted scene.

A Monte Carlo simulation of the jitter is out of reach. Assuming at least 100 pointing offsets taken in a Normal distribution $\mathcal{N}(0, 150^2)$ for each of the 5 images leads to 10^7 spectral images to build. Knowing that each image consists of tens of tangent altitudes, we would end up with about 10^9 calls to the radiative transfer model. The privileged approach was then to work out the 32 cases illustrating a $\pm 1\sigma$ bias for each image[1]. Then, the 32 measurement vectors of mis-registered data were inverted in one step by using the weighting functions of the final iteration of the error-free retrieval.

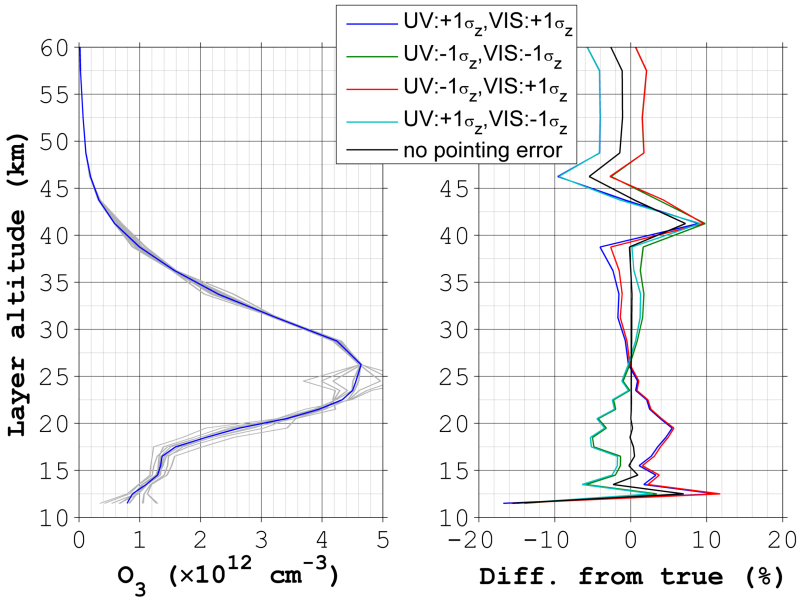

Figure 3.11 – Left: The gray curves are the retrieved ozone profiles from the 32 bad pointing scenarii. The blue curve is the true ozone profile. Right: The relative difference from the true profile is computed for 4 particular situations wherein the images taken by the same channel have encountered the same pointing offset. The legend provides description of the situations. The black curve is for the error-free measurements.

Fig. 3.11 shows how the 32 pointing perturbation scenarii behave

[1]With five independent spectral images, there are 2^5 different pointing offset schemes as each image can suffer from an upward or a downward bias

compared to the true profile (left panel). Some situations lead to extremely large errors at the altitude of the peaking ozone concentration. Of course, at this altitude, images taken at visible wavelengths are to blame. Analysis of the data shows that the worst case scenarii are met when the pointing offset goes in opposite directions among the triplet images. The case is summarized in Table 3.10. However, such an unfavorable situation is very unlikely: over timescales of a few seconds (the typical time for an observation), the demonstrated performance in terms of pointing jitter for the PROBA platform is about 70 m (2σ) at the tangent point [31], and it is a high frequency effect which will average out during the image exposure time.

`Table 3.10` – Worst case situation for pointing errors on the triplet images inducing large bias at the peaking O_3 altitude.

Point. Err. $\boldsymbol{I}(\lambda_s)$	**Point. Err.** $\boldsymbol{I}(\lambda_{w1})$ **and** $\boldsymbol{I}(\lambda_{w2})$	**Ret. Err. at 25km**
-150 m	+150 m	-20%
+150 m	-150 m	+20%

The right panel of Fig. 3.11 is a focus on the relative difference of four situations which have more chances to occur (same bias among images acquired in the same channel). Two of them (blue and green curves) illustrate a general pointing bias affecting both UV and VIS channels. The two others (red and cyan curves) show what to expect in case of inter-channel mis-alignement. A carefull look at the behaviour of these curves reveals without surprise that test cases sharing the same bias for the visible wavelengths travel together up to the upper stratosphere. After the region of lower sensitivity (40 km), they split and form new pairs linking test cases showing identical bias among their UV images. One can also notice that the sign of the relative difference is consistent with the ozone profile gradient: below the peaking altitude, an upward pointing bias (in the VIS channel) sees higher O_3 content than it should (the relative difference is computed as $(\boldsymbol{x}^{\text{ret}} - \boldsymbol{x}^{\text{true}})/\boldsymbol{x}^{\text{true}} \times 100$). Above the ozone layer, the same upward bias yields lower concentrations.

3.6.7 Spectral uncertainty

For an atmospheric limb remote sensing instrument, an error on the knowledge of the measurement wavelength can introduce a bias in the retrieved profile: in the process of matching the measured radiance with the simulated one (and hence with the atmospheric state), the forward

model will make a mistake on the actually encountered absorption cross-section which will end up with an error in the profile estimation.

In ALTIUS, the central wavelength of the AOTF passband is determined by the RF signal applied to the transducer. As the relationship follows approximately a F^{-1} law, steps in frequency corresponding to 0.1 nm (the 2σ spectral requirement) vary from 5 to 50 kHz from one side to the other of the spectral domain (see Chapter 4). The onboard signal generator should be capable of delivering this level of accuracy. The major issue is actually the crystal temperature. An uncertainty of 1°C easily leads to a 0.1 nm error. This is due to the thermal dependence of the refractive indices and elastic properties. A reliable temperature sensor and a good AOTF thermal model will be mandatory to meet the spectral requirements.

In order to simulate the impact of the wavelength mis-registration, the measurement vector was computed for slightly different doublet and triplet wavelengths. A value of 0.05 nm corresponding to the 1σ uncertainty requirement was used (Table 3.7). It is likely that the temperature error will consist of a bias equally affecting all the wavelengths selected by a given AOTF. Therefore, in one situation, 0.05 nm were added, while in another, 0.05 nm were subtracted from the target wavelengths. The weighting functions remained unchanged.

We emphasize in Fig. 3.12 the same combination of offsets as found in the right panel of Fig. 3.11. Here, the smoothness and the flatness of the ozone absorption cross-section around the triplet wavelengths remove completely any sensitivity of the retrieval to small wavelength mis-registrations over most of the stratosphere. Ignoring the spikes at 40-45 km (which have been discussed above), the spectral error starts to be non-negligible around the stratopause (still below 1% though). This is well explained by the position of the strong doublet wavelength: right in the steep slope of the O_3 Hartley band (see Fig. 3.6). Any wavelength error yields a wrong assessment of the ozone specific optical thickness and hence, of its concentration.

3.6.8 Final O_3 retrieval performance budget

The payload and spacecraft requirements have been used to derive the precision (random error) and the accuracy (systematic error) with which O_3 profiles can be retrieved, based on the information carried by five bright limb spectral images recorded by ALTIUS (two in the UV channel and three in the VIS channel). The data have been simulated with a multiple scattering radiative transfer code (MODTRAN 5), and con-

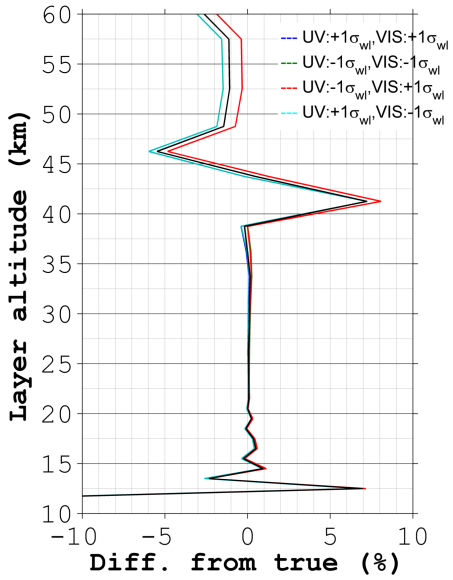

Figure 3.12 – Colored lines are retrieved ozone profiles based on measurements suffering from wavelength mis-registration. The spectral bias is the same for the images acquired by the same channel, but affect both channels independtly.

volved to the expected instrument spectral resolution. The measurement noise was provided by the SNR requirements.

The inversion algorithm followed a Bayesian approach, and the retrieved ozone profiles ($\boldsymbol{x}^{\text{ret}}$) are those maximizing the conditional probability $\text{P}(\boldsymbol{x}|\boldsymbol{y})$, with $\boldsymbol{y}$ being the measurement vector. The retrieval precision was assessed by a Monte Carlo method. Impacts on retrieval accuracy were determined from pointing and spectral biasses.

Table 3.11 summarizes the retrieval errors for selected altitudes and at a 1σ confidence level. Numbers at 45 km were estimated as if the choice of wavelengths was slightly more optimal: by averaging errors found below and above this region of lower sensitivity. The vertical resolution of the retrieved profile (understood as the FWHM of the averaging kernels) is also indicated.

Similar studies exist for OSIRIS [54] and SCIAMACHY [92], such that the ALTIUS figures can be critically reviewed against proven instruments. Obviously, their work is based on the actual performance of the instruments (whereas the above simulations make use of the requirements, i.e. the projected performance). Table 3.12 provides error

Table 3.11 – Final error budget for the O_3 retrieval by ALTIUS in limb scattering geometry. The total error is the quadratic sum of the three error sources.

		Layer altitude (km)								
Error type	**Perf.**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**
	Res. (km)	1	1	1	1.5	2	2.5	2.5	3	3.5
precision	SNR (%)	7	5	2	1.5	2.5	3	3.5	2.5	2.5
accuracy	Pointing (%)	2	4	1	1	2	2.5	2.5	3	3
accuracy	Spectral (%)	0	0	0	0	0	0.2	0.4	0.5	0.6
	Total (%)	7.3	6.4	2.2	1.8	3.2	3.9	4.3	3.9	4.0

estimates (1σ) for the pointing and SNR performance. We took care of using values for mid-latitude measurements. In general, ALTIUS is comparing very well with respect to its predecessors, especially regarding precision.

Table 3.12 – Error contributions to the total error budget of ozone profile measurements by OSIRIS and SCIAMACHY.

	Layer altitude (km)								
	15	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**
OSIRIS									
Res. (km)	2.2	2.2	2.2	2.2	2.2	2.3	2.4	2.8	3.3
SNR (%)	5	4	4	6	10	13	13		
Pointing (%)	7	4	1.5	5	8	9	5.5		
SCIAMACHY									
Res. (km)	4.5	4	4	4.3	4.5	4	4	3.7	4
SNR (%)	18	16	13	15	15	12	11	10	11
Pointing (%)	4	3.5	1	1.5	1.5	2	3.5	4	5

3.7 NO_2 retrieval in bright limb

3.7.1 NO_2 measurements

Nitrogen dioxyde exhibits a highly structured absorption spectrum at short visible wavelengths (Fig. 3.13). Instruments with moderate resolution (< 3 nm) can resolve most of the large oscillations found around 440 nm. This makes the measurement approach based on wavelengths ratio as attractive as for ozone. However, there is one more argument in this case: the doublet will be made of very close wavelengths, reducing further the interference by weakly wavelength-dependent processes such

as Rayleigh scattering.

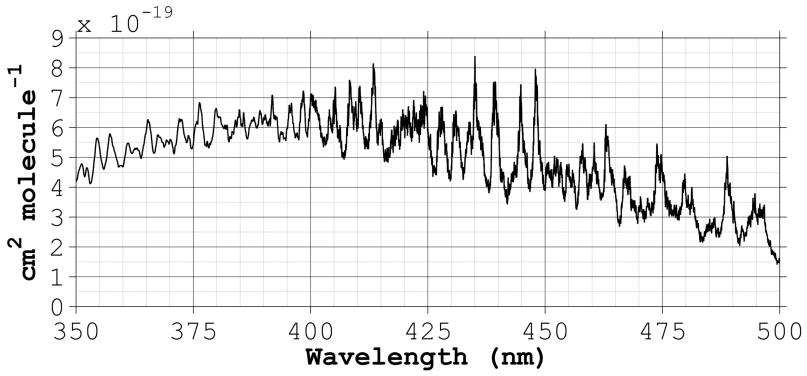

Figure 3.13 – Absorption cross section of NO_2 at 294 K [107].

Starting in the 80's, SAGE II demonstrated NO_2 retrievals [28] in solar occultation with two wavelengths: at 448 and 453 nm. Launched in 1998, POAM III used a slightly different doublet: 439.6 and 442.2 nm [78]. More recently, a similar approach was attempted with the limb scattering instrument OSIRIS [12], based on a quadruplet of wavelengths: a strong one at 448.23 nm, and three weak ones at 447.04, 449.81 and 450.21 nm.

For these simulations, we selected the POAM doublet because it is located in a region of fairly constant ozone absorption, and close to the global maximum of the NO_2 absorption cross section. The ozone interference can be problematic in real conditions, when both molecules need to be retrieved simultaneously. Minimizing the effects of ozone absorption at the NO_2 measurement wavelengths is therefore a good practice. The vertical region of interest covers almost the entire stratosphere, as the peak of nitrogen dioxyde concentration is usually found a few km higher than the ozone maximum. The measurement parameters for this species are summarized in Table 3.13.

	λ (nm)	**Altitude range (km)**	**Norm. altitude (km)**	**Heritage**
Doublet	$\lambda_s = 439.6$ $\lambda_w = 442.2$	10–45	50	POAM III, [78]

Table 3.13 – Parameters used for the NO_2 measurements simulations.

3.7.2 Statement of the inverse problem

MODTRAN 5 computed the scattered radiance at the tangent points seen by the VIS channel pixels and at the desired wavelengths. The reference NO_2 profiles were taken from the global assimilation model BASCOE [42]. The true state consists of a NO_2 profile located above 45°N 20°W, on the 15th of June 2011 at 12AM (less favourable conditions as NO_2 content is depleted by photodissociation under sunlight). The a priori was taken at the same point and time but one year earlier. The division of the atmosphere into homogeneous layers follows the US standard grid.

The radiances were convolved at the expected instrument spectral resolution, i.e. 2.4 nm at 440 nm (Fig. 3.4). Then, the doublet was computed according to eq. (3.1) at each probed altitude (Fig. 3.14). Clearly, the contrast proposed by the two adjacent wavelengths is not as large as what can be obtained with ozone. Its associated SNR was derived by propagating the ALTIUS spectral images noise requirements with eq. (3.11). The NO_2 doublet SNR is plotted in the right panel.

Figure 3.14 – Left: Doublet computed from eq. (3.1) in agreement with the NO_2 measurement parameters (Table 3.13) for a standard atmosphere. Right: Doublet SNR obtained from spectral images fulfulling the ALTIUS radiometric noise requirements (Table 3.8).

The NO_2 inversion algorithm strictly follows the same formalism as

for ozone. The only difference is that the measurement vector is now consisting of the doublet values only:

$$\boldsymbol{y} = (D_1 \cdots D_m)^T . \tag{3.17}$$

The same for its error covariance matrix: $C_{\boldsymbol{y}} = C_{\boldsymbol{D}}$.

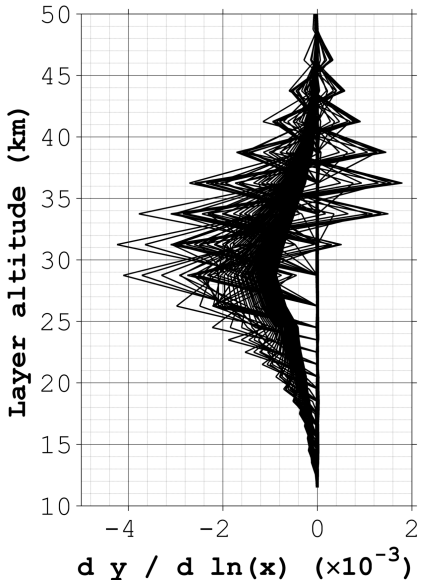

Figure 3.15 – Weighting functions representing the response of the measurement vector to a 5% perturbation of the NO_2 profile.

3.7.3 Retrieved profile

As for ozone, three steps were needed to converge towards the MAP solution, each of them involving new weighting functions calculation. Those of the final steps are shown in Fig. 3.15. They indicate a good response of the measurements over the entire stratosphere. The impact of the radiometric noise was assessed by inverting 1000 perturbated measurement vectors. As for ozone, the perturbation consisted of changing each element of $\boldsymbol{y}$ by a quantity taken in a $\mathcal{N}(0, \sigma_{\boldsymbol{y}}^2)$ distribution.

The left panel of Fig. 3.16 shows the *a priori* profile (red), the true NO_2 profile (blue) and the noise-free retrieved one (magenta). The grey curves are the solutions to the inverse problem for the 1000 noisy measurement vectors. The right panel gives the expected precision in function of altitude. It corresponds to the standard deviation of the population of retrieved profiles.

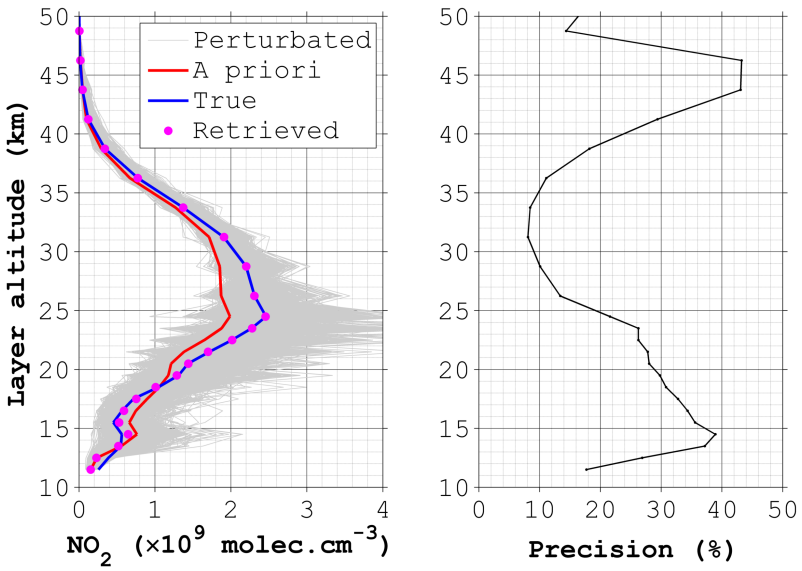

Figure 3.16 – Left: Retrieved NO_2 profile (dots) from two spectral images matching the ALTIUS SNR requirements. The gray lines are the solutions obtained for 1000 noisy measurements. Right: Relative precision of the retrieved profile.

As can be seen, the true profile is well captured by the noise-free inversion. This means that there is enough information contained in the measurements. However, despite large SNR requirements, the measurement noise causes significant uncertainties on the retrieved profile. This is due to several reasons:

- as illustrated by the amplitude of their respective weighting functions, the ozone signature is much larger than NO_2 in the stratosphere: at their peak altitude, their respective extinction coefficient (the product of the absorption cross-section with the concentration) is in a proportion of 20:1 in favour of O_3,
- the choice of wavelengths could be optimized by taking them a little longer, this would allow for an easier penetration into the lowest layers of the stratosphere,
- the 2.5 nm resolution at 440 nm is already sensibly smoothing the structures in the NO_2 spectrum, reducing the contrast between the doublet wavelengths.

If better precision would be requested, we should probably consider

the retrieval of daily zonal means of NO_2 profiles (averaged along latitude circles). The Sun-synchronous orbit of ALTIUS offers about 14 observations per day at almost any latitude and constant local time. The validity of this approach has already been demonstrated with OSIRIS products ([19]).

The more difficult detection at the lowest altitudes is confirmed by the averaging kernels (Fig. 3.17). The retrieval algorithm shows increasing sensitivity up to 25 km where it remains optimal until 40 km. In the right panel, their width is plotted as a function of their peaking altitude and compared to the grid model vertical resolution.

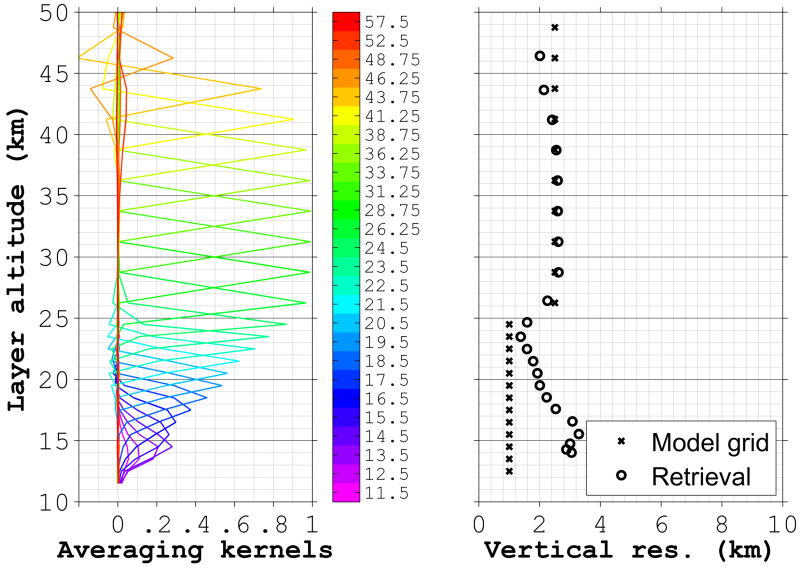

Figure 3.17 – Left: Averaging kernels of the NO_2 inversion algorithm. Right: Vertical resolution of the retrieved profile.

3.7.4 Pointing uncertainty

As both measurement wavelengths belong to the VIS channel, the assessment of the pointing requirement on the retrieval performance was done by shifting together both images up and down. The pointing misregistration was taken as the 1σ requirement: 150 m. The retrieved profiles and the induced bias are plotted in Fig. 3.18. The effect is quite small in the target range between 20 and 40 km: a few percents at maximum, and a symmetric response in function of the sign of the pointing offset.

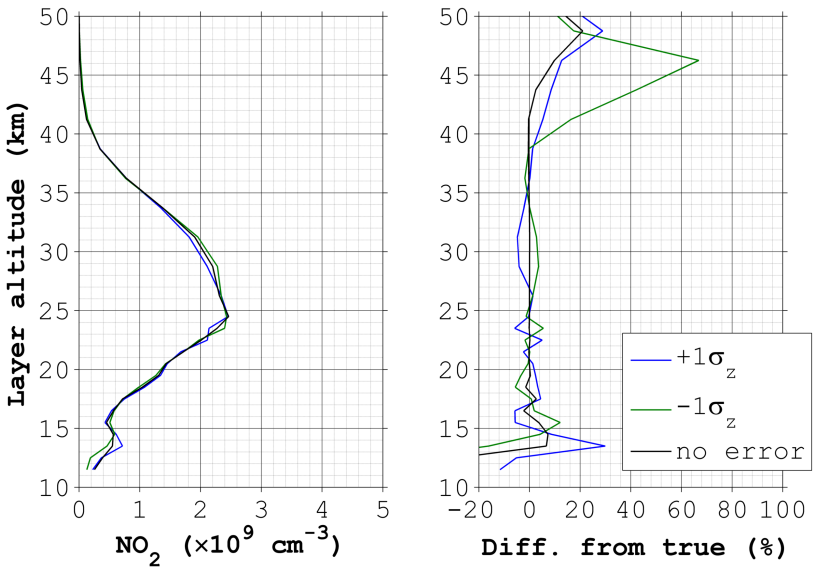

Figure 3.18 – Left: Retrieved profiles for measurements corrupted by a tangent height registration error of +150 m and −150 m (blue and green curves respectively). The black curve is the true NO_2 profile. Right: Relative difference with the true profile. Here the black curve shows the error-free case.

3.7.5 Spectral uncertainty

The stringent spectral requirement was also tested in a similar way as for the pointing. We considered that the error affected both images in the same way. As the effect for the 1σ (0.05 nm) offset was too small, we assessed the 2σ (0.1 nm) situation as well. Both cases are plotted in Fig. 3.19. Again, over the target spatial range, little impact is expected if the spectral uncertainty requirement is fulfilled. The bias however responds linearly with the doubling of the error.

3.7.6 Final NO_2 retrieval performance budget

The results obtained by the propagation of different error sources through the retrieval simulations are gathered in Table 3.14. They represent the 1σ confidence level. The vertical resolution is also indicated.

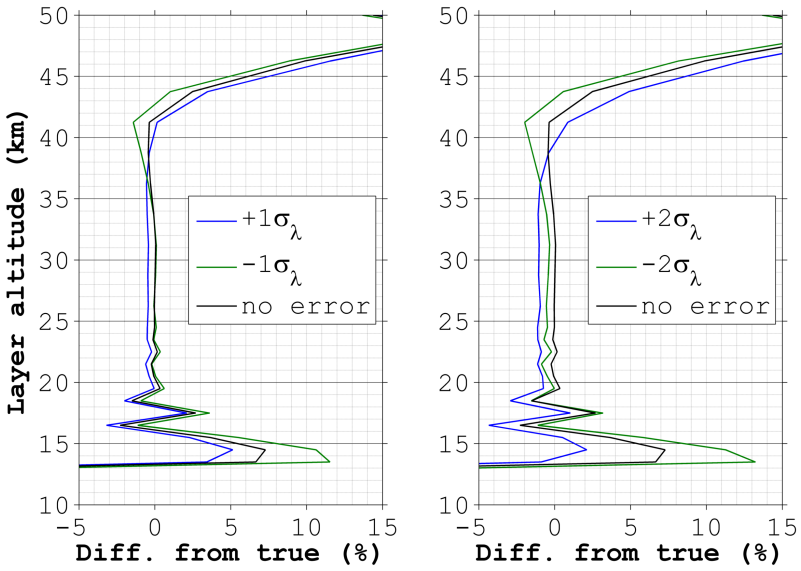

Figure 3.19 – Left: Impact of a wavelength misregistration of ±0.05 nm on the NO_2 retrieval accuracy. Right: The same for a ±0.1 nm error.

Table 3.14 – Final error budget for the NO_2 retrieval by ALTIUS in limb scattering geometry. The total error is the quadratic sum of the three error sources.

		Layer altitude (km)						
Error type	**Perf.**	**15**	**20**	**25**	**30**	**35**	**40**	**45**
	Res. (km)	2	1.4	1.6	2.5	2.5	2.5	2.3
precision	SNR (%)	37	29	19	9	10	24	43
accuracy	Pointing (%)	4	3	0	4	0	6	22
accuracy	Spectral (%)	2	0.2	0.2	0.2	0.2	0.5	1
	Total (%)	37	29	19	10	10	25	48

Chapter 4

Acousto-optical tunable filters

4.1 Foundations

In 1914, Léon Brillouin predicted the scattering of light by acoustic waves in his first scientific communication after graduating from the Ecole Normale Supérieure in Paris. Eight years later, he published a second theoretical paper in this field [18], but experimental confirmation only occured in 1932 with the parallel works of Debye and Sears [34], and Lucas and Biquard [77]. Since then, acousto-optic (AO) interaction became a tool for measuring several parameters of various media: elastic and photoelastic constants, visualization of acoustic fields, sound attenuation, etc...

However, until the appearance of lasers in 1961, no significant application of light-sound interaction emerged. With the appearance of coherent light sources, increasing efforts in transducer technology, and better characterization of crystal acoustic properties, essentially three families of AO devices rose: modulators, deflectors and tunable filters.

Modulators are used to control the intensity of optical beams, usually with periodic pulses. They can be used in telecommunication systems, as Q-switches, etc... On the other hand, deflectors take a monochromatic light source and steer it over a wide angular range by sweeping the acoustic frequency. They can be used as optical gates, in displays, etc... These two types of AO devices were mainly developed in parallel in the 1960's as they rely on similar configurations (laser source and large angular spread of the acoustic beam).

The tunable filters came a little later [56] with completely different applications and principles: from an incident collimated beam of white light, only a fraction of its spectrum is scattered into another propagation direction, depending on the applied acoustic frequency. The goal

was to reach a good spatial separation together with a narrow bandwidth and as many resolvable spots as possible. Further improvements were obtained by taking the full potential of birefringent crystals [24], making devices with large angular aperture possible.

Since then, most efforts have been spent on improving the AO device performances (with materials of higher figure of merit), or seeking for new applications. In the following sections we will not discuss the deflectors and the modulators any further. Instead, we will focus on the tunable filters for their evident suitability in spectral imaging applications.

4.2 Theoretical background

There are several excellent textbooks describing optical and acoustic wave propagation in solids [120], [122], [5]. As acousto-optical tunable filters (AOTF) rely on light and sound coupling inside crystals, a summary of the underlying physics is useful for understanding their parameters and limitations. Moreover, as it will be described later, some laboratory experiments led to re-considering fundamental aspects of the filter (mainly thermal effects). A precise description of the involved physical processes is therefore mandatory for a comprehensive analysis.

In this section, we will first focus on the description of optical and acoustic waves propagating inside crystals, with some emphasis put on birefringent media and shear acoustic deformations as they are mandatory for achieving the best AOTF performances. Then we will present the light-sound coupled equations and study their solutions to see which are the important parameters for a good spectral imager. Finally, we will derive most of the AOTF characteristics by describing the AO effect as a Bragg diffraction process involving photons, phonons and momentum conservation.

4.2.1 Optical waves in birefringent crystals

The constitutive equation of a dielectric medium relates the electric displacement vector $\mathbf{D}$ to the electric field $\mathbf{E}$ of an optical wave through the relative permittivity tensor $\boldsymbol{\epsilon}$ (relative to the free space permittivity constant ϵ_0):

$$\mathbf{D} = \epsilon_0 \boldsymbol{\epsilon} \mathbf{E}. \tag{4.1}$$

The permittivity measures how the electric field induces local electric dipole moments in the medium. As in general the medium does not

behave identically in all directions, this quantity is a rank-2 3×3 tensor. Along the principal axes of crystals with orthogonal symmetry (cubic, tetragonal, hexagonal, trigonal and orthorhombic classes), it forms a diagonal matrix. In this frame, the electrical energy density stored in the crystal is given by:

$$U = \frac{1}{2}\mathbf{E}.\mathbf{D} = \frac{1}{2\epsilon_0}\left(\frac{D_x^2}{\epsilon_x} + \frac{D_y^2}{\epsilon_y} + \frac{D_z^2}{\epsilon_z}\right), \tag{4.2}$$

where eq. (4.1) has been used. We immediately see that the components of the electrical displacement vector form the principal axes of an ellipsoid at each point of constant electrical energy density.

Assuming a classical representation of the electromagnetic (EM) wave of the form

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{A}(\mathbf{r})e^{i(\omega t - \mathbf{k}.\mathbf{r})}, \tag{4.3}$$

with $\mathbf{A}$ taking the role of any of the time- and space-dependent Maxwell quantities, one can derive an expression connecting the relative permittivity of the medium with the refractive index experienced by the optical wave. This is done by combining two of the Maxwell's equations, $\dot{\mathbf{D}} = \nabla \times \mathbf{H}$ and $-\mu_0\dot{\mathbf{H}} = \nabla \times \mathbf{E}$, with eq. (4.1) in order to obtain:

$$\boldsymbol{\epsilon}\mathbf{E} = n^2\mathbf{E} - (\mathbf{n}.\mathbf{E})\mathbf{n}. \tag{4.4}$$

Here, $\mathbf{n}$ is the index vector. It indicates the propagation direction and its modulus gives the effective refractive index: $n = c/v$. Eq. (4.4) is an eigenvalue problem. Choosing a frame in which $\boldsymbol{\epsilon}$ is diagonal (i.e. the crystallographic axes), one can write for each of the vector components:

$$k^2\frac{D_j}{\epsilon_j} - \frac{\omega^2}{c^2}D_j = k_j\epsilon_0\,(\mathbf{k}.\mathbf{E}) \tag{4.5}$$

with j denoting the principal axes x, y and z. Since $\nabla.\mathbf{D} = -i\mathbf{k}.\mathbf{D} = 0$, and $c^2 = 1/(\epsilon_0\mu_0)$, we end up with the Fresnel equation:

$$\frac{k_x^2\epsilon_x}{n^2 - \epsilon_x} + \frac{k_y^2\epsilon_y}{n^2 - \epsilon_y} + \frac{k_z^2\epsilon_z}{n^2 - \epsilon_z} = 0. \tag{4.6}$$

This equation accepts two solutions for the refractive index n seen by a monochromatic wave propagating in the direction of a wave vector $\mathbf{k}$. In other words, for a given propagation direction, there are two possible phase velocities. Inserted into the eigenvalue problem (eq. 4.4), they will give the corresponding polarization directions which will be orthogonal by definition of eigenvectors. It follows also that in the crystallographic frame, $\sqrt{\epsilon_x}$, $\sqrt{\epsilon_y}$, and $\sqrt{\epsilon_z}$ are the principal refractive indices of the crystal.

As pointed out by eq. (4.2), in order to comply with a constant electrical energy density independent of the wave vector (otherwise the transparency would be **k**-dependent), all the possible polarization directions of the electric field are mapping out an ellipsoid. By definition, $\mathbf{D}\perp\mathbf{k}$ and the plane perpendicular to **k** intersects this ellipsoid drawing an ellipse. This geometrical shape is consistent with what comes out of the Fresnel equation (eq. (4.6)): for a given **k**, only two associated phase velocities are allowed. Their corresponding refractive indices are the lengths of the principal axes of the polarization ellipse.

A normalized representation of the electrical energy density ellipsoid (eq. (4.2)) is the *optical indicatrix* (also referred to as *index ellipsoid*):

$$\frac{x^2}{\epsilon_x}+\frac{y^2}{\epsilon_y}+\frac{z^2}{\epsilon_z}=1. \qquad (4.7)$$

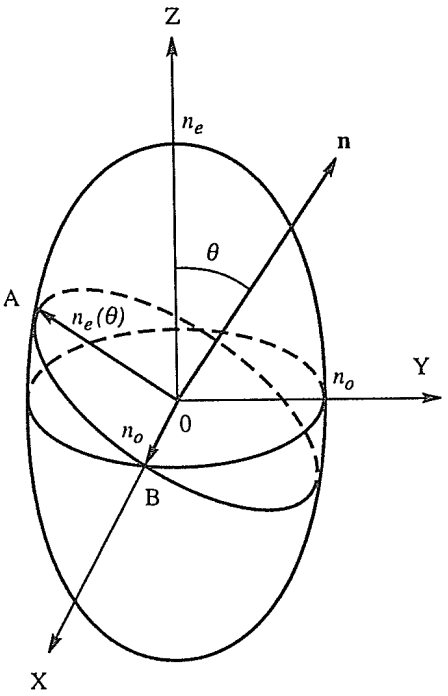

Figure 4.1 – Index ellipsoid of a positive uniaxial crystal with the optical wave vector making an angle θ with the optical axis. The ellipse of the polarizations is also represented. (Reproduced from [120])

This form is only valid in the principal axes of the crystal. As will be seen later, the general form expressed in arbitrary axes is

$$\eta_{ij}x_ix_j=1, \qquad (4.8)$$

where summation is assumed over repeated indices $i,j=1,2,3$, and $\boldsymbol{\eta}=\boldsymbol{\epsilon}^{-1}$ is the relative impermeability tensor.

Usually, the crystals used for AOTF fabrication are uniaxial: their index ellipsoid is partially degenerated as they show a circular symmetry about the optical axis (the only independent direction). Going back to the principal frame, it implies that $\epsilon_x=\epsilon_y=n_o^2$ and $\epsilon_z=n_e^2$. Hence, the optical indicatrix of a uniaxial crystal takes the following form:

$$\frac{x^2+y^2}{n_o^2}+\frac{z^2}{n_e^2}=1 \qquad (4.9)$$

By convention, the crystal is said positive or negative if $n_e > n_o$ or $n_e < n_o$ respectively. The index ellipsoid of a positive uniaxial crystal is illustrated in Fig. 4.1. It is clear that for any angle θ (as defined in Fig. 4.1), one polarization will always see the ordinary index $n_1 = n_o$, while the other will experience a refractive index n_2 dependent on θ as expressed below:

$$n_2^2(\theta) = \frac{n_o^2 n_e^2}{n_e^2 \cos^2\theta + n_o^2 \sin^2\theta}. \tag{4.10}$$

This expression will be involved in the computation of many AOTF parameters.

4.2.2 Acoustic waves in crystals

Acoustic waves are time-varying deformations of a medium. When they propagate through a crystal, its particles are displaced from their equilibrium position, creating internal restoring forces. The particle displacement field is a time-dependent vector quantity representing the distance between the moved and the rest positions. For single frequency acoustic vibrations for instance, it may be represented as:

$$\mathbf{u}(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}) \cos(\Omega t - \mathbf{K}.\mathbf{r}), \tag{4.11}$$

where $\mathbf{r}$ is the particle position vector. However, as $\mathbf{u}$ does not allow to discriminate between deformation and rigid motion, the strain concept is more relevant.

Based on the differential of the displacement vector, the strain matrix assesses the deformation of the medium as the difference of displacement between adjacent particles in all three dimensions. For small displacements, its elements are given by the following expression in rectangular coordinates:

$$S_{ij}(\mathbf{r}, t) = \frac{1}{2}\left(\frac{\partial u_i}{\partial r_j} + \frac{\partial u_j}{\partial r_i}\right) \tag{4.12}$$

with $i, j = x, y, z$. As the strain matrix is symmetric, up to six components are independent. The following naming convention is widely accepted (*abbreviated subscript notation*):

$$\mathbf{S} = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix} = \begin{pmatrix} S_1 & \frac{1}{2}S_6 & \frac{1}{2}S_5 \\ \frac{1}{2}S_6 & S_2 & \frac{1}{2}S_4 \\ \frac{1}{2}S_5 & \frac{1}{2}S_4 & S_3 \end{pmatrix}. \tag{4.13}$$

In some circumstances, when a vector-matrix formalism is more convenient, the six independent strain matrix elements can be computed as

follows:

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} \\ 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{pmatrix} . \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \nabla_{\mathbf{s}}.\mathbf{u} \tag{4.14}$$

A particle displacement gives rise to restoring forces (stresses), assuming elastic deformation. These forces $\mathbf{T}(\mathbf{r}, t)$ apply on the surfaces of infinitesimal volume elements at position $\mathbf{r}$. Using Newton's law of dynamics, the combined action of the body forces ($\mathbf{F}(\mathbf{r}, \mathbf{t})$, due to the acoustic wave) and the stresses induce the movement of the crystal particles. Taking the limit for inifinitesimal volumes leads to the equation of motion for a vibrating medium:

$$\nabla.\mathbf{T} = \rho \frac{\partial^2 \mathbf{u}}{\partial^2 t} - \mathbf{F}, \tag{4.15}$$

where ρ is the mass density at equilibrium.

Hooke's law states that for small elastic deformations, the strength of the restoring forces is linearly proportional to the strains. The proportionality coefficients are called *elastic stiffness constants* and are properties of the medium. In the most general case (crystals with the lowest symmetry), each stress component depends on all the strain components:

$$T_{ij} = c_{ijkl} S_{kl} \tag{4.16}$$

with $i, j, k, l = x, y, z$ and T_{ij} is the ith component of the restoring force acting on the $+j$ face of an infinitesimal volume element. In the above equation the summing convention is assumed on the repeating indices. The c_{ijkl} coefficients are the elements of the rank-4 stiffness tensor $\bar{\mathbf{c}}$. Its inverse is the *compliance* tensor $\bar{\mathbf{s}}$. Dropping the subscripts, eq. (4.16) reduces to the following tensor product,

$$\mathbf{T} = \bar{\mathbf{c}} : \mathbf{S}, \tag{4.17}$$

where the double dots indicate the product of a rank-4 with a rank-2 tensor.

Again, there is an advantage in using abbreviated subscript notation: when the stress components are symmetric, the maximum number of independent elements of $\bar{\mathbf{c}}$ or $\bar{\mathbf{s}}$ falls down to 36 such that one can write $c_{ijkl} = c_{IJ}$ with $I, J = 1, \ldots, 6$ according to the same rule as for the

strain components[1]. With this convention, eq. (4.16) can be rewritten as follows:

$$T_I = c_{IJ} S_J. \tag{4.18}$$

Combining Hooke's law (eq. 4.17) with the particle dynamic equation (eq. 4.15), one can obtain the general acoustic wave equation. Defining the particle velocity $\mathbf{v} = \partial \mathbf{u}/\partial t$, the time-derivative of eq. (4.14), (4.15) and (4.17) lead to

$$\nabla . \bar{\mathbf{c}} : \nabla_{\mathbf{s}} . \mathbf{v} = \rho \frac{\partial^2 \mathbf{v}}{\partial t^2} - \frac{\partial \mathbf{F}}{\partial t}. \tag{4.19}$$

In absence of system excitation ($\mathbf{F} = 0$), and for plane waves described by classical exponential functions, the differential operators of the left-hand side reduce to simple linear equations. Therefore, the overall problem of finding acoustic wave velocities and their respective polarization is a matter of solving an eigenvalue problem.

Eq. (4.19) applies to isotropic and anisotropic crystals. It is usually referred to as the *Christoffel equation*. Solutions can be found for all crystal classes and all propagation directions. In this work, two crystals belonging to the tetragonal class have been used in AOTF configuration. It will be shown later that the AOTF performance increases at slower acoustic velocities. It will also be shown that shear waves (whose particle displacement is perpendicular to the wave vector $\boldsymbol{K}$) are mandatory for designing wide angular aperture devices. Both criteria flow down from a carefull assessment of the Christoffel equation solutions.

4.2.3 Elasto-optic effect and coupled wave equations

As we have seen, the optical wave propagation in crystals fully depends on the optical indicatrix of the medium, i.e. crystal axes and electrical permittivity tensor elements define the encountered refractive indices and the polarization eigenmodes. The elasto-optic (EO) effect, a coupling between internal elastic forces and the medium optical properties, can only be understood if such elastic forces affect the optical indicatrix.

In our case, the strain of the acoustic wave must impact the permittivity $\boldsymbol{\epsilon}$ (the impermeability $\boldsymbol{\eta}$) of the crystal. As the acoustically-induced particle displacement can be highly anisotropic, the change of impermeability at a given position takes all the strain components into account through the use of the elasto-optic (or photo-elastic) tensor $\bar{\boldsymbol{p}}$:

$$\Delta \eta_{ij} = p_{ijkl} S_{kl}. \tag{4.20}$$

[1] As an example, $c_{xxyz} = c_{14}$, or $c_{zxxz} = c_{55}$.

As in the previous sections, the indices i, j refer to x, y, z. The abbreviated subscript notation enables a more convenient vector-matrix formalism:

$$\Delta\eta_I = p_{IJ} S_J, \tag{4.21}$$

where p_{IJ} is a 6×6 matrix.

Only triclinic crystals have 36 independent elements. The other classes show a higher degree of symmetry and often present many zero elements. KDP and TeO_2 crystals belong to the tetragonal family which possesses 7 distinct p_{IJ} elements:

$$\bar{\boldsymbol{p}} = \left(\begin{array}{ccc|ccc} p_{11} & p_{12} & p_{13} & 0 & 0 & 0 \\ p_{12} & p_{11} & p_{13} & 0 & 0 & 0 \\ p_{31} & p_{31} & p_{33} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & p_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{66} \end{array}\right). \tag{4.22}$$

Longitudinal strains have their $S_J \neq 0$ for $J = 1, 2, 3$ only. Taking eq. (4.22) into account, it is clear that from eq. (4.21), only $\Delta\eta_{1,2,3}$ will be non-zero. This means that the change in the index ellipsoid is only about the length of the principal axes, not about its orientation.

On the other hand, a shear strain ($S_{4,5,6} \neq 0$) will give birth to the $\Delta\eta_{4,5,6}$ components which will make off-diagonal terms arise in the quadratic form of eq. (4.8). In that case, the index ellipsoid is rotated and light energy is spread over both re-oriented polarization eigenmodes. This makes the phase matching between the two orthogonal optical polarizations feasible.

We saw that the elastic wave changes the local impermeability (eq. (4.20)). By its connection with the permittivity ($\boldsymbol{\eta} = \boldsymbol{\epsilon}^{-1}$), we will see that it adds an extra term to the local electric field. In absence of strain, the electric displacement vector is given by

$$\mathbf{D} = \epsilon_0 \boldsymbol{\epsilon} \mathbf{E} = \epsilon_0 \left(1 + \boldsymbol{\chi}\right) \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P}, \tag{4.23}$$

with $\boldsymbol{\chi}$ being the dielectric susceptibility of the crystal. In the presence of elastic forces, some extra-susceptibility is added to the system and eq. (4.23) becomes:

$$\mathbf{D} = \epsilon_0 \left(1 + \boldsymbol{\chi} + \boldsymbol{\Delta\chi}\right) \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P} + \mathbf{P}^{\mathrm{AO}}. \tag{4.24}$$

$\mathbf{P}^{\mathrm{AO}}$ is the acoustically-induced dipole.

Similarly to eq. (4.20), the elements of $\boldsymbol{\Delta\chi}$ are the $\Delta\chi_{ij} = \chi^{\mathrm{AO}}_{ijkl} S_{kl}$. It can be shown that the relationship between the susceptibility tensor

and the photo-elastic tensor is as follows: $\chi_{ijkl} = -\epsilon_{im}\epsilon_{jn}p_{mnkl}$. In the frame of the crystal principal axes, where $\boldsymbol{\epsilon}$ is diagonal, it reduces to $\chi_{ijkl} = -n_i^2 n_j^2 p_{ijkl}$.

Inserting eq. (4.24) into the Maxwell equations $\dot{\mathbf{D}} = \nabla \times \mathbf{H}$ and $-\mu_0 \dot{\mathbf{H}} = \nabla \times \mathbf{E}$, and using the identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla.\mathbf{E}) - \nabla^2\mathbf{E}$, one obtains the following expression (assuming $\nabla.\mathbf{E} = 0$):

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\boldsymbol{\epsilon}\ddot{\mathbf{E}} = \frac{1}{c^2\epsilon_0}\ddot{\mathbf{P}}^{\mathrm{AO}}. \tag{4.25}$$

This expression is a classical wave equation in which the right-hand side term acts as a source for the electric field. One interpretation is that the acoustically-induced susceptibility in the medium is responsible for the appearance of a second optical wave as long as it is a solution of the wave equation. In fact, there can be several optical waves created (Raman-Nath regime), but in the following, we will only address the *Bragg regime*: the AO interaction generates a diffracted light beam out of the incident optical energy provided that the acoustic and optical beams propagation directions, power and frequency are adequately chosen.

Solutions to the wave equation (eq. (4.25)) are not so difficult to obtain if one works with plane monochromatic waves. However, we will not describe all the calculation steps as they are not mandatory for understanding the further developments. Assuming that the acoustically-induced strain is described by

$$\mathbf{S}(\mathbf{r},t) = \frac{1}{2}\hat{\mathbf{s}}\left(S(\mathbf{r}).\exp\left(i\Omega t - i\mathbf{K}.\mathbf{r}\right) + \text{c.c.}\right), \tag{4.26}$$

and the incident and diffracted optical waves are described respectively by

$$\mathbf{E_i}(\mathbf{r},t) = \frac{1}{2}\,\hat{\mathbf{e}}^{\mathbf{i}}\left(E_i(\mathbf{r}).\exp\left(i\omega_i t - i\mathbf{k_i}.\mathbf{r}\right) + \text{c.c.}\right), \tag{4.27}$$

and

$$\mathbf{E_d}(\mathbf{r},t) = \frac{1}{2}\,\hat{\mathbf{e}}^{\mathbf{d}}\left(E_d(\mathbf{r}).\exp\left(i\omega_d t - i\mathbf{k_d}.\mathbf{r}\right) + \text{c.c.}\right), \tag{4.28}$$

one can obtain a set of coupled wave equations by first inserting the strain wave (eq. (4.26)) and the incident electric field (eq. (4.27)) into the expression for the acoustically-induced polarization: $\mathbf{P}^{\mathrm{AO}} = \epsilon_0\,\boldsymbol{\Delta\chi}\,\mathbf{E}$. At this moment, the situation is as follows: the strains induced by the acoustic wave give rise to an extra susceptibility $\boldsymbol{\Delta\chi}$. The incident electric field $\mathbf{E_i}$ leans on it to make $\mathbf{P}^{\mathrm{AO}}$ exist. As a source term in the wave equation (eq. (4.25)), $\mathbf{P}^{\mathrm{AO}}$ gives birth to the diffracted electric field $\mathbf{E_d}$. The immediate effect is that $\mathbf{E_d}$ also couples with $\mathbf{S}$ and

generates a dipole which in turn stimulates $\mathbf{E_i}$. In that sense, two coupled equations arise and one must solve the differential equations with some boundary conditions to obtain the amplitude of both optical waves.

Without loss of generality, we can assume that we are interested in measuring how the amplitude of the optical waves varies along the incident direction. In this case, the solutions of the coupled equations take the following form:

$$E_i(z) = E_0 \exp\left(-i\frac{\Delta k}{2}z\right)\left(\cos Tz + i\frac{\Delta k}{2T}\sin Tz\right), \quad (4.29)$$

$$E_d(z) = iE_0 \exp\left(i\frac{\Delta k}{2}z\right)\frac{\Gamma_{di}}{T}\sin Tz. \quad (4.30)$$

In these formulas, the following notations have been used: Δk is the modulus of the phase mismatch vector $\mathbf{\Delta k} = \mathbf{k_i} - (\mathbf{k_d} - \mathbf{K})$ projected onto $\mathbf{k_i}$, and $T = \left(\Gamma_{di}\Gamma_{id} + \Delta k^2/4\right)^{1/2}$ with $\Gamma_{di} = -n_i^2 k_d p S \cos\theta_i/(4\cos\theta_d)$ and $\Gamma_{id} = -n_d^2 k_i p S^*/4$.

One important parameter in these expressions is the effective photo-elastic coefficient $p = \hat{e}_i^i \hat{e}_j^d p_{ijkl} \hat{s}_{kl}$ which measures the photo-elastic coupling for the specific polarization directions of the strain and the incident and diffracted light. In some configurations, p can simply be equal to 0 (as in the case of the optical axis being collinear with a shear wave in TeO_2 f.i.) and the AO coupling cannot take place. Appendix C discusses the cross-orientation of the incident and diffracted beams polarization, and provides an example for the computation of p in a KDP AOTF.

The efficiency of the AO interaction is measured by comparing the optical energy lying in the diffracted beam at the end of the interaction region (L) to that of the incident beam. This is the *diffraction efficiency* (DE) which takes a familiar $\sin^2(x)/x^2$ shape:

$$\mathrm{DE} \equiv \frac{|E_d(L)|^2}{|E_i(0)|^2} = |\Gamma_{di}L|^2 . \left(\frac{\sin TL}{TL}\right)^2 . \quad (4.31)$$

At Bragg incidence, i.e. when the wave vectors of all three coupled waves (the acoustic wave and the incident and diffracted light beams) exactly add up such that $\Delta k = 0$, we can compute the best DE one can get from a given configuration:

$$\mathrm{DE}_{\Delta k \to 0} = \sin^2\left(\frac{\pi L}{\lambda}\sqrt{\frac{M_2 P_a}{2H_\mathrm{T}L_\mathrm{T}}\frac{\cos\theta_i}{\cos\theta_d}}\right), \quad (4.32)$$

where the strain amplitude has been replaced by the acoustic power P_a, a more useful quantity obtained from a relationship involving the medium

mass density ρ, the acoustic wave phase velocity v and the transducer height (H_T) and length (L_T):

$$P_a = \frac{\rho v^3 S^2 H_\mathrm{T} L_\mathrm{T}}{2}. \tag{4.33}$$

The M_2 parameter in eq. (4.32) is the *acousto-optic figure of merit* which is completely determined by the crystal properties and the waves propagation directions:

$$M_2 = \frac{n_i^3 n_d^3 p^2}{\rho v^3}. \tag{4.34}$$

As can be seen, the filter DE is improved for long and narrow transducers driven at higher power (but too much power will lower the efficiency by drawing it to the right, decreasing, side of the $\sin^2$). Also, the performance will always be better at shorter wavelengths. However, the next section will show that these parameters also control several other important aspects of an AOTF (separation angle, passband, frequency range,...) and often trade-off studies will be necessary.

4.2.4 Basic AOTF parameters

Working with coupled wave equations is in general not really convenient. As AO interaction is the coupling between light (photons) and sound (phonons) inside a medium, another useful approach for characterizing AOTF is therefore to rely on the wave vector diagram which illustrates the conservation of momentum: $\mathbf{k_d} = \mathbf{k_i} \pm \mathbf{K}$.

The situation for birefringent AO coupling in TeO_2 is illustrated in Fig. (4.2), where the wave vector modulus of the incident ($\boldsymbol{k}_i$) and diffracted ($\boldsymbol{k}_d$) light, and of the acoustic wave ($\mathbf{K}$) of frequency F are:

$$k_i = \frac{2\pi n_i}{\lambda}, \quad k_d = \frac{2\pi n_d}{\lambda}, \quad K = \frac{2\pi F}{v}. \tag{4.35}$$

Such configuration is usually referred to as *off-axis* anisotropic interaction ([24],[121]) because it generalizes the concept of collinear interaction. In this geometry, the significant advantage lies in the relative insensitivity of the coupling strength to small variations of θ_i (*non-critical* phase matching), carrying the opportunity for large angular aperture filters.

4.2.4.1 Tuning curve

The 001 axis is the optical axis of the uniaxial crystal along which any light polarization state sees the same *ordinary* refractive index n_o. Con-

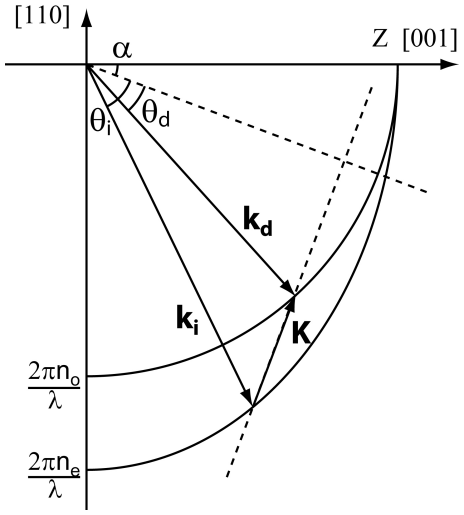

Figure 4.2 – Momentum vector diagram illustrating the birefringent AO interaction in TeO_2.

sistently with the discussion of section 4.2.1, any departure from this axis will introduce some birefringence between different polarization states.

The main axes of the elliptic cross-section between the plane perpendicular to the wave vector and the optical indicatrix, as illustrated in Fig. (4.1), are n_o and $n_e(\theta)$. In the situation of Fig. (4.2), the incident light is polarized along the *extraordinary* axis and has its refractive index $n_i(\theta_i + \alpha)$ given by eq. (4.10), where α is the angle between the acoustic wave vector and the 110 axis. On the other side, the diffracted beam will see $n_d = n_o$, independently of θ_d.

The 110 axis is the direction along which an acoustic wave with shear deformation along the $\bar{1}10$ axis has its slowest phase velocity [4]. The 110 and 001 axes form what is usually called the $t - Z$ plane, and the phase velocity for a $\bar{1}10$-polarized shear wave varies in function of the *cut-angle* α as:

$$v^2(\alpha) = \frac{c_{11} - c_{12}}{2\rho} \cos^2 \alpha + \frac{c_{44}}{\rho} \sin^2 \alpha. \tag{4.36}$$

Extreme values for v are 616 ms^{-1} along 110 and 2104 ms^{-1} along 001.

Working out trigonometric relationships in the triangles of the wave vector diagram, the following expression arise:

$$K = k_i \sin \theta_i - \sqrt{k_d^2 - k_i^2 \cos^2 \theta_i}. \tag{4.37}$$

Using the definitions of eq. (4.35), eq. (4.37) becomes the *tuning curve* of the AOTF, i.e. the relationship providing the frequency of the acoustic

wave which enables its coupling with an incident light beam of wavelength λ for a given material and interaction geometry:

$$F = \frac{v(\alpha)}{\lambda}\left(n_i(\theta_i+\alpha).\sin\theta_i - \sqrt{n_o^2 - n_i^2(\theta_i+\alpha).\cos^2\theta_i}\right). \qquad (4.38)$$

Again, as this example illustrates an extraordinary-polarized incident light (e-light), $n_i(\theta_i+\alpha)$ is given by eq. (4.10).

In order to get a better feeling on this important relationship, let us define some test cases with a TeO_2 crystal. First, a good formula for the computation of the refractive indices is needed for the spectral and temperature ranges of interest. Uchida [103] published a comprehensive characterization of TeO_2 optical properties, including the fitting of refractive index measurements at 293K with a two-oscillators Sellmeier formula over the visible range. His data were merged with those taken at longer wavelengths measured by other groups and a global fit over all datasets [46] led to the following expressions (valid from 0.4 to 3.5 μm):

$$n_o^2(\lambda) = 1 + \frac{3.71789\lambda^2}{\lambda^2 - 0.19619^2} + \frac{0.07544\lambda^2}{\lambda^2 - 4.61196^2}, \qquad (4.39)$$

$$n_e^2(\lambda) = 1 + \frac{4.33449\lambda^2}{\lambda^2 - 0.20242^2} + \frac{0.14739\lambda^2}{\lambda^2 - 4.93667^2}. \qquad (4.40)$$

Fig. 4.3 shows the dispersion of n_o and n_e over the visible range as computed from eq. (4.39)-(4.40). As can be seen, the birefringence varies in function of λ: at 400 nm, $\Delta n \equiv |n_e - n_o| = 0.1854$, whereas at 1 μm, $\Delta n = 0.14398$.

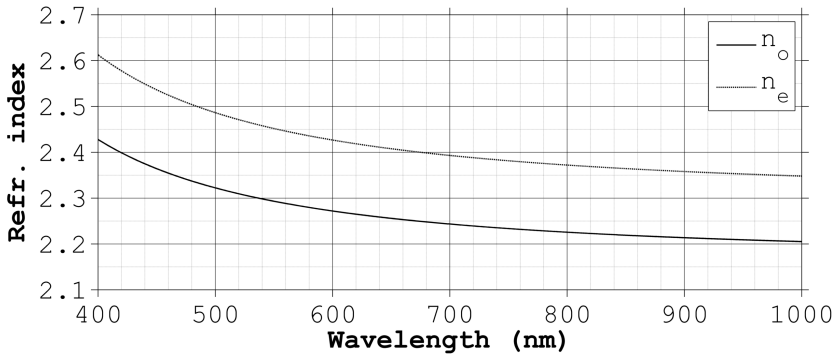

Figure 4.3 – Refractive indices of TeO_2 at 293K. n_e is always larger than n_o.

For the test cases, we will consider two wavelengths emitted by ion lasers: 476.5 nm (Ar) and 632.8 nm (HeNe). The cut-angle α (the angle between the acoustic wave vector **K** and axis 110), will take values

between 6° and 16°, by steps of 2°. Using eq. (4.36), the corresponding acoustic speed $v(\alpha)$ can be obtained. Finally, knowing the wavelengths, the refractive indices and the acoustic phase velocity, the tuning relationship can be solved for a wide range of incident angles. Results are shown in Fig. (4.4).

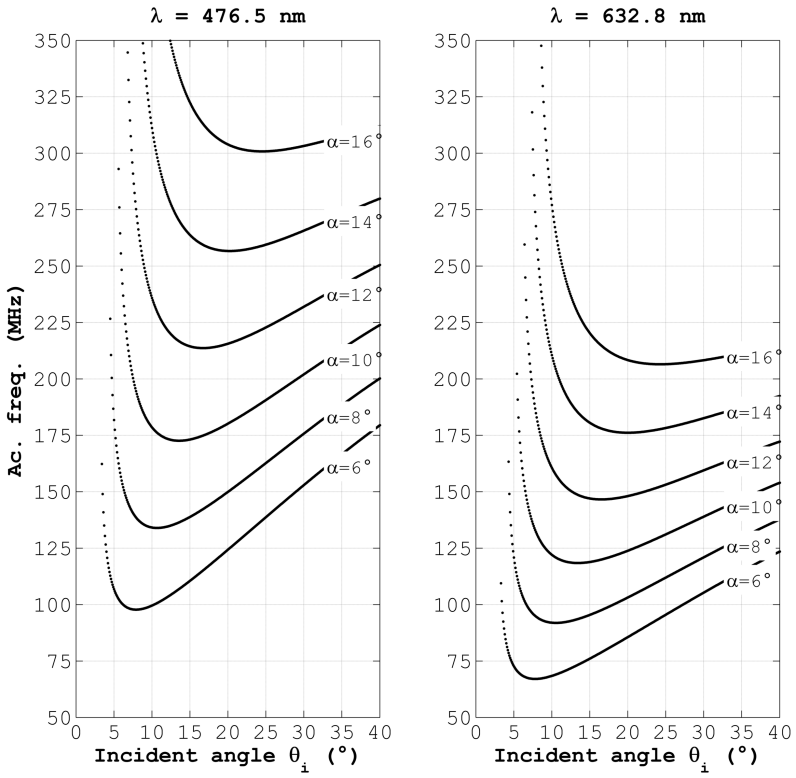

Figure 4.4 – Relationship between the acoustic frequency and the optical incident angle for efficient AO interaction in a TeO_2 crystal. One can notice that the trough of the curve becomes wider as the cut-angle increases. In TeO_2, the flatness is reached at $\alpha = 18.9°$ (not represented).

All these curves show a minimum characterized by $\partial F/\partial\theta_i = 0$, which is the region where AOTF's operate: with $\partial F/\partial\theta_i$ being minimized, the filter can accept a cone of light with a non-negligible angular size while still being driven at a single frequency. If desired, larger angular apertures can be obtained by chosing larger values of α at the price of higher driving frequencies. In the following, the values corresponding to the minima of the $F - \theta_i$ curves will be referred to as F_B and θ_B where the B stands for *Bragg* diffraction.

Another important aspect to notice, is that from one wavelength to another, for a given cut-angle, θ_B stays the same. The angles α and θ_B are therefore the only parameters to freeze when designing an AOTF. The other filter properties arise from this initial choice, such as the tuning curve $F(\lambda)$ (eq. (4.38)). Several tuning curves are plotted in Fig. 4.5.

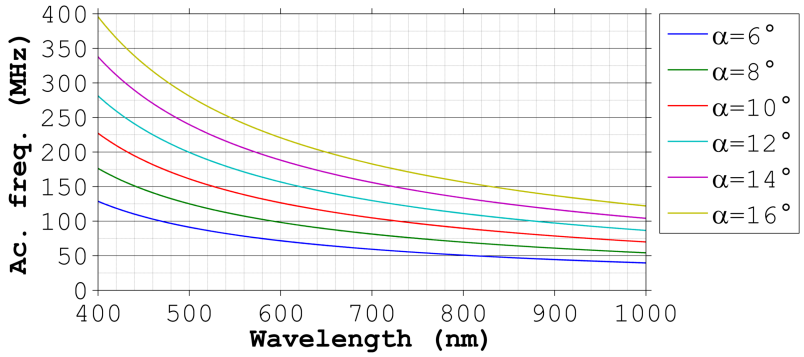

Figure 4.5 – Tuning curves for a TeO_2 AOTF computed for several cut-angles.

4.2.4.2 Separation angle

When spectral imaging is the target, the role of the filter is to suppress a large part of the incident spectrum and let the desired wavelength go through. The AOTF has a strong advantage compared to filters because the acousto-optically diffracted beam travels with a rotated polarization and in a slightly diverted direction with respect to the incident light. The removal strategy of the unwanted spectrum can rely on both aspects.

State of the art polarizers can achieve extinction ratios as high as 10^5:1. Though attractive, this extinction level is probably extremely difficult to achieve such that the cleanest signals will be obtained if the acceptance angle is also kept smaller than the separation angle. With this approach, incident and diffracted beams will not overlap, which makes the optical design easier (as baffles can be used to block one of the beams). This spatial selection is a key aspect but often a limiting design factor as well.

According to the wave vector diagram (Fig. 4.2), we have $n_i \cos\theta_i = n_o \cos\theta_d$. Knowing θ_i, it is easy to obtain the separation angle between incident and diffracted beams inside the crystal: $\Delta\theta \equiv |\theta_i - \theta_d|$. At the crystal boundary, the refraction law increases the angular distance: $\Delta\theta_{\text{air}} = n_o.\Delta\theta$. Fig. 4.6 shows the separation angle in air for our test

cases at Bragg incidence. As it can be seen, the large birefringence of TeO_2 makes angular distances of several degrees easily achievable. The dispersion of n_o makes $\Delta\theta_{\text{air}}$ dependent on λ. But by tilting the output facet of the crystal, this dependence can be largely removed over a broad spectral range.

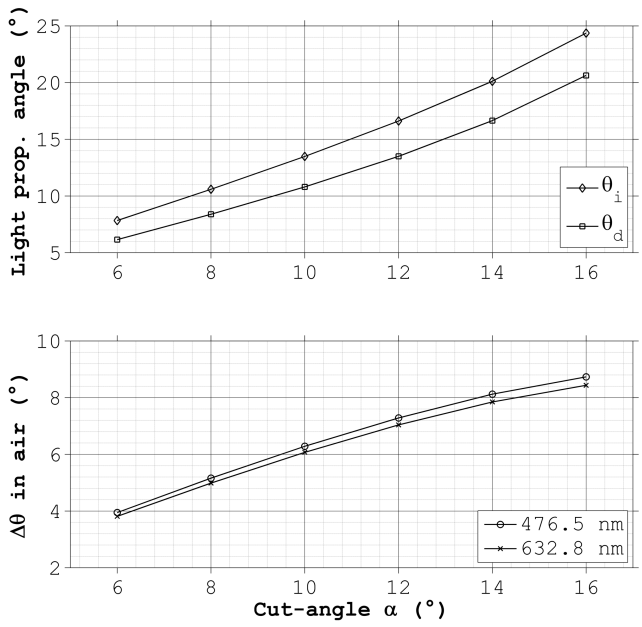

Figure 4.6 – Top: Evolution of the Bragg matching incident (e-light) and diffracted (o-light) angles as a function of the cut-angle α in a TeO_2 AOTF. Bottom: Separation angle at the crystal output facet in air.

The restriction of a smaller acceptance angle than the separation angle in order to facilitate extinction of the unwanted order of diffraction has a side-effect: it limits the étendue (throughput) of the optical system. A reduced étendue means a lower luminosity for the instrument. Another AOTF parameter impacting the amount of transmitted photons is the filter passband. From the point of view of the capability to resolve the spectral signatures of absorbing species, the narrower is of course the better. On the other hand, the amount of photons sent into the diffracted order is proportional to the convolution of the incident light spectrum with the filter transfer function (the diffraction efficiency curve of eq. (4.31)). A trade-off has to be performed to maximize both aspects.

4.2.4.3 Diffraction efficiency

The AOTF passband is usually measured as the full width at half maximum (FWHM) of the DE sinc^2 main lobe (eq. (4.32)). The steps to derive the filter spectral transfer function (STF) are not straightforward. For this, we need to compute the DE as a function of the mismatch vector $\boldsymbol{\Delta k}$. As an example, we will make this exercise for one of our test cases, with $\lambda = 632.8$ nm and $\alpha = 10°$. We will also assume the same interaction geometry as illustrated in Fig. 4.2. As we work in an off-axis AOTF configuration (i.e. with minimized angular sensitivity), the correct incidence angle for e-light is the one minimizing the $F - \theta_i$ curve on Fig. 4.4. Computing the refractive index values from eq. (4.39), (4.40) and (4.10), and using eq. (4.36) to obtain $v(\alpha) = 708$ ms^{-1}, we find $\theta_i = 13.4°$. From the wave vector diagram, we have $\theta_d = \arccos n_i \cos\theta_i / n_o = 10.8°$. The matching acoustic frequency is given by eq. (4.38): $F_B = 118.44$ MHz.

If, for whatever reason, one or several of these parameters are changed, the AO interaction efficiency will be reduced as the mismatch vector increases. In the following, we assume that F is the only varying parameter. This is typical of a measurement procedure when one whishes to measure the STF: a laser beam is sent through the AOTF, while the acoustic frequency is swept over a few MHz. In the same time, the incident and diffracted beams intensities are monitored and a curve DE(F) can be drawn.

For consistency with the theoretical developments, the mismatch vector is computed along the incident direction:

$$\begin{aligned} \Delta k &= \boldsymbol{\Delta \mathbf{k}}.\frac{\mathbf{k_i}}{k_i} \\ &= (\mathbf{k_i} - (\mathbf{k_d} - \mathbf{K})).\frac{\mathbf{k_i}}{k_i} \\ &= k_i - k_d \cos\Delta\theta - K \sin\theta_i. \end{aligned} \tag{4.41}$$

Following our assumption of creating the mismatch by the sole acoustic wave frequency detuning, only K will take different values in the above expression.

The next step is to freeze the transducer parameters. By fixing its height H_T, length L_T, and the driving RF signal power P_a, eq. (4.33) allows to compute the strain amplitude S. For a transducer of $H_T = 10$ mm and $L_T = 15$ mm, and by driving it with 0.3 W, we find $S = 4.3 \times 10^{-5}$.

The remaining quantity is the most difficult to compute. It is the

effective photo-elastic coefficient p. This coefficient measures how the relative orientations of the shearing strain and the polarizations of the incident and diffracted optical waves do couple in an efficient way via the photo-elastic tensor:

$$p = \hat{e}_i^i \hat{e}_j^d p_{ijkl} \hat{s}_{kl}. \tag{4.42}$$

The $\hat{}$ symbol denotes a unit vector indicating the polarization ($\hat{\boldsymbol{e}}^i$, $\hat{\boldsymbol{e}}^d$) and shearing ($\hat{\boldsymbol{s}}$) directions in the crystal axes frame. Eq. (4.42) involves several manipulations and we will only provide the final result. The full developments can be found in [120]. For a $\bar{1}10$-polarized shear strain propagating at an angle α from the 110 axis, and for incident e-light and diffracted o-light propagating at $\theta_i + \alpha$ and $\theta_d + \alpha$ from 001 axis respectively, the effective photo-elastic coefficient is given by:

$$p = \frac{1}{2}.\left(p_{11} - p_{12}\right).\cos\alpha.\cos(\theta_i + \alpha) + p_{44}.\sin\alpha.\sin(\theta_i + \alpha). \tag{4.43}$$

For TeO_2, despite a significant spread of values found in the litterature for the p_{ij} coefficients, we take the numbers from [112]: $p_{11} - p_{12} = -0.226$, $p_{44} = -0.044$. In our case, we find $p = -0.105$.

Finally, the quantities Γ_{di} and Γ_{id}, which have been introduced in the previous section, can be computed for our interaction geometry:

$$\Gamma_{di} = \frac{-n_i^2 k_d\, p\, S \cos(\theta_i + \alpha)}{4\cos(\theta_d + \alpha)}, \tag{4.44}$$

$$\Gamma_{id} = \frac{-n_d^2 k_i\, p\, S}{4}. \tag{4.45}$$

And the diffraction efficiency for a given mismatch vector follows from eq. (4.31):

$$\mathrm{DE}(\Delta k) = |\Gamma_{di} L|^2 \left(\frac{\sin\left(L\sqrt{\Gamma_{di}\Gamma_{id} + \frac{\Delta k^2}{4}}\right)}{L\sqrt{\Gamma_{di}\Gamma_{id} + \frac{\Delta k^2}{4}}} \right)^2. \tag{4.46}$$

Here, L is not equivalent to L_T, the transducer length. It represents the total length of AO interaction, i.e. taking the propagation angles into account. It also accounts for the walk-off of the acoustic beam: in anisotropic media, the group velocity is most of the time not parallel to the wave vector.

By definition, the group velocity is the derivative of the phase velocity with respect to the direction of propagation: $\boldsymbol{v_g} = \boldsymbol{\nabla_K}\, \boldsymbol{v}$. Conversely,

$\boldsymbol{v_g}$ is the normal to the slowness curve, the inverse of the phase velocity: $\vartheta(\alpha) \equiv K/\Omega = 1/v(\alpha)$. It takes the following expression in the $\bar{1}10$ plane in TeO_2:

$$\vartheta(\alpha) \equiv \frac{1}{v(\alpha)} = \left(v_{110}^2 \cos^2\alpha + v_{001}^2 \sin^2\alpha\right)^{-1/2}. \tag{4.47}$$

The ellipse of eq. (4.47) and the relevant angles are illustrated in Fig. 4.7. The walk-off angle Ψ is the angle between the ellipse radius $\vartheta(\alpha)$ and the group velocity $\boldsymbol{v}_g$. Following the drawing notation, we

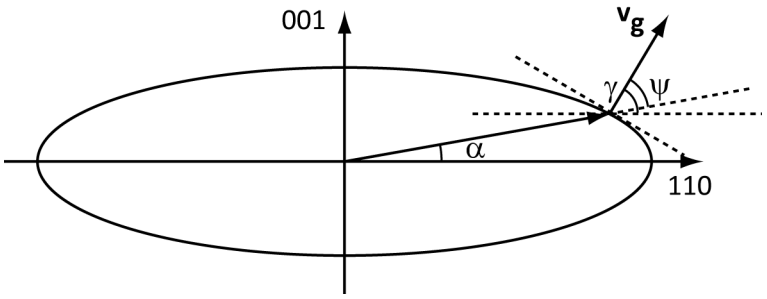

Figure 4.7 – Slowness curve for a $\bar{1}10$-polarized shear acoustic wave propagating in the $\bar{1}10$ plane in TeO_2.

have $\tan\Psi = \tan(\gamma - \alpha)$. With some manipulations, this equation can be rewritten as

$$\tan\Psi = \frac{(v_{001}^2 - v_{110}^2)}{2v^2} \sin 2\alpha. \tag{4.48}$$

Due to the large ellipticity of the slowness curve of this particular acoustic propagation mode, the walk-off of the acoustic beam can be quite severe: for $\alpha = 10°$, we have $\Psi = 54°$.

With the above expression for the walk-off angle, we can evaluate the AO interaction length. Fig. 4.8 illustrates the walk-off of the acoustic beam in the interaction plane. Based on this representation, the law of sines provides the following relationship:

$$L = \frac{L_T \cos\Psi}{\cos(\Psi - \theta_i)}. \tag{4.49}$$

Accordingly, for a transducer length of 15 mm, the interaction region will reduce in this case to 11.6 mm.

We can now compute the diffraction efficiency of the AOTF for the current test case settings: $\lambda = 632.8$ nm, $L_T = 15$ mm, $H_T = 10$ mm, $\alpha = 10°$ and $P_a = 0.3$ W. The result is shown in Fig. 4.9. This

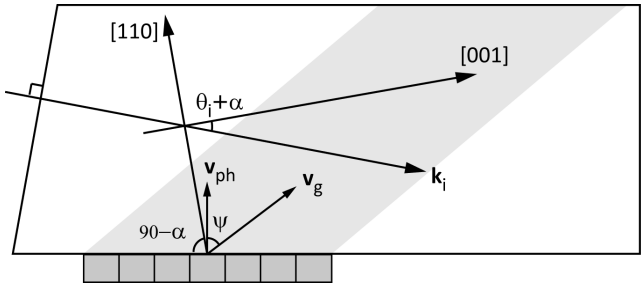

Figure 4.8 – Illustration of the acoustic beam walk-off and how the interaction length is impacted.

representation is valid only at Bragg incidence, i.e. $\theta_i = 13.4°$ for e-light. We see that with these parameters, the 100% efficiency is almost achieved.

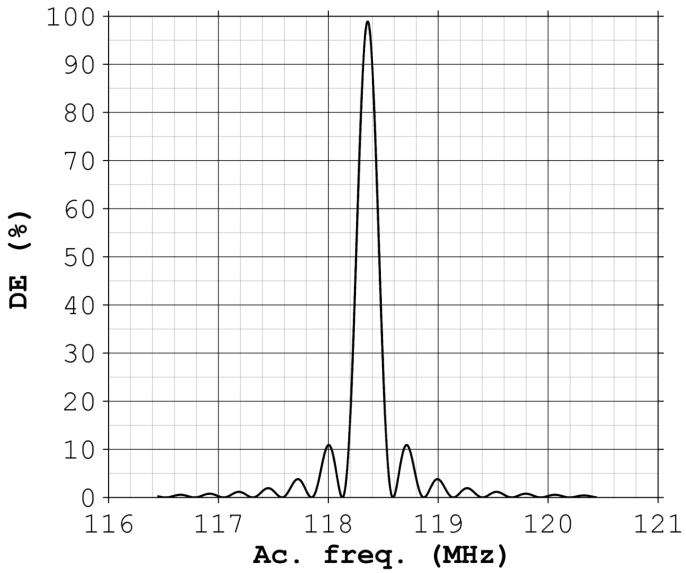

Figure 4.9 – Theoretical DE curve computed for $\lambda = 632.8$ nm, $L_T = 15$ mm, $H_T = 10$ mm, $\alpha = 10°$, $P_a = 0.3$ W.

4.2.4.4 Passband

The starting point for this discussion on drawing the DE curve was to estimate the FWHM of the main lobe as a relevant quantity for passband $\Delta\lambda$ and hence luminosity aspects. In the current configuration, we find $\Delta F = 0.212$ MHz from eq. (4.46), which corresponds

to $\Delta\lambda = \Delta F/F_B.\lambda_B = 1.13$ nm. Most of the time, however, an approximated expression for $\Delta\lambda$ is used and the full DE calculation is unnecessary. Following [112], it takes the form of:

$$\Delta\lambda = \frac{0.8.\lambda^2.\cos(\Psi - \theta_i)}{\Delta n.L_T.\cos\Psi.\sin^2(\theta_i + \alpha)}. \tag{4.50}$$

With this formula, the same settings lead to a passband of $\Delta\lambda = 1.14$ nm, in very close agreement with exact theory. Extending the range of computation to the complete visible domain and for all cut-angles, we obtain the curves drawn in Fig. 4.10. Clearly, the AOTF spectral resolution follows a λ^2 law, notwithstanding a small correction due to the wavelength-dependent birefringence Δn.

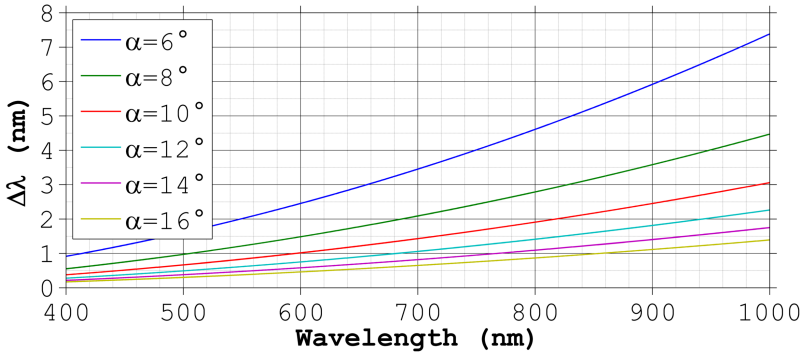

Figure 4.10 – Passband of a TeO_2 AOTF computed for several cut-angles over the visible range.

4.2.4.5 Driving power

Finally, as we are interested in getting the highest efficiency from the filter, it is worth having a look at the parameters controlling the peak DE when the essential AOTF features (α, θ_B, F_B) are frozen. We start from eq. (4.32) and modify it to account for the acoustic beam walk-off:

$$\mathrm{DE}_{\Delta k\to 0} = \sin^2\left(\frac{\pi}{\lambda}\sqrt{\frac{M_2.P_a.L_T}{2.H_T}\frac{\cos\theta_i}{\cos\theta_d}}\frac{\cos\Psi}{\cos(\Psi-\theta_i)}\right). \tag{4.51}$$

This function is plotted in Fig. 4.11 for $\alpha = 10°$, $L_T = 15$ mm and $H_T = 10$ mm at our two test wavelengths in function of the driving power P_a. As expected by the $\sin^2$ function, one must be cautious with the applied power as it can be detrimental to the efficiency (and to the

transducer) if set too high. The operational point should stay around the first maximum, which clearly exhibits a strong wavelength dependence.

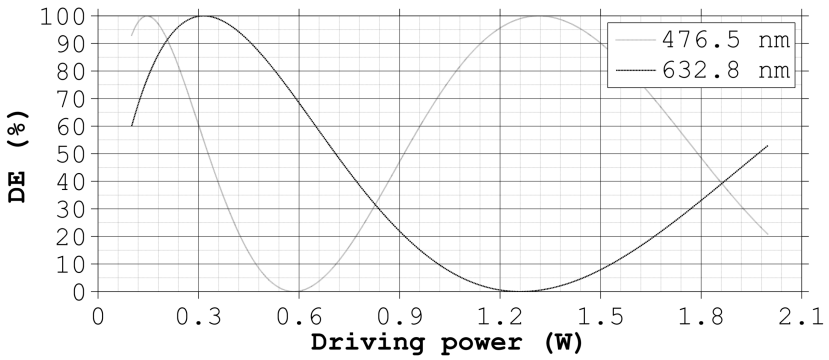

Figure 4.11 – DE in function of the driving power in Bragg matching condition ($\Delta k = 0$) with a transducer of $L_T = 15$ mm and $H_T = 10$ mm.

In the context of an AOTF-based spectral imager, the dependence of the maximum DE in function of λ is crucial for the radiometric consistency of measurements involing several wavelengths. According to eq. (4.51), the acoustic power settings should follow:

$$P_a(\lambda) = \frac{H_T}{L_T}\left(\frac{\lambda}{2}\frac{\cos(\Psi - \theta_i)}{\cos\Psi}\right)^2 \frac{2\cos\theta_d}{M_2.\cos\theta_i}. \tag{4.52}$$

This relationship is represented in Fig. 4.12 for the same cut-angle and transducer dimensions. Eq. (4.52) can be used during the AOTF

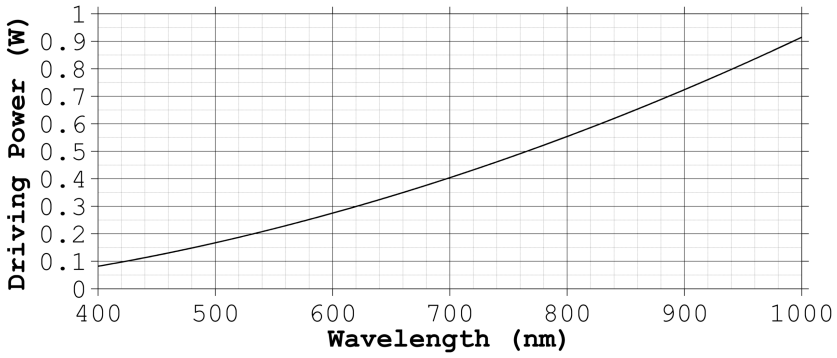

Figure 4.12 – Curve showing the driving power to be applied to the transducer in order to preserve the maximum DE over the tuning range. (TeO_2 AOTF, $\alpha = 10°$, $L_T = 15$ mm, $H_T = 10$ mm)

design process too, in the sizing of the transducer for instance when the available power is the limiting factor.

With this summary of the underlying theory of birefringent acousto-optic interaction and the derivation of the most important formula allowing to understand and control many key aspects of an AOTF, we end up with a comprehensive package for anyone wishing to develop an AOTF-based instrument. As we have seen, after the selection of the crystal, the choice of the cut-angle α has a strong impact on the overall behaviour of the AOTF. A small angle will allow to drive the filter at smaller frequencies, but will offer a reduced separation angle for a broader passband. Selecting a larger cut-angle will benefit to the spectral resolution and the angular aperture, but the AOTF will have to be driven at much higher frequencies. All these aspects have to be taken into account at the design level. And it is often a matter of trading off the optical throughput, the spectral resolution, the driving power and the tuning range. Unfortunately, they are all entangled in the elasto-optic effect and a full model is necessary to take decisions on single parameters.

4.3 Application to NO_2 absorption cross-section measurement

A widespread method for the remote sensing of atmospheric species in the UV-VIS range, makes use of a grating spectrometer coupled to a scanning system (SCIAMACHY [15], OSIRIS [73],...). From the pure spectroscopic point of view, it has a number of advantages against which the AOTF cannot compete: the spectrum over a broad spectral window is recorded at once and the spectral resolution is usually well below 1 nm with a constant sampling.

Though not being its natural operational mode (as the acousto-optic interaction only occurs for one wavelength and one frequency at a time), an AOTF can also record continuous spectra, provided that acquisition time is not a constraint. The method is simple: by sweeping the acoustic wave frequency with reasonably small steps compared to the passband, one can record the diffracted beam intensity as a function of the frequency. Using the filter tuning curve (like in Fig. 4.5), the frequency scale can be converted to the wavelength domain.

The rationale behind this is that by looking at the nicely structured absorption spectrum of a relevant species, it provides a much better feeling on the spectral resolution of an AOTF-based instrument. Sec-

ond, it is an excellent performance test for the tuning accuracy (how the driving electronics can be trusted) as hundreds of small steps are performed. Finally, it is one of the possible methods for the instrument wavelength scale calibration as the well-known spectral structures can serve as absolute references.

Among the family of atmospheric species, NO_2 is probably the best choice for this exercise: it has an important role in the global ozone chemistry, it is relatively abundant in the stratosphere (>0.1 ppb between 25 and 40 km, [19]), and it presents a quite structured absorption spectrum [107] whose differential structures are compatible with a ~ 1 nm passband.

For these reasons, we tested a TeO_2 AOTF in a laboratory experiment involving an absorption cell filled with NO_2. The following paragraphs describe the AOTF under testing and the experimental set-up. The results will be compared to a reference spectrum and analysed. This experiment is a very unconventional test case for an AOTF, and to our knowledge, it is the first time such experiment is reported.

4.3.1 Commercial TeO_2 AOTF

In the frame of the ALTIUS visible channel breadboarding activity, two identical AOTF's made of TeO_2 were selected in the products catalogue of the *Gooch & Housego* company (G&H). Their key parameters were undisclosed by G&H at the time of the purchase. Only a limited number of basic specifications were readily available: the optical aperture, the overall dimensions of the housing, rough estimations of the passband, the spectral range and the tuning range, the transparency.

This insufficient knowledge was completed by some preliminary laboratory tests. Using an Ar laser from the Institute of Condensed Matter and Nanosciences (IMCN) of the Université Catholique de Louvain (UCL), we could retrieve all the key parameters of the AOTF's. The method was based on carefully aligning the AOTF with the vertically-polarized laser beam (at perpendicular incidence), and finding the matching acoustic frequency (the one maximizing the diffraction efficiency) for each laser wavelength (five Ar wavelengths were used). These points were fitted with a tuning curve (eq. (4.38)) in order to determine the cut-angle α and the incident angle θ_i. Only the transducer length could not be retrieved with good confidence. The reason is that G&H is using a proprietary apodization technique to reduce the amplitude of the spectral transfer function sidelobes playing on the *apparent* length of the transducer. However, the efficiency of the technique is limited to a

Parameter	Spec.	Parameter	Spec.
Linear aperture	10×10 mm^2	Cut-angle α	7.65°
Transducer length L_T	15 mm	θ_i (o-light)	8°
Transducer height H_T	10 mm	θ_d (e-light)	10°
Tuning range	60-150 MHz	$\Delta\theta_{\text{air}}$ @600 nm	4.5°
Spectral range	425-800 nm	Passband $\Delta\lambda$ @600 nm	2 nm

Table 4.1 – Main parameters of the G&H TeO_2 AOTF's.

portion of the tuning range (the longer wavelengths in our case). The immediate consequences are that the passband $\Delta\lambda$ deviates from eq. (4.50) and the optimal driving power as determined from eq. (4.52) do not hold anymore. Despite this uncertainty, an average transducer length could be determined. The AOTF specifications are summarized in Table 4.1.

4.3.2 Experimental setup

The measurements took place at the Katholieke Universiteit Leuven, in the laboratory of Pr. Shaun Carl, with the help of his assistant Victor Khamaganov. A long-path absorption gas cell was prepared and aligned with a deuterium lamp. A collimating lens was used to make the light beam parallel throughout the cell. Lenses were used on the other side of the cell to match the AOTF optical aperture and focus the diffracted order onto a power meter. The AOTF driving electronics and the reading of the photodiode were controlled by a LabView software developed by Bert Van Opstal.

The cell was 120 cm long. It was connected to a gas tank containing a mixture of He (95%) and NO_2 (5%). The cell could be filled with different pressures in static or flow conditions. The results which will be discussed below were obtained with a total pressure of 10 Torr and a constant flow. Using the ideal gas law

$$P.V = N.k.T, \tag{4.53}$$

with $P = 66.7$ Pa the partial pressure of NO_2, k the Boltzmann constant (1.38×10^{-23} m^2 kg s^{-2} K^{-1}), and $T = 294$ K, the room temperature. The gas concentration in the cell was estimated to $n = N/V = 1.64 \times 10^{16}$ molecules cm^{-3}.

The measurement strategy consisted of two steps. First, a reference measurement is recorded with the empty cell. It actually consists of the deuterium lamp spectral irradiance ($I_0(\lambda)$) convolved with the spectral

transfer function of the overall system (AOTF, lenses and cell windows). Practically, the acoustic frequency was swept by steps corresponding to 0.1 nm in the wavelength space. At each step, the diffracted beam power was averaged over 0.5 seconds, then corrected for the background signal and stored. After completion of this reference scan, the cell was filled at the nominal pressure and the same acquisition sequence was run.

As in the atmosphere, the extinction of light by absorption from NO_2 follows the Beer-Lambert principle (see Chapter 1):

$$I(\lambda) = I_0(\lambda).\exp\left(-\sigma(\lambda)\, l\, n\right). \tag{4.54}$$

In this expression, $\sigma(\lambda)$ is the absorption cross-section that we want to measure, l is the absorption path length (the cell length), and n is the gas concentration.

4.3.3 Results

Fig. 4.13 shows the experimental data superimposed on the reference spectrum recorded with a Fourier transform spectrometer at a 0.1 cm^{-1} resolution [106]. The agreement is very good except around 486 nm, where the strong D_β emission line from the hydrogen Balmer series corrupted the data by saturating the sensor (thus a lamp effect independent of the AOTF).

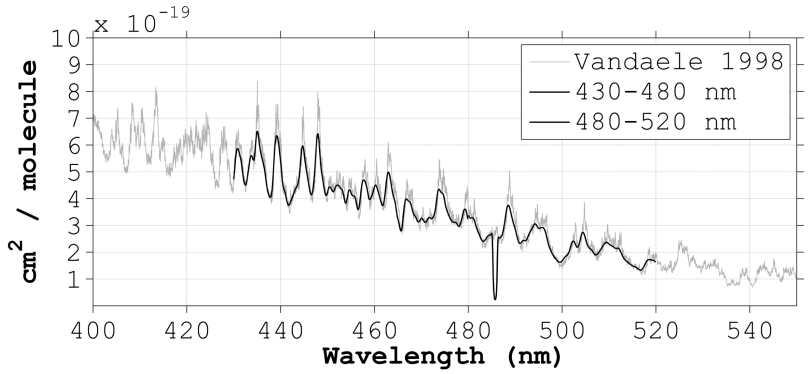

Figure 4.13 – NO_2 absorption cross-section measured with an AOTF (black curve) and compared to a reference spectrum (grey curve).

Clearly, a remote sensing instrument based on an AOTF showing similar specifications as in Table 4.1 allows to resolve the characteristic structures of the NO_2 absorption cross-section. This was anticipated as such a filter possesses a passband of about 1 nm in the range of interest, a sufficiently small value for this molecule. But as the measurement

principle relied on a quite unusual way of operating the AOTF (hundreds of close steps), this was yet to be demonstrated.

The resolving capabilities of such filter make it well suited for differential acquisition. Taking advantage of the NO_2 structures around 440 nm, one can retrieve NO_2 concentration from the ratio of the signal acquired at one strongly and one weakly absorbing wavelengths (see Chapter 5 or section 3.7). This method was already successfully applied in the 80's by the solar occultation instrument SAGE II [32]. It was also proven to work with the more recent limb scattering instrument OSIRIS [12]. In Chapter 5 we will show that it is also a valid approach in a ground-based remote sensing application.

4.4 KDP-based AOTF for hyperspectral imaging in UV

The unique properties of TeO_2 have drawn much of the research on acousto-optics in the last decades. Naturally, the transparency range of this material (0.35-5 μm) led to many applications targeting the visible and infrared domains. Despite interest for UV spectroscopy and hyperspectral imaging, the small number of birefringent materials transparent at shorter wavelengths and having good AO performance delayed the developments in this spectral region. For an atmospheric remote sensing instrument however, extending its sensitivity range down to UV wavelengths is of high importance: the upper-stratospheric and mesospheric ozone is only accessible to measurements performed at wavelengths shorter than 330 nm (peak absorption at 250 nm), and other trace gases like BrO and OClO exhibit characteristic absorption features between 300 and 350 nm.

Table 4.2 lists the major candidates for the interaction medium of a UV AOTF. Their UV cut-off wavelength, birefringence at 250 nm and AO figure of merit M_2 (defined by eq. (4.34)) for non-critical phase matching configuration are provided. While a small value for Δn helps keeping the matching acoustic frequency on the low side (eq. (4.38)), a large M_2 value indicates a good filter efficiency with low driving power (eq. (4.52)).

Considering the ease of manufacturing, quartz is of course the first choice. Consequently, the first reported UV noncollinear AOTF was made of quartz [63]. However, its low figure of merit makes substantial diffraction efficiency difficult to achieve. KDP, a well-known nonlinear optical material, was introduced a few years later in the field of acousto-

Material	Cut-off λ (μm)	Δn (at 0.25 μm)	M_2 ($\times 10^{-18} s^3 g^{-1}$)
SiO_2 (quartz)	0.12	0.01	0.59
Al_2O_3 (sapphire)	0.15	0.01	0.3
$NH_4H_2PO_4$ (ADP)	0.13	0.05	5.1
MgF_2	0.11	0.01	0.5
KH_2PO_4 (KDP)	0.2	0.05	4.6

Table 4.2 – Selected properties of birefringent materials transparent in the UV.

optics [91]. Since then, it is certainly the material which retained the most attention with early application to ozone monitoring [111] and demonstration of hyperspectral imaging [53].

Through the collaboration with Pr. V. Voloshinov from the Lomonosov Moscow State University (MSU), we could acquire practical knowledge with a KDP AOTF. The crystal itself appears to have suffered from metallic ion contamination during the growing phase. This caused a shift of the cut-off wavelength from 0.2 to 0.31 μm.

We obtained several encouraging results with this crystal. The least of them not being the acquisition of true spectral images or the measurement of the O_3 cross-section in the Huggins band (320-350 nm). However, we will restrict ourselves to the description of the main parameters of the filter, and to a finding which retained our attention: the influence of the temperature on the AOTF passband and its tuning curve.

4.4.1 Main AOTF parameters

Until this point, most of the encountered expressions and quantities were either generic to acousto-optics, either specific to TeO_2 filters. We will first quickly update some relationships to the KDP case, and then do the calculations for the particular AOTF under testing.

KDP is a negative ($n_o > n_e$) uniaxial crystal belonging to the tetragonal group $\bar{4}2m$ at room temperature. A particular mode of sound propagation in this group is the pure shear acoustic wave polarized along Y and propagating in the XZ plane[2] [5]. Depending on its angle α with the X axis, the sound velocity (v) is given in this configuration by eq.

[2] Unlike TeO_2 which usually needs to be rotated to deliver its best AO performances (interaction in the 110 plane, usually referred to as t-Z plane), here we can work with the usual axes denominations: X, Y and Z.

(4.55):

$$v(\alpha) = \left(\frac{c_{66}}{\rho} \cos^2 \alpha + \frac{c_{44}}{\rho} \sin^2 \alpha \right)^{1/2}, \tag{4.55}$$

where ρ is the KDP mass density (2338 kg m^{-3}), and c_{44} and c_{66} are two stiffness coefficients. Extreme values for the shear wave in the XZ plane are $v_X = 1.65 \times 10^3$ m s^{-1} and $v_Z = 2.34 \times 10^3$ m s^{-1}.

The elastic anisotropy induces non-collinearity of the phase and group velocities. Consistently with eq. (4.48), the walk-off angle is in this case given by:

$$\tan \Psi = \frac{(v_Z^2 - v_X^2) \sin 2\alpha}{2v^2}. \tag{4.56}$$

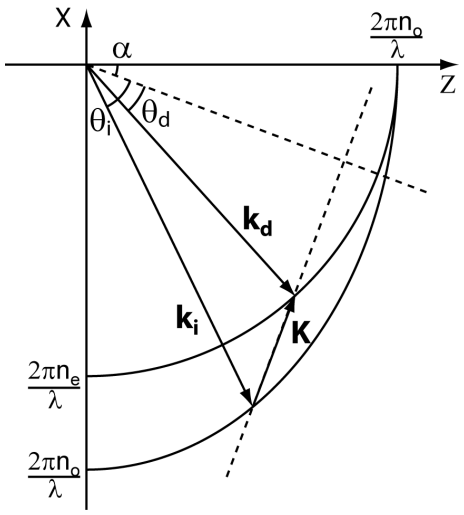

Figure 4.14 – Momentum vector diagram illustrating the birefringent AO interaction in KDP.

The wave vector diagram illustrating the non-critical AO phase matching in KDP is shown in Fig. 4.14, with the incident (o-light) and diffracted (e-light) optical wave vectors $\mathbf{k_i} = 2\pi n_o/\lambda$ and $\mathbf{k_d} = 2\pi n_d/\lambda$, and the acoustic wave vector $\mathbf{K} = 2\pi F/v$. The refractive index n_d of the diffracted beam follows eq. (4.10) with $\theta = \theta_d + \alpha$. The dispersion law for n_o and n_e at room temperature was measured down to 200 nm by Zernike [125] and fitted to a polynomial expression (eq. (4.57)-(4.58) with λ in μm). Results for the UV range are plotted in Fig. 4.15.

$$n_o^2(\lambda) = 2.259276 + \frac{0.01008956}{\lambda^2 - 0.0129426} + \frac{13.00522.\lambda^2}{\lambda^2 - 400}, \tag{4.57}$$

$$n_e^2(\lambda) = 2.132668 + \frac{0.008637494}{\lambda^2 - 0.0122810} + \frac{3.2279924.\lambda^2}{\lambda^2 - 400}. \tag{4.58}$$

Some trigonometric manipulations in Fig. 4.14 allow to derive the KDP

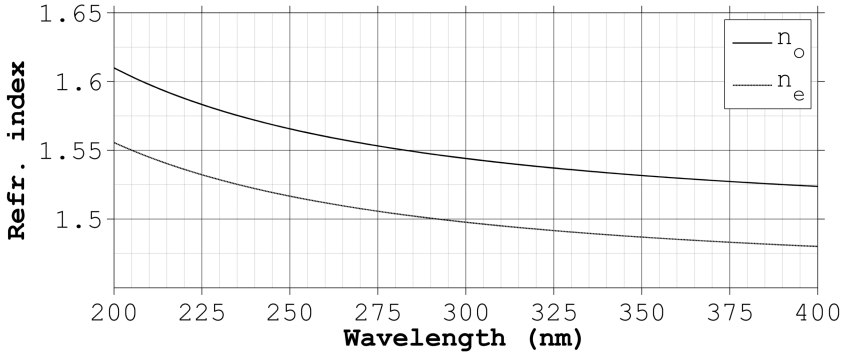

Figure 4.15 – Refractive indices for KDP at room temperature over the UV range. In this crystal, n_o is always larger than n_e.

AOTF tuning curve:

$$F = \frac{v(\alpha)}{\lambda}\left(\sqrt{n_o^2 - n_d^2\cos^2\theta_d} - n_d\sin\theta_d\right). \tag{4.59}$$

Conversion from θ_d to θ_i arise from the equality: $n_o\cos\theta_i = n_d\cos\theta_d$.

The last important difference between TeO_2 and KDP is the effective photo-elastic coefficient p which is provided here without the computation steps [110]:

$$p = p_{44}.\sin\alpha.\sin(\theta_d+\alpha) + p_{66}.\cos\alpha.\cos(\theta_d+\alpha), \tag{4.60}$$

with $p_{44} = -0.034$ and $p_{66} = -0.068$.

The KDP AOTF we tested was cut such that the acoustic wave vector is making an angle $\alpha = 9°$ with respect to the X axis. From the above expressions (eq. (4.55)-(4.59)), we computed the following parameters:

KDP AOTF, α= 9°					
θ_i	=	9.5°	$\Delta\theta$	=	1.1°
θ_d	=	8.4°	Ψ	=	8.7°

Knowing these values, the tuning curve of the filter can be drawn (Fig. 4.16).

The transducer has a length of 32 mm (L_T) and a width of 10 mm (H_T). With these parameters, the passband of the filter as approximated by eq. (4.50) can be determined. This is plotted in Fig. 4.17.

Finally, eq. (4.52) can be used to predict the required driving power allowing the diffraction efficiency to reach 100%. As the power at longer

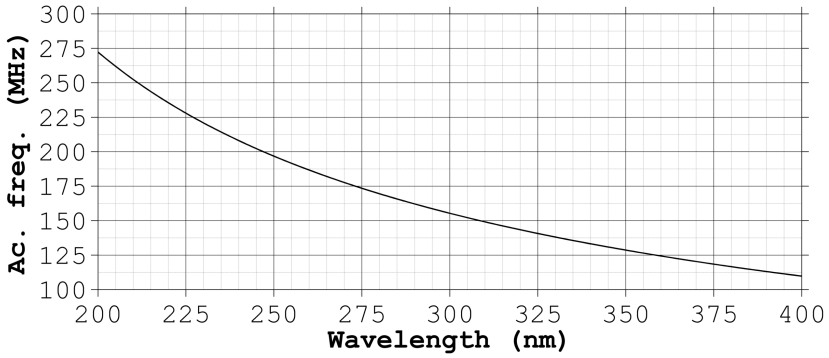

Figure 4.16 – Tuning curve for a KDP AOTF with cut-angle $\alpha = 9°$.

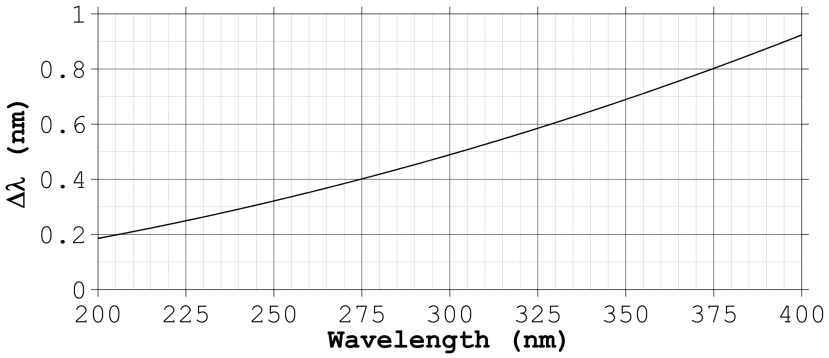

Figure 4.17 – Passband of a KDP AOTF with $\alpha = 9°$, $L_T = 32$ mm and $H_T = 10$ mm.

wavelengths can exceed reasonable values sustainable by the transducer (3-4 W), we included in Fig. 4.18 the power corresponding to a 50% DE.

This power estimated from eq. (4.52) is really related to the power density in the acoustic beam, neglecting all causes that can reduce the effective value in the medium, mainly acoustic attenuation and impedance mismatch. One quantity which is quite straightforward to measure is the 50 Ω electrical impedance mismatch at the transducer level. As the transducer must be driven over a wide range of frequencies, generally one octave, designing an electrical circuit offering a constant 50 Ω impedance over the entire range is very difficult. When there is a mismatch, part of the signal amplitude is reflected such that the transmitted power is lower.

Using calibrated electronic equipment, we measured the reflected-to-

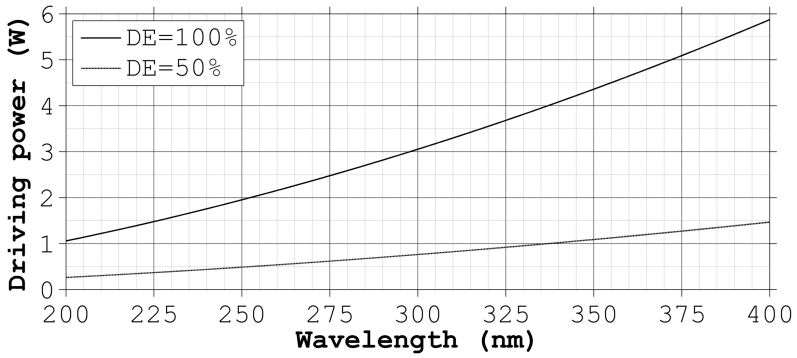

Figure 4.18 – Driving power needed to reach 100% or 50% of the DE at all wavelengths for a KDP AOTF with $\alpha = 9°$, $L_T = 32$ mm and $H_T = 10$ mm.

forward power ratio ($r = P_r/P_f$). Large values of these ratio indicate a bad impedance matching. As can be seen in Fig. 4.19, the region showing a reflection ratio smaller than 20% goes from 125 to 310 MHz. Associated to the tuning curve, this corresponds to all the wavelengths below 360 nm. Unfortunately, the transparency defect cuts all the wavelengths below 310 nm such that the optimal operational range of this AOTF is limited to a window of 50 nm. Longer wavelengths (shorter frequencies) operations are still feasible, but at the price of high power losses.

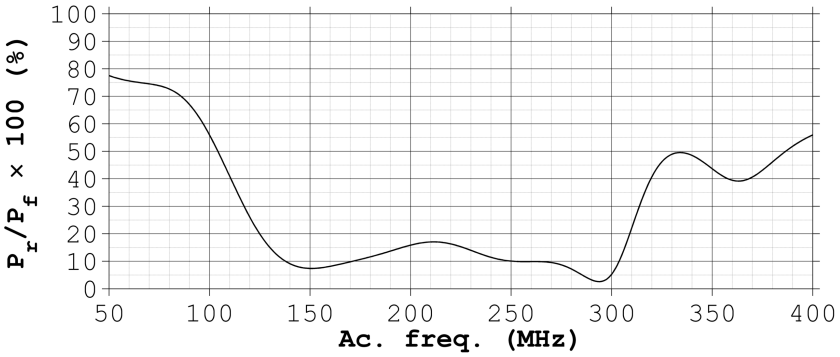

Figure 4.19 – Reflected-to-forward power ratio at the KDP AOTF transducer level.

From a practical point of view, we compensated these reflection losses by increasing the forward power: denoting the desired acoustic power by P_a, we have $P_f = P_a/(1 - r)$. As an example, from Fig. 4.18, we know

that at 325 nm, we need $P_a = 1$ W to reach a 50% DE. Accounting for the impedance mismatch, we will have to drive the transducer with a RF signal of 1.11 W.

4.4.2 Temperature dependence

The relatively high driving powers are the cause of an unexpected effect which was noticed during simple calibration tests. In order to verify the agreement with AO theory and to assess the DE of the KDP AOTF, we went to Pr. Wenseleers's laboratory at the Universiteit Antwerpen where a krypton laser was available. We used it to produce a monochromatic and vertically polarized light beam at the wavelength 350.74 nm (incident o-light). The AOTF was placed such as to have a perpendicular incidence of the light onto the crystal input face. In this configuration, acousto-optic theory predicts that Bragg matching will occur around 130 MHz (Fig. 4.16). Two photodiodes (Thorlabs S120VC) were placed 136.5 cm away from the back side of the AOTF in order to measure the zero and first orders of diffraction simultaneously. A temperature sensor was pasted on the upper face of the crystal close to the transducer side. Another one monitored the laboratory temperature.

In a test for verifying the power dependence of the DE, the transducer was driven by continous RF signals, setting the acoustic power to 0.5, 1, 1.5, 2, 2.5 and 3 W. After each power change, we let the temperature stabilize for more than 10 minutes before performing the measurements. These measurements consisted of sweeping the acoustic frequency (F) between 128 and 131 MHz by steps of 50 kHz. By monitoring the zero and first order intensity (I_0 and I_1), the shape of the AOTF spectral transfer function $T(F)$ can be drawn using:

$$T(F) = \frac{I_1(F)}{I_1(F) + I_0(F)}. \tag{4.61}$$

The measured transfer function for the 6 driving powers is plotted in Fig. 4.20 (top panel). The temperature measured by the sensor at the crystal surface after stabilization is plotted in function of the driving power in the bottom panel of the same figure.

Clearly, rising the driving power induces a shift of the overall transfer function towards smaller frequencies. A displacement of more than 1 MHz is observed when applying 3 W instead of 0.5 W. This is larger than the bandwidth itself. A second aspect concerns the peak DE which seems not to follow the increase of power: its amplitude levels off at 3 W to reach 27%. The shape of the transfer function is also affected:

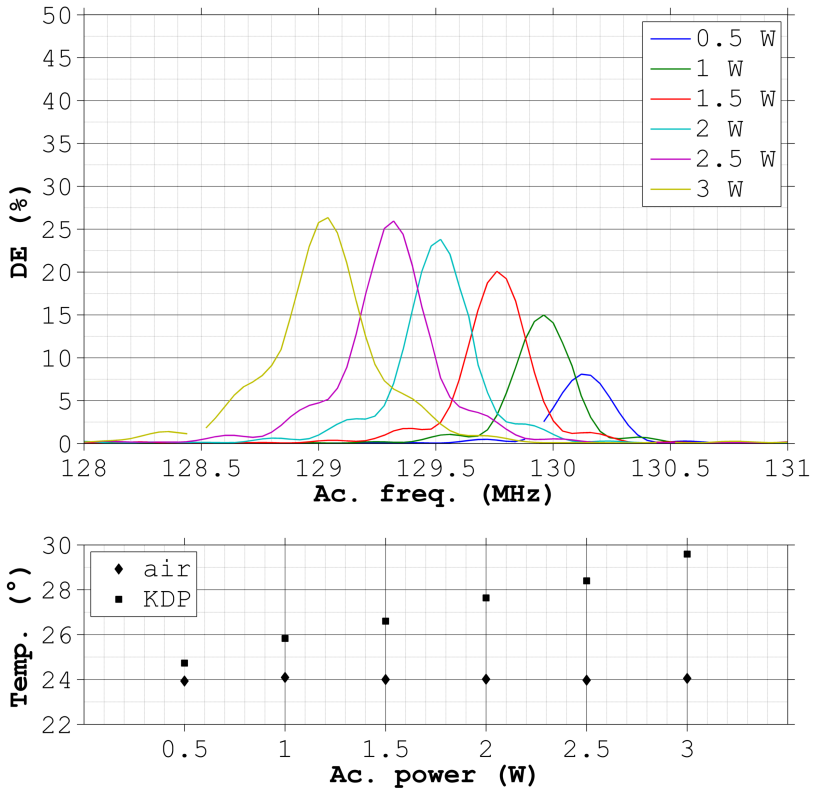

Figure 4.20 – Top: Measured DE of the KDP AOTF in function of F and for six values of P_a. Bottom: Measured temperature in function of the driving power in air and on the crystal surface, close to the transducer.

the higher the power, the more it deviates from the classical sinc^2 shape with the main lobe overlapping the side lobes. Finally, the temperature measured at the crystal surface exhibits a linear dependence on the driving power. The slope is approximately +1.85°C per watt.

In order to emphasize the influence of the crystal temperature on the transfer function, we modified the setup by coupling the signal generator to a pulse generator. The transducer was then driven by square pulses of 70 ms every second (corresponding to a duty cycle of 7%). In this configuration, the crystal was barely heated. We performed the same measurements, in this case synchronizing the photometers acquisition with the RF pulses. The observed transfer functions are plotted in Fig. 4.21. As can be seen, the changes observed previously are clearly minimized: the spectral shift is ten times less, the peak DE is increasing with the driving power to reach 42% at 3 W, and the sinc^2 shape is

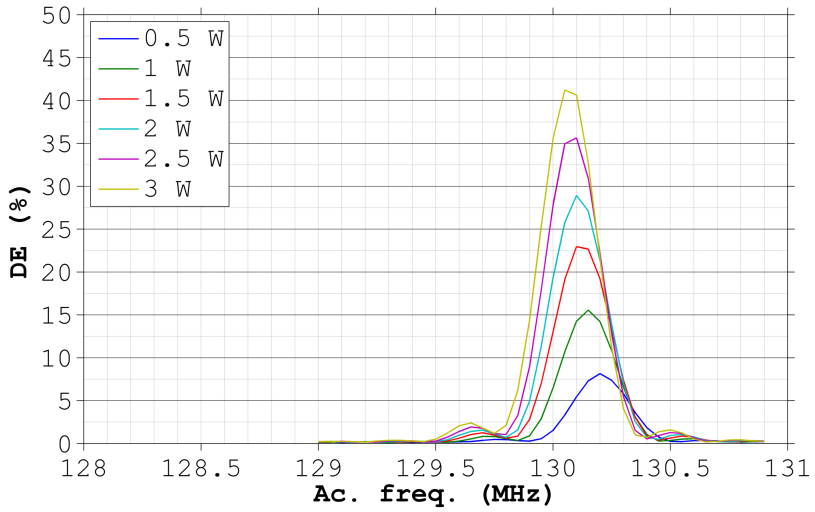

Figure 4.21 – AOTF spectral transfer function measured at different driving powers with pulsed signals.

preserved.

The consequences of continuously driving a KDP-based AOTF at powers up to several watts can be severe, especially for applications dealing with moderate to high spectral resolution (at the nanometer level and below). First, when turning the AOTF on, the matching frequency will shift towards smaller values until the thermal equilibrium is reached. Second, if not taken into account, this shift will result in important wavelength misregistration. For example, a difference of about -1.3 MHz was observed between both extrema. In the wavelength scale, this translates to an error of 2.9 nm at 350 nm. Considering that the bandwidth of the filter in pulsed regime (i.e. closer to theoretical predictions) is about 0.8 nm (see FWHM of the 3 W curve in Fig. 4.21), this effect should not be overlooked. Finally, the spectral resolution is degraded: the wings of the transfer function are taking too much importance with respect to the main lobe such that the FWHM is not a good measure of the bandwidth anymore. The diffracted order intensity will remain approximately identical — this can be seen by comparing the area below the curves in continuous and pulsed regimes — but more acoustic frequencies or more optical wavelengths will satisfy the Bragg condition, therefore reducing the selectivity of the AOTF.

A correlation between the shift of the tuning curve and the crystal temperature is obvious. The crystal gets heated by the acoustic wave

through energy transfer from the shear wave to the thermal phonons (Akhieser damping). A direct consequence is that the elastic and optical properties of the material change. By using temperature-dependent stiffness coefficients and refractive indices, it should be feasible to model the shift of the tuning curve with temperature. As a side product, we will have access to the inner temperature of the AOTF.

A few authors reported temperature-dependent stiffness coefficients in KDP [79], [26]. At temperatures sensibly higher than the Curie point (-150°C), all stiffness coefficients exhibit a linear dependence. Taking the data from the most cited paper [79], the following relations were obtained by a simple linear regression (stiffness expressed in 10^{10} Pa and temperature in °C):

$$c_{44}(T) = -5.0625 \times 10^{-4}.T + 1.2855, \quad (4.62)$$

$$c_{66}(T) = -4.3271 \times 10^{-4}.T + 0.6326. \quad (4.63)$$

Several authors have measured the refractive indices of KDP across its transparancy range at a single temperature ([125], [66]) or multiple ones ([108], [89]), and fitted them to a double-oscillator Sellmeier equation for most of them. Sometimes the dispersion of the temperature gradient $\mathrm{d}n/\mathrm{d}T$ was also studied [89] and [6]. One author [47] fitted data points from four published datasets to derive temperature-dependent Sellmeier coefficients. Assessing the quality of these fits with respect to the published measurement points, it was found that it actually is the best model especially at the shortest wavelengths where other fits can lead to errors larger than 10^{-2}. However, the agreement is slightly better for n_o than n_e. Their temperature-dependent Sellmeier equation is reproduced below (eq. (4.64) with the coefficients given in Table 4.3 (T in K, λ in μm)):

$$n^2(\lambda, T) = A(T) + \frac{B(T)}{1 - C(T)/\lambda^2} + \frac{D(T)}{1 - E/\lambda^2}. \quad (4.64)$$

The temperature-dependent stiffness coefficients and refractive indices were used in eq. (4.55) and (4.10), and the shift of the Bragg matching frequency for diffraction of light at 350.74 nm was found to be $\mathrm{d}F/\mathrm{d}T = -0.104$ MHz K^{-1} which corresponds to a shift of the central wavelength of +0.23 nm K^{-1}. Our experiment showed that the shift in frequency in function of acoustic power was $\mathrm{d}F/\mathrm{d}P = -0.437$ MHz W^{-1}. Together, these two slopes lead to the heat generation rate of the KDP AOTF under the acoustic strain: $\mathrm{d}T/\mathrm{d}P = 4.2$ K W^{-1}. It is a measure of the heating of the crystal by the acoustic wave in the

	n_o	n_e
$A(T)$	$31.85 \times T + 1.44896$	$-11.52 \times T + 1.42691$
$B(T)$	$-141.14 \times T + 0.84181$	$-61.39 \times T + 0.72722$
$C(T)$	$-2130 \times T + 0.01280$	$3104 \times T + 0.01213$
$D(T)$	$57.5 \times T + 0.90793$	$-19.8 \times T + 0.22543$
E	30	30

Table 4.3 – Temperature-dependent Sellmeier coefficients for KDP refractive indices. (Reproduced from [47])

steady state thermal conduction regime. Compared to the measured heating rate at crystal surface (1.85 K W^{-1}), the increase of temperature inside the acoustic column as a function of the driving power is doubled, i.e. at 3 W, the temperature increased by 6 K at the surface, whereas it raised by 13 K in the AO interaction region, according to the temperature-dependent model.

The measurement of the tuning curve displacement therefore constitutes an indirect means of temperature determination at the heart of the medium. Without this knowledge, using surface temperature and thermal conductivity coefficients in complicated thermal analysis would be mandatory in order to estimate the inner temperature.

Chapter 5

AOTF-based hyperspectral imager — Application to the remote sensing of NO_2 in industrial smokes

In the previous chapters, we have discussed the advantages of an AOTF-based hyperspectral imager for the remote sensing of atmospheric species from a space borne instrument. We also described the working principle of an AOTF and illustrated some aspects such as the spectral resolution or the thermal effects with laboratory experiments.

An important step in the ALTIUS project has been to manufacture an optical breadboard of the VIS channel using commercial elements, including the TeO_2 AOTF which was purchased from Gooch & Housego (see Chapter 4). This breadboard was fully characterized by the OIP company in a laboratory environment. We brought it to the next level by proposing a field campaign in which its remote sensing capabilities would be checked against *actual* measurements.

To our knowledge, it was the first time an AOTF-based spectral imager was used to detect the presence of a weak absorber in outdoor conditions. In line with the general ALTIUS concept, the imaging capability played here an important role. The detection of NO_2 in turbulent exhaust smoke at a distance of several kilometers could only be achieved thanks to the excellent optical and spectral performances of the breadboard. While the instrument managed to highlight the presence of NO_2 in a changing scene, the quantification of its concentration turned out not to be straightforward, especially with the difficult environmental conditions (wind, clouds, pollution, distance,...).

In the following, we provide a comprehensive description of the breadboard itself and the outdoor experiment. The data analysis reveals all the difficulties to retrieve the NO_2 content in this apparently simple viewing geometry. The approach developed to subtract the background

signal could be useful in future similar campaigns.

This work has been published in Applied Optics in 2012, under the title *Tunable acousto-optic spectral imager for atmospheric composition measurements in the visible spectral domain*, [36].

5.1 Instrument description

The breadboard is a simplified version of the visible channel of the ALTIUS instrument as designed during the feasibility study of the project. It contains commercial refractive optical parts instead of mirrors, a detector, and a TeO_2 AOTF (for more details on AOTF principle, please refer to Chapter 4). It is designed to simulate ALTIUS measurements in the bright limb observation mode. The characteristics of the breadboard are shown in Table 5.1.

Parameter	**Value**	**Parameter**	**Value**
Entrance pupil (mm)	8.73	Spectral domain (nm)	400-800
Total focal length (mm)	125	AOTF aperture (mm)	10×10
FEO focal length (mm)	100	AOTF acceptance angle (°)	2.4
BEO magnification	1.25	Spectral resolution (nm)	0.6-3.5
Field of view (°)	5.73×5.73	Detector pixels array	512×512
F/#	14.32	Pixel size (μm)	24×24

Table 5.1 – Visible channel breadboard characteristics.

The breadboard consists of a front end (FEO) and back end optics (BEO), as drawn in Fig. 5.1. The FEO is designed as a telecentric confocal system, optimized for high uniformity. The BEO is a relay system which images the AOTF selected order onto the detector. The large f-number of the breadboard ensures a large depth of focus at the detector focal plane position. The detector is a commercial PIXIS 512B from Princeton Instruments including a back illuminated CCD array, sensitive from the UV to the NIR spectral range and cooled down to -70°C. The Glan-Taylor polarizers in the FEO and BEO are cross-oriented to achieve highest extinction ratios of the rejected spectrum (because an AOTF rotates the polarization of the diffracted light by 90°). Their use in combination with the BEO pupil ensures efficient stray light reduction.

The instrument opto-mechanical assembly is composed of an optical

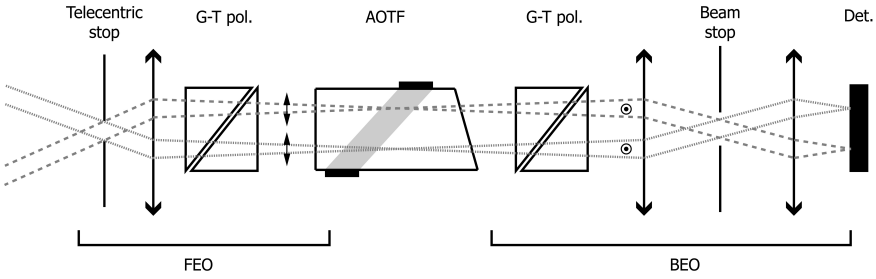

Figure 5.1 – Breadboard optical design: the front end optics (FEO) is composed of a telecentric stop, a lens doublet, a calcite Glan-Taylor polarizer, and the TeO_2 AOTF; the back end optics (BEO) is composed of a Glan-Taylor polarizer, a lens doublet, an aperture stop, a second lens doublet and the pixel array. Polarization state of the incident beam and the selected order is represented after each polarizer. The acoustic wave propagating through the crystal is represented by the pale gray region. Neither the relative dimensions of the optical elements nor the focal lengths are representative of the real dimensions.

baseplate (manufactured at BIRA-IASB) and holders made from black anodized aluminum. On top of this structure commercial Thorlabs assemblies are mounted for integration and alignment of the optics. Provisions to tilt lenses and to align the assembly in X-, Y- and Z-directions allow for compensation of possible misalignments for optimum optical performance. The detector is mounted on a bracket that can be moved in the Z-direction (height) by using adjusting screws. This adjustable bracket is mounted on two motorized translation stages for X- and Y-movement. This assembly ensures accurate positioning of the detector CCD and allows for the correction of the chromatic aberrations within the 400-800 nm operational spectral range.

The AOTF and the detector are controlled by LabView software. It allows for a wide range of different acquisition sequences and saves the images together with the corresponding housekeeping data.

The AOTF we used was a Gooch & Housego model number TF625-350-2-12-BR1A driven by a 16-channel RF driver R64060-150-10DFS-16X1-CDI3. This AOTF was compatible with the requirements of a spectral imaging system; in particular, the aperture was large (10×10 mm) so as to maximize the light throughput. The design has been described elsewhere [115, 99]. From laboratory measurements, it was found that a diffraction efficiency (DE) close to 100% could be reached over the full wavelength range. Using laser sources available in the laboratory of nanoscopic physics (NAPS) in the Université Catholique de Louvain (Louvain-la-Neuve, Belgium), we measured a spectral bandwidth rang-

ing from 0.65 nm at 457.9 nm to 3.5 nm at 800 nm, in agreement with theoretical predictions for a filter designed with a cut-angle of 7.65°.

5.2 Smokestack experiment

During the 4th week of January 2012, an airborne campaign in Toulouse (France) was prepared in collaboration with the Service des Avions Français Instrumentés pour la Recherche en Environnement (SAFIRE). The measurement campaign consisted of a 3 hours flight during which limb scattered light and a solar occultation could be observed. Unfortunately, those measurements were not conclusive as fog appeared on the aircraft window very early during the ascent. However, on the 24th of January (the day before the flight), the airplane inside which the breadboard was mounted was driven outside of its hangar for testing. Fortunately, the Toulouse's waste incinerator smokestack can be seen from the tarmac of the Francazal airport at a distance of about 3.5 km. By orienting the airplane in such a way that the breadboard could observe the smokestack, spectral snapshots of the smoke plume could be taken. It was an unexpected test for our instrument with rather limited control on the experimental setup but it was definitely an excellent opportunity to demonstrate the capabilities of a fast AOTF-based spectral imager.

5.2.1 Experimental conditions

The instrument was mounted on a rack fixed in the airplane cabin in front of a quartz window through which the scene can be observed. The Francazal airport is located at 43.5413°N, 1.3616°E and the smokestack is located at 43.5569°N, 1.3991°E. The distance between the instrument and the stack is about 3485 m and the mean line of sight azimuth angle with respect to North is 60°. At this distance, the measured instrument FOV of 0.097×0.097 rad^2 embraces a scene of 338×338 m^2 (0.44×0.44 m^2 per pixel).

Fig. 5.2 shows an image taken by the breadboard with the AOTF tuned to the central wavelength of 645 nm (offering better visibility than shorter wavelengths). Due to the telecentric optical design, all perspective effects have disappeared, which gives the impression that the stack belongs to the building in the foreground, which is actually much closer than the incinerator. The image was taken at 10:20:54 (UT) with an integration time of 1 second. At that time, the sky was partly cloudy with moderate wind (estimated to about 10 km/h from the smoke

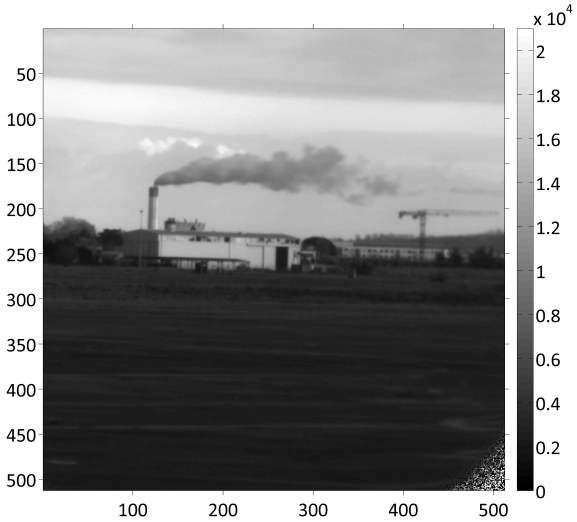

Figure 5.2 – Image taken by the breadboard with the AOTF tuned at the central wavelength of 645 nm. The grayscale indicates the number of counts per pixel. The integration time was 1 second. Straylight has been removed and the image was corrected for flat field non-uniformities which is the reason why the lower right corner is noisy (region of rather poor diffraction efficiency). Axes indicate pixel number.

displacement in successive pictures). Contrary to the impression given by this grayscale picture, the smoke actually appeared white, mainly because the cleaning processes of combustion smokes involve water vapor flushing.

The experimental conditions were not optimal for the measurement of NO_2 concentrations by means of backscattered sunlight: a partly cloudy sky prevented us from taking reference measurements at zenith and the plume was dense, contained lots of different gases and particles, and was blown by a moderate wind.

By comparison, the experimental conditions described in the reference paper by Lohberger [74] about imaging differential optical absorption spectroscopy (DOAS) were more favorable : they observed the exhaust gases of a gas cogeneration plant with a perfectly blue sky, in the absence of wind, and whose smokes were transparent. It was actually the only possible experimental condition for them because the imaging system used in their study was a scanning spectrometer which took several minutes to complete the picture. On the other side, they managed to measure accurately the NO_2 concentration at the outlet of the stack with uncertainties of about 10% only (but neglecting temperature effects

on the differential NO_2 absorption cross-section).

5.2.2 Measurement principle

This instrument is designed to capture fast snapshots with moderate spectral resolution. It is not designed to sample large spectral windows the way other DOAS instruments do because it would take too much time, especially when observing moving targets. We therefore chose to apply the simplest measurement principle: by quickly taking pictures at two well-chosen wavelengths (a strongly absorbing one (λ_s) and a weakly absorbing one (λ_w)), one can emphasize the presence of a weak absorber in the light path by detecting a variation of light intensity as long as other species do not spectrally interfere.

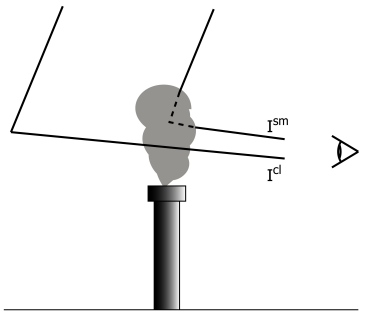

Figure 5.3 – Illustration of two photons trajectories. One is backscattered by the plume, the other one by the neutral air density (and passes close to the plume but not through). A simple single-scattering process is assumed. One notices the difference in path length in the boundary layer.

Fig. 5.3 illustrates the radiative transfer problem and the assumptions made in the frame of this simple test case: the photons reaching the instrument have crossed the entire atmosphere to reach the boundary layer where they are scattered in our FOV. Some of them have penetrated the plume, others have been scattered far behind the stack. The main difference is that the former have experienced strong extinction by the smoke, while the latter have traveled a larger distance in the boundary layer (if we assume single scattering approximation). Describing the extinction of light intensity by a Beer-Lambert law (see section 1.3), one can write for the *smoke* pixels:

$$I^{\mathrm{sm}}(\lambda) = I_0(\lambda).\exp\left(-\tau^{\mathrm{atm}}_{\mathrm{NO_2}}(\lambda) - \tau^{\mathrm{sm}}_{\mathrm{NO_2}}(\lambda) - \tau_{\mathrm{other}}(\lambda)\right), \qquad (5.1)$$

and for the *clear air* pixels :

$$I^{\mathrm{cl}}(\lambda) = I_0(\lambda).\exp\left(-\tau^{\mathrm{atm}}_{\mathrm{NO_2}}(\lambda) - \tau^{\mathrm{trop}}_{\mathrm{NO_2}}(\lambda) - \tau_{\mathrm{other}}(\lambda)\right), \qquad (5.2)$$

where I_0 is the solar spectral radiance out of the atmosphere, τ_{NO_2} is the optical thickness measuring the extinction of light by NO_2 molecules along its path ($\tau^{atm}_{NO_2}$ denotes the atmospheric path common to the clear and smoke pixels, $\tau^{trop}_{NO_2}$ relates to the additional traveled distance in the boundary layer for clear pixels, and $\tau^{sm}_{NO_2}$ is for the smoke path), and τ_{other} is the optical thickness of other atmospheric extinction processes (f.i. Rayleigh and Mie scattering).

The choice of the two wavelengths is critical for the method. By choosing them close, the difference in aerosol extinction can be neglected. Moreover, as other absorbing trace gases are present (with possibly comparable or larger extinction capability), the selected spectral range should be kept away from differential structures in their spectrum. By doing so, one can estimate the extinction due to NO_2 within the smoke by computing the following ratios:
for the smoke pixels,

$$\begin{aligned} \ln \frac{I^{sm}(\lambda_w)}{I^{sm}(\lambda_s)} &= \ln \frac{I_0(\lambda_w)}{I_0(\lambda_s)} + \left(\tau^{atm}_{NO_2}(\lambda_s) - \tau^{atm}_{NO_2}(\lambda_w)\right) \\ &\quad + \left(\tau^{sm}_{NO_2}(\lambda_s) - \tau^{sm}_{NO_2}(\lambda_w)\right) + \left(\tau_{other}(\lambda_s) - \tau_{other}(\lambda_w)\right), \end{aligned} \tag{5.3}$$

and for the clear air pixels,

$$\begin{aligned} \ln \frac{I^{cl}(\lambda_w)}{I^{cl}(\lambda_s)} &= \ln \frac{I_0(\lambda_w)}{I_0(\lambda_s)} + \left(\tau^{atm}_{NO_2}(\lambda_s) - \tau^{atm}_{NO_2}(\lambda_w)\right) \\ &\quad + \left(\tau^{trop}_{NO_2}(\lambda_s) - \tau^{trop}_{NO_2}(\lambda_w)\right) + \left(\tau_{other}(\lambda_s) - \tau_{other}(\lambda_w)\right). \end{aligned} \tag{5.4}$$

The final term is assumed to vanish in both equations since $\lambda_s \approx \lambda_w$. The extinction of light by NO_2 within the smoke follows from replacing eq. (5.4) in eq. (5.3) to find:

$$\tau^{sm}_{NO_2}(\lambda_s) - \tau^{sm}_{NO_2}(\lambda_w) = \ln \frac{I^{sm}(\lambda_w)}{I^{sm}(\lambda_s)} - \ln \frac{I^{cl}(\lambda_w)}{I^{cl}(\lambda_s)} + \left(\tau^{trop}_{NO_2}(\lambda_s) - \tau^{trop}_{NO_2}(\lambda_w)\right). \tag{5.5}$$

By definition, the optical thickness τ_{NO_2} is the product of the absorption cross section of NO_2 (σ_{NO_2}) by the slant column density (SCD_{NO_2}) which is the integrated gas concentration along the light path : $\tau_{NO_2}(\lambda) = \sigma_{NO_2}(\lambda).SCD_{NO_2}$. The NO_2 smoke SCD is then computed by eq. (5.6):

$$\begin{aligned} SCD^{sm}_{NO_2} &= \frac{1}{\sigma_{NO_2}(\lambda_s) - \sigma_{NO_2}(\lambda_w)} \times \\ &\quad \left[\ln \frac{I^{sm}(\lambda_w) I^{cl}(\lambda_s)}{I^{sm}(\lambda_s) I^{cl}(\lambda_w)} + \left(\tau^{trop}_{NO_2}(\lambda_s) - \tau^{trop}_{NO_2}(\lambda_w)\right)\right]. \end{aligned} \tag{5.6}$$

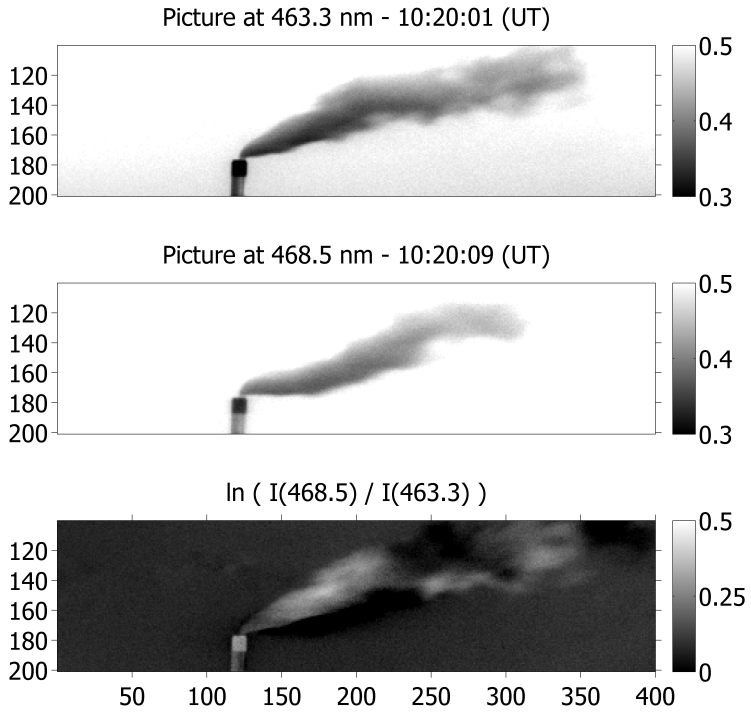

Figure 5.4 – The first two images (cropped from the initial ones) capture the plume at 463.3 nm and 468.5 nm. The grayscale units are arbitrary. The last image shows $\ln(I(468.5)/I(463.3))$ where one can observe that most of the signal in the main body of the plume and its tail comes from the poor correspondence of the two plume images. Clearly, as the plume has moved down in the second picture, the ratio is sometimes computed between a brighter background pixel and a plume pixel, instead of two plume pixels. Only the very beginning of the plume is consistent from one image to the other.

It should be noted that while all the collected photons have crossed the entire atmosphere, the upper air masses (higher than the boundary layer) do not play any role in this measurement method. In particular, the effect of the varying temperature of these air masses on the absorption cross-section does not have to be considered.

5.2.3 Data acquisition and correction

A standard doublet of wavelengths for NO_2 measurements are 448 nm and 453 nm (see discussion in section 3.7). Unfortunately, the instrument could not be calibrated down to these wavelengths and another

pair at $\lambda_s = 463.3$ nm and $\lambda_w = 468.5$ nm was selected instead. They have the advantage of being close ($\Delta\lambda = 5.2$ nm), corresponding to a negligibly different O_3 absorption cross section amplitude, and benefiting from a narrow 0.65 nm FWHM AOTF spectral response function.

Between 10:12 AM and 10:20 AM, 6 sequences of two pictures were acquired with an integration time of 2 seconds. The wavelength was tuned to 463.3 nm and 468.5 nm alternatively with an average dwell time of 6 seconds between each picture, with a partial change of the shape of the plume during this interval. Among the 6 sequences, the last one shows the best overlap of the plumes, as illustrated in Fig. 5.4. Stray light and flat field measurements were also performed during and after the acquisition sequence.

Flat field. The smokestack pictures acquisition was followed by flat field measurements. The technique is based on imaging uniform extended light source at each target wavelength in the same environmental conditions (temperature, humidity, ...) as the measurements. The acquired picture forms the instrument imaging transfer function which includes the effects of the optics aberrations, the non-constant DE over the AOTF aperture and the detector pixel response non-uniformity (PRNU). Any picture ($P(\lambda)$) can then be corrected by performing a pixel-to-pixel division by the normalized flat field response ($F(\lambda)$).

The spatially uniform light source was provided by the cloudy sky. With a sufficiently large number of pictures of moving clouds in the FOV, one can get as close as desired to a perfect flat scene. In our case, the cloud coverage was almost 100% with a moderate wind. 20 consecutive pictures were averaged to construct the true flat field.

As the retrieval method uses ratios of spectral pictures (see eq. 5.6), particular attention was paid to not introducing any bias along with the subsequent ratio of flat fields. In absence of coïncident calibrated measurements of the zenith cloudy sky radiance, radiative transfer simulations (using MODTRAN 5, [7]) provided the missing information. It was found that the radiances at 463.3 nm and 468.5 nm are almost equal for similar conditions (urban pollution, stratus clouds, same time, date and location, convolved by the instrumental spectral response function), requiring no further correction of the computed ratios. Though, a 1% 1σ standard deviation on the flat field ratio correction factor is considered in the total error budget.

Stray light. Due to the combined effect of an imperfect cross-orientation of the two polarizers surrounding the AOTF and the scattering of

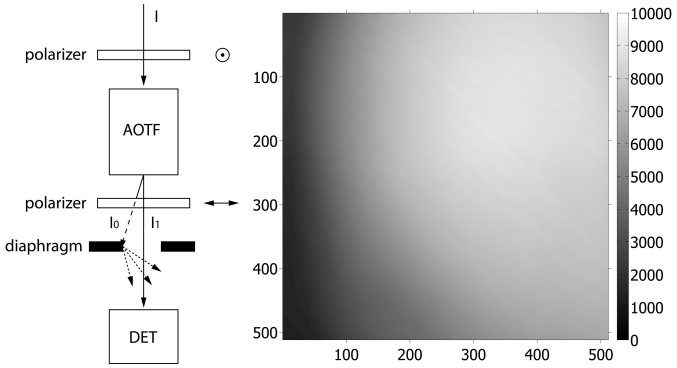

Figure 5.5 – Left drawing shows a part of the optical design as seen from above. Stray light is generated by the scattering of the unwanted order (I_0) by the diaphragm. The right picture is an image acquired when observing in the direction of the smokestack with the AOTF turned off. Integration time was 2 seconds.

the residual undiffracted beam by the BEO diaphragm, pictures are contaminated by a non-negligible amount of stray light. In order to remove its contribution to the total intensity, some pictures must be acquired with the AOTF turned off (P^{OFF}). By doing so, nothing else than residual undiffracted photons can reach the detector, which happens also when the AOTF is turned on but superimposed to the image made from the diffracted beam. Fig. 5.5 shows how residual undiffracted photons contaminate the picture, together with an image taken with the AOTF turned off.

Altogether, the image correction follows a 2-step procedure: first, the stray light is removed by subtraction (which is also applied to the flat field pictures), then normalized by the mean flat field:

$$P^{\mathrm{corr}}(\lambda) = \left(P^{\mathrm{ON}}(\lambda) - P^{\mathrm{OFF}}\right) / \left(\overline{F}^{\mathrm{ON}}(\lambda) - \overline{F}^{\mathrm{OFF}}\right) \tag{5.7}$$

5.2.4 Results

In order to solve eq. 5.6 and to find the NO_2 smoke SCD, $I^{\mathrm{cl}}(\lambda)$ and $\tau^{\mathrm{trop}}_{\mathrm{NO_2}}(\lambda)$ must be determined. From the spectral convolution of the NO_2 absorption cross-section with the instrument response function, we have $\sigma_{\mathrm{NO_2}}(463.3) = 5.45 \pm 0.22 \times 10^{-19}$ cm^2 molecule^{-1} and $\sigma_{\mathrm{NO_2}}(468.5) = 3.49 \pm 0.14 \times 10^{-19}$ cm^2 molecule^{-1} (data from Vandaele et al. at 294 K, [107], neglecting the small uncertainties due to the different temperatures of the local air masses).

As can be observed in the third image of Fig. 5.4, the quantity $\ln(I(468.5)/I(463.3))$ is almost constant and close to 0 for the air imaged by the pixels looking at the left side of the stack (the *clear air* pixels). By computing the natural logarithm of the average ratio of intensities registered between rows 100 and 200 and columns 50 and 100, we find $\ln(I^{\mathrm{cl}}(468.5)/I^{\mathrm{cl}}(463.3)) = 0.0783 \pm 0.0010$ (which seems consistent with the presence of a background atmospheric NO_2 absorption in a polluted region).

The last term of eq. 5.6 corrects the *smoke* and *clear air* intensity ratio for their difference of light path length in the boundary layer. In our case, $\tau^{\mathrm{trop}}_{\mathrm{NO_2}}(463.3)$ and $\tau^{\mathrm{trop}}_{\mathrm{NO_2}}(468.5)$ can only be determined indirectly. By definition, $\tau_{\mathrm{NO_2}}(\lambda) = \sigma_{\mathrm{NO_2}}(\lambda).n_{\mathrm{NO_2}}.L$, where $n_{\mathrm{NO_2}}$ is the NO_2 concentration in the air and L is the optical path length. Information on the hourly concentrations of pollution-related species like NO_2 could be obtained from the Observatoire Régional de l'Air en Midi-Pyrénées (ORAMIP[1]) which is the official institution responsible for the air quality monitoring in the region of Toulouse. From their in-situ measurements in different places of the city, an average NO_2 mass concentration of $55 \pm 7 \mu$g m^{-3} at this time of the day could be estimated. The fact that the first 15 km of the instrument line of sight cover the city area is an argument for using urban measurement stations only.

The estimation of L is less straightforward. We assume that L is the visual range (visibility) as defined by the Koschmieder formula [68]: $\mathrm{vis}_{550} = 3.912/\beta(550)$, where $\beta(\lambda)$ is the light extinction coefficient at a given wavelength in the air mass, mainly due to aerosols. A number of studies have shown significant correlation between particulate matter (PM_{10} or $PM_{2.5}$, particles whose diameter is smaller than 10 or 2.5 μm) and the total aerosol optical thickness (AOT, represented by $\tau_{\mathrm{a}}(\lambda) = \beta(\lambda).H$, where H is the light path length). Chu et al. [27] derived a relation between PM_{10} concentrations and the AOT at 550 nm from co-registered ground-based AERONET data (AErosol RObotic NETwork, [58]) : $PM_{10} = 54.7 \times \tau_{\mathrm{a}} + 8$, ($R^2 = 0.82$).

The ORAMIP data show a PM_{10} mass concentration of $22 \pm 1 \mu$g m^{-3} on average at the time of our measurements, hence yielding $\tau_{\mathrm{a}}(550) = 0.26 \pm 0.07$. Assuming that most of the aerosols observed by the AERONET Sun photometer are found in the boundary layer (box distribution profile), and according to the NOAA meteorological model READY (Real-time Environmental Applications and Display sYstem[2]) which gives a boundary layer thickness of $H = 700$ m at the time and

[1] http://www.oramip.org

[2] http://ready.arl.noaa.gov

place of our measurements, we find a mean aerosol extinction coefficient at 550 nm in the boundary layer: $\beta(550) = \tau_a(550)/H = 0.37 \pm 0.09$ km^{-1}.

Using the Angström law to derive $\beta(466)$ (mean measurements wavelength) with a mean Angström coefficient of 1.55 ± 0.34 (taken from measurements by Mélin and Zibordi [84] in similar pollution conditions in Ispra, Italy), we find $\beta(466) = 0.37.(466/550)^{-1.55} = 0.48 \pm 0.12$ km^{-1}.

From these values, we estimate a visual range of 8.2 ± 2.1 km at our wavelengths. Subtracting the 3.5 km distance between the instrument and the stack from the visual range, we finally obtain : $\tau^{\text{trop}}_{NO_2}(463.3) - \tau^{\text{trop}}_{NO_2}(468.5) = 0.067 \pm 0.033$.

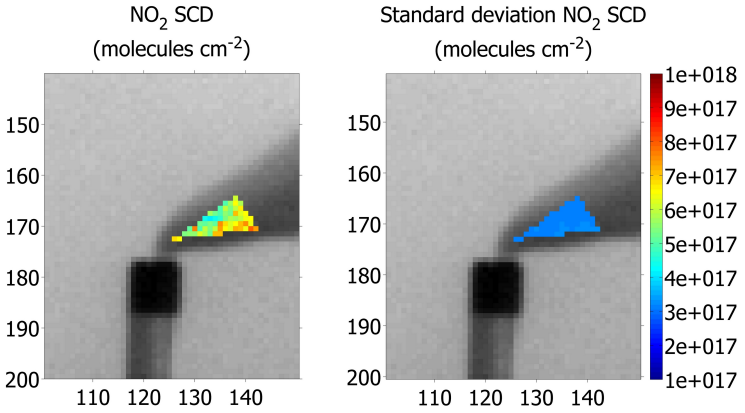

Figure 5.6 – Left panel : NO_2 slant column density at the beginning of the plume, as computed with eq. 5.6 for individual pixels. Right panel : standard deviation of the computed SCD value. The large uncertainty at the pixel level arises from a relatively low signal to noise ratio of the pictures and, to a lesser extent, to uncertainties on the optical thickness in the boundary layer.

Eq. 5.6 can now be solved and the NO_2 SCD in the plume close to the stack outlet is represented by a colormap superimposed to the initial grayscale picture in Fig. 5.6. The location of the small subset of selected pixels for the SCD computation corresponds to the region where the plume remains stable within the 6 second interval (see Fig. 5.4). The standard deviation on the SCD values at the pixel level are shown in the right panel of Fig. 5.6. The relatively large uncertainty on the SCD is due to the somewhat reduced signal to noise ratio of the initial pictures (a consequence of the small integration time) and to the uncertainty on the optical thickness in the boundary layer (last term

of eq. 5.6). However, observing a bell-shaped distribution of the SCD values, we can decrease the uncertainty by computing a mean value for the SCD at the beginning of the plume : $6.0 \pm 0.4 \times 10^{17}$ molecules cm^{-2}.

This value does not take into account the effect of the temperature on the differential absorption cross-section of NO_2 ([55], [106]). From computations on different datasets ([107], [21]) convolved by our instrument spectral transfer function, we estimated a relative effect of -0.3% K^{-1} in the cross-section amplitude difference between 463.3 nm and 468.5 nm. However, no valuable information could be found on the temperature of the plume released in the atmosphere by such incinerators. A few tens of degrees above the ambiant temperature can quickly lead to an underestimation of more than 10% in the SCD. This aspect was eluded in Lohberger's work and in absence of reliable information we cannot do much more than mention it.

In order to get an insight of the mass concentration of NO_2 within the plume, the optical path length inside the smoke must be determined. The large optical thickness of this white plume is in favor of an accurate treatment involving multiple scattering. Though, as the accurate determination of NO_2 emissions of such incinerator is not the target of this work, we obtained an estimated value by making two rough assumptions. The first one is to suppose an effective optical path length equal to twice the radius of the disk section of a conical envelope of the plume (i.e. photons undergo a single scattering at the core of the plume). The second one is to assume a homogeneous distribution of the NO_2 molecules within the plume conical volume. In the middle of the triangle-shaped subset of selected pixels (see Fig. 5.6), the diameter of the cone is approximately 9 pixels, equivalent to a length of 4 m. The concentration of NO_2 at the beginning of the plume would then be about 115 mg m^{-3}. This estimation, although a rough one, is at least in line with the official threshold[3] of 200 mg m^{-3}, and with the emission rates published by the incinerator administrator (Veolia): for the 24^{th} of January, 151.8 mg m^{-3} of NOx (= NO+NO_2) were emitted on average.

[3]Journal officiel de la République française, Lois et Décrets, 1er décembre 2002, N°280, 19778-19789, NOR : DEVP0210351A.

Conclusion and future research

In the beginning of the years 2000's, there were plenty of atmospheric sounders orbiting around the Earth, several of them collecting high vertical resolution data for the retrieval of the concentration profile of key atmospheric trace gases. Among other measurement techniques, the limb scattering geometry offers a dense spatio-temporal sampling and some sensitivity to important minor constituents such as ozone, nitrogen dioxide, aerosols, water vapor, methane and other UV-VIS-NIR absorbers. Missions like OSIRIS, SCIAMACHY or even SAGE III have met great successes, and to the present day, OSIRIS and the recently-launched OMPS-LP are still providing limb-scatter spectra.

As it has been shown in recent dataset reprocessing exercises, the disagreement among the instruments is large enough at tropical and high latitudes to prevent any detection of a statistically significant positive trend in ozone profiles, definitely not below a threshold of 5% per decade. Yet, reducing the uncertainties on this essential climate variable is urgent as the atmospheric modelling community is predicting a recovery of the ozone layer above Antarctica at a rate of 2% per decade.

Regarding the limb scattering technique, one of the most important sources of uncertainty is the tangent height registration. Under sunlight, the Earth atmospheric limb is a diffuse light source which has no more distinct feature than an exponentially decreasing radiance profile. Past and current space experiments relying on that geometry have used grating spectrometers to record limb radiance spectra. But the one-dimensional nature of their acquisitions, combined with the platform attitude uncertainties and, for the instruments designed with a horizontal slit, a scanning system inaccuracy, often lead to tangent height misregistration of the order of half a kilometer. Knowing that for some part of the ozone profile, a mistake of 500 meters yields an error of 20% in the retrieved concentration, one can understand the need for a better

pointing knowledge.

This is the context which gave birth to the ALTIUS mission: a spaceborne instrument mainly designed for the acquisition of spectral images of the bright limb, but also capable of solar, lunar, stellar and planetary occultations. The spectral imaging feature plays here the central role: by taking 2-D snapshots, the pointing uncertainty will be minimized as the full scene is acquired at once (no need for scanning systems, background features like stars helping for pointing calibration). The overall concept includes the use of an AOTF responsible for the selection of the spectral information.

Part of this work has been dedicated to the demonstration that the instrumental concept is suitable for the measurement at high vertical resolution of concentration profiles. In Chapter 3, we have presented some simulations of the performance of ALTIUS in retrieving two molecules: ozone and nitrogen dioxide profiles. This mandatory exercise for any new experiment required the use of a radiative transfer code, an instrument simulator including its tolerances on the spectral, pointing and radiometric capabilities, and an inversion algorithm. This work enabled the comparison with other missions and the identification of possible limitations (or sometimes excessively strong requirements) of the current design. During previous phases of the project, more species have been studied, but we chose to show the two most important ones.

Good simulation results cannot offer as much confidence in a concept as experimentation. Quite early in the project, an optical breadboard of the visible channel was manufactured. While its first goal was to assess the quality of an AOTF-based hyperspectral imager built from commercial pieces, most interesting things were accomplished outside of the clean room. With this instrument, we managed to reproduce reference measurements of the nitrogen dioxide absorption cross-section. In Chapter 4 we described the setup and how the AOTF was operated in a very unconventional way.

But the highest achievement with this breadboard was obtained outside, when we looked at the dense and turbulent plume released by the smokestack of a waste incinerator plant. Using the full potential of the AOTF-based hyperspectral imager (tuning and acquisition speed, high image quality and sufficient spectral resolution), we managed to detect the presence of NO_2 in the plume and retrieve a relatively precise value for its concentration. As for the cross-section, this measurement technique is very unconventional and there exists only a very limited number of similar experiments, none of them with a tunable hyperspectral imager. Given the environmental conditions (wind speed, cloudiness, dis-

tance, ambient pollution), the chances of success were not so high, but in Chapter 5, we described the measurement campaign, the breadboard principle and the data processing yielding the NO_2 concentration.

In parallel, the investigations on the performance of an AOTF operating in the UV have led to the identification of an unexpected effect: the change of the filter transfer function and tuning curve caused by the heating of the crystal under sound propagation. In Chapter 4 we showed the link between the acoustic wave power and the filter characteristics modification. The measurements could be partially interpreted using temperature-dependent elastic and optical coefficients.

Finally, in a separate study on what could be achieved by a true imager in solar occultation geometry, we have examined what could ALTIUS do if, for any undefined reason, it was left with its imaging capability in solar occultation mode. In absence of any spectral information, it would become a camera taking snapshots of the refracted Sun as it sets or rises behind the atmosphere. By using a set of orthogonal polynomials, we have discussed in Chapter 2 how to retrieve pressure profiles from the analysis of the apparent flattening of the solar disk. In parallel, we have shown that this method is well suited for the automatic flagging of the presence of any high altitude cloud when performing occultations, a detection which is not straightforward for non-imaging systems.

We have tried in this work to establish the first milestones of a promising instrumental concept for the atmospheric remote sensing. We believe that the ALTIUS concept is a sound proposition to provide high-quality vertically resolved information on selected trace gases, and that this is mainly due to this AOTF-based hyperspectral imager concept. Apart from the pursuit of the normal work within the ALTIUS project (further simulations, design consolidation, ...), many further developments are waiting to consolidate this remote sensing instrumental concept.

First, as the instrument is readily available, we should develop the range of application of the visible breadboard. With its capabilities in the visible range, it could prove useful whenever image and spectral information are necessary: NO_2 emissions from ships, industries or city traffic for instance, but also aerosol detection thanks to the tunability of the AOTF. Inspired by the ALTIUS concept, a team from the University of Saskatchewan is currently preparing a balloon campaign with a similar instrument, but in the NIR domain. A collaboration could lead to a tandem instrument spanning the VIS-NIR range and capable of ozone, aerosol and NO_2 measurements in the stratosphere.

Similarly, the pursuit of collaboration with the group of Pr. Voloshinov could lead to a UV breadboard, extending the remote sensing pos-

sibilities to UV species like SO_2 or ozone in the higher stratosphere. In volcanology, SO_2 cameras form a strongly developing area. Commercial equipments already exist, but either they use grating spectrometers and they encounter difficulties with pointing knowledge and data interpretation, either they are based on interference filters whose spectral resolution is often quite poor and the central wavelength is not always optimal. An instrument designed around a UV AOTF could turn to be very useful for this community.

Appendix A

Error covariance of limb-scatter measurements

The ALTIUS instrument has no strong requirement on absolute radiometric accuracy. Its spectral imagers will collect photons and deliver rough data as arrays of values representing the counted electrons. The exact relationship between the number of photons entering the optical port and the number of counted electrons can only be determined after an absolute calibration campaign, generally performed before launch. This is often a pretty difficult and time-demanding task. It also implies to identify in-flight absolute calibration strategies in order to maintain the level of accuracy under harsh space environment (UV aging, outgassing contamination, vibrations, ...).

Solar and stellar occultation instruments perform what is usually referred to as *self-calibrated* measurements: the irradiance spectrum is first measured above the atmosphere, then it is used to normalize the consecutive spectra recorded along the sunset (attenuated by atmospheric extinction). In that case, as the reference signal is acquired within short timescale with respect to the real measurements, no calibration is needed. This argument holds for ALTIUS occultation measurements too.

The diffuse nature of the atmospheric bright limb makes the calibration problem more complex. For an imager with discrete spectral coverage, the necessity is to optimize the particle signatures with a limited number of measurement wavelengths. One approach is to build ratios of spectral images combining high and low sensitivity to the target species. This is the followed strategy for the O_3 and NO_2 measurements (see sections 3.6 and 3.7). Provided that there is good knowledge of the relative detection efficiency among the spectral images, it alleviates the need for absolute radiometric calibration.

Furthermore, as the multiple scattering processes at the lowest tan-

gent altitudes (multiple Rayleigh scattering, cloud reflections, wavelength dependent albedo, ...) can alter the signal observed at higher altitudes, a normalization among the rows of a same image is almost systematically performed. Typically, one row looking 5 km above the highest usefull measurement normalizes all the other ones below. This makes the retrieval algorithm less sensitive to radiative transfer model errors.

A.1 Limb scattering measurement uncertainty

The process of counting photons (or electrons) is not error-free. Its probability density function (pdf) follows a Poisson distribution. For an expectation value μ of the number of counts, the probability to actually count x is given by

$$P_\mu(x) = e^{-\mu}\,\frac{\mu^x}{x!}. \tag{A.1}$$

The Stirling formula, $x! \approx \sqrt{2\pi x}\exp(x\ln x - x)$, allows to write it in a different form:

$$P_\mu(x) \approx \frac{1}{\sqrt{2\pi x}}\,e^{x(\ln\mu - \ln x)+x-\mu}. \tag{A.2}$$

For total counts $\mu \gg 1$ (and therefore $x \gg 1$ as well), we can show that the Poisson distribution is well approximated by a normal distribution of expectation μ and equal variance ($\sigma^2 = \mu$). Defining $f(x) = x(\ln\mu - \ln x) + x - \mu$, we can see that $f(x)$ has one minimum in $x = \mu$. By a Taylor expansion around this minimum, we have

$$f(x) \approx -\frac{(x-\mu)^2}{2\mu}. \tag{A.3}$$

As the term $1/\sqrt{2\pi x}$ varies slowly when x gets large, it can be replaced by $1/\sqrt{2\pi\mu}$ when trying to evaluate eq. (A.2) around $x = \mu$. Replacing eq. (A.3) in eq. (A.2), we find

$$P_\mu(x) \approx \frac{1}{\sqrt{2\pi\mu}}\,e^{-\frac{(x-\mu)^2}{2\mu}}. \tag{A.4}$$

The hypothesis of replacing $1/\sqrt{2\pi x}$ by $1/\sqrt{2\pi\mu}$ holds even stronger when considering how quick the exponential term in eq. (A.4) vanishes when x deviates from μ.

The primary source of uncertainty on the number of counts registered by the pixels during their exposure time is therefore related to the width of the normal distribution described by eq. (A.4). This is called

the *shot noise* and it amounts to $\sqrt{\mu}$. We will assume that the counting process of a given pixel is independent of its neighbours. Hence, a radiance profile, i.e. a stack of pixels from the same image, is a vector I whose elements I_i correspond to different tangent altitudes h_i and follow each a $\mathcal{N}(\mu_i, \mu_i)$ distribution (according to eq. (A.4)). They are also independent from each other such that their sum would be described by a normal distribution $\mathcal{N}(\sum \mu_i, \sum \mu_i)$.

However, as explained before, the preferred approach in limb scattering measurements, is to build ratios of pictures whose rows have been normalized by a high altitude row. Unfortunately, the ratio of normally distributed quantities is not normally distributed and we have to solve the problem of determining the variance of the ratio vector elements. Moreover, one reference measurement is used to normalize the entire radiance profile, which introduces correlation among the ratio vector elements as well. Their covariance matrix is not a simple diagonal matrix anymore and the off-diagonal terms must be evaluated.

One method for computing the image ratio variance and covariance has been described in section 3.6.2. It is based on the linearization of the ratio expression. The measurement error is *propagated* to the final quantity. It is a widely adopted technique, even in papers adressing precision estimates specifically [13].

We introduce here another approach, which has the advantage of not requiring the derivative computation. It is based on working out the statistical moments of the logarithm of the image ratio.

A.2 Bright limb image ratio error covariance by the method of statistical moments

We start with the general definition of a ratio as it is encountered in the bright limb remote sensing community [45], [113], [93], [35], [12]. In order to emphasize the absorption of a species, a measurement at a strongly absorbed wavelength (I^s) is divided by one or more measurements performed at weakly absorbed wavelengths (I^{wj}, with $j = 1, \ldots, k$). If $k > 1$, a common practice is to weight the logarithm of these measurements by a factor $1/r_j$ such that $\sum 1/r_j = 1$. See section 3.6.1 for more details. Remembering that in bright limb, the radiance originating from a target layer (I_i) (not all the atmosphere is a target for every species) is divided by the light scattered at an upper altitude (I_n), we can write the following expression for one element (one tangent altitude) of the

ratio vector:

$$R_i = \frac{I_i^s}{\prod_{j=1}^{k} \sqrt[r_j]{I_i^{wj}}} \cdot \frac{\prod_{j=1}^{k} \sqrt[r_j]{I_n^{wj}}}{I_n^s}, \tag{A.5}$$

with $i = 1, \ldots, m$, the tangent altitude number. As already mentionned, any single quantity is an independent measurement of pdf $\mathcal{N}(\mu, \mu)$. The ratio, however, does not follow a normal distribution, and the normalization by I_n introduces uncertainty correlation among the elements of R.

Taking the log of expression (A.5) makes the problem at least linear:

$$\ln R_i = \ln I_i^s - \ln I_n^s + \sum_{j=1}^{k} \frac{1}{r_j} \left(\ln I_n^{wj} - \ln I_i^{wj} \right). \tag{A.6}$$

Moreover, for large values of μ, the normal distribution is quite narrow which makes a linear approximation of the log function reasonable:

$$\ln I \approx \ln \mu + \frac{1}{\mu}(I - \mu). \tag{A.7}$$

This expression helps to derive the expectation $\mathrm{E}[\cdot]$ and the variance $\mathrm{Var}[\cdot]$ of the logarithm of a normally distributed parameter:

$$\begin{aligned}
\mathrm{E}[\ln I] &= \mathrm{E}[\ln \mu] + \mathrm{E}[\frac{1}{\mu}(I - \mu)], \\
&= \ln \mu, \qquad\qquad (\text{A.8}) \\
\mathrm{Var}[\ln I] &= \mathrm{E}[(\ln I)^2] - \mathrm{E}[\ln I]^2, \\
&= \mathrm{E}[(\ln \mu + \frac{1}{\mu}(I - \mu))^2] - (\ln \mu)^2, \\
&= (\ln \mu)^2 + \mathrm{E}[\frac{1}{\mu^2}(I - \mu)^2] - (\ln \mu)^2, \\
&= \frac{1}{\mu^2} \left(\mathrm{E}[I^2] + \mathrm{E}[I]^2 - 2\mu \mathrm{E}[I] \right), \\
&= \frac{1}{\mu^2} \left(\mathrm{E}[I^2] - \mathrm{E}[I]^2 \right), \\
&= \frac{1}{\mu^2} \mathrm{Var}[I] = \frac{1}{\mu}. \qquad\qquad (\text{A.9})
\end{aligned}$$

As all the terms of eq. (A.6) are uncorrelated, we can simply add their respective variances to obtain $\mathrm{Var}[\ln R_i]$. Using eq. (A.9) and the property $\mathrm{Var}[a.X] = a^2.\mathrm{Var}[X]$, we find

$$\mathrm{Var}[\ln R_i] = \frac{1}{\mu_i^s} + \frac{1}{\mu_n^s} + \sum_{j=1}^{k} \frac{1}{r_j^2} \left(\frac{1}{\mu_n^{wj}} + \frac{1}{\mu_i^{wj}} \right). \tag{A.10}$$

Due to the normalization by high altitude measurements, the covariance matrix Cov[·] of $\ln R$ is not diagonal. Starting from the definition of the covariance, we have:

$$\begin{aligned}
\mathrm{Cov}[\ln R_i, \ln R_j] &= \mathrm{E}[(\ln R_i - \mathrm{E}[\ln R_i]).(\ln R_j - \mathrm{E}[\ln R_j])], \\
&= \mathrm{E}[\ln R_i \,.\, \ln R_j] - \mathrm{E}[\ln R_i] \,.\, \mathrm{E}[\ln R_j], \\
&= \mathrm{Var}[\ln I_n^s] + \sum_{j=i}^{k} \frac{1}{r_j^2} \mathrm{Var}[\ln I_n^{wj}], \\
&= \frac{1}{\mu_n^s} + \sum_{j=i}^{k} \frac{1}{r_j^2} \frac{1}{\mu_n^{wj}}, \qquad \text{(A.11)}
\end{aligned}$$

where the covariance of the $\ln I_i$ terms have disappeared due to their independence.

Now that the first statistical moments of $\ln R$ are known, we would like to go back to the initial quantity by taking the exponential. Defining the variable $Q_i = \ln R_i$, which has an expectation $\mathrm{E}[Q_i] = \ln \mu_{R_i}$ (from eq. (A.8)), we can again use the argument that for a large number of counts, the spread of the pdf is small and a first order expansion is a valid approximation. Therefore,

$$\exp Q_i \approx e^{\ln \mu_{R_i}} + e^{\ln \mu_{R_i}} . (\ln R_i - \ln \mu_{R_i}). \qquad \text{(A.12)}$$

One can verify from the above equation that

$$\mathrm{E}[\exp Q_i] = \mu_{R_i}, \qquad \text{(A.13)}$$

and that from the same reasoning that led to eq. (A.10) and (A.11), we find

$$\mathrm{Var}[R_i] = \mu_{R_i}^2 \, \mathrm{Var}[\ln R_i], \qquad \text{(A.14)}$$

and

$$\mathrm{Cov}[R_i, R_j] = \mu_{R_i} \, \mu_{R_j} \, \mathrm{Cov}[\ln R_i, \ln R_j]. \qquad \text{(A.15)}$$

Appendix A: Covariance of LS measurement error

Appendix B

The MAP log-normally distributed solution

In Appendix A, we described how to compute the error covariance matrix of an image ratio from the expectation and the variance of the probability distribution function of the measurements. In the Bayesian formalism of the inversion problem presented in the first chapter, the error covariance matrix of the measurement vector is an important quantity. We are now interested in the pdf of the solution to the inverse problem.

We already emphasized in section 3.6.3 that in order to reduce numerical instabilities that could arise from the very large numbers representing the species concentration, we were working with the logarithm of these quantities. This is the case for the weighting functions $\mathbf{K}$ and for the a priori constituent profile $\mathbf{x_a}$. In that case, the maximum a posteriori solution (MAP) is also given in term of the logarithm of the concentration:

$$\ln \hat{\mathbf{x}} = \ln \mathbf{x_a} + \left(\mathbf{K^T}\, \mathbf{C_y}^{-1}\, \mathbf{K} + \mathbf{C_a}^{-1} \right)^{-1} \mathbf{K^T}\, \mathbf{C_y}^{-1}\, (\mathbf{y} - \mathbf{K}\, \ln \mathbf{x_a})\,. \tag{B.1}$$

As this expression is derived under the assumption of normally-distributed quantities, $\ln \hat{\mathbf{x}}$ has a log-normal pdf whose variance $\hat{\sigma}^2$ is given by the diagonal elements of the retrieval error covariance matrix

$$\hat{\mathbf{S}} = \left(\mathbf{K}^T\, \mathbf{C_y}^{-1}\, \mathbf{K} + \mathbf{C_a}^{-1} \right)^{-1}. \tag{B.2}$$

To be consistent with this formalism, the final concentration profile is not simply given by $\exp(\ln \hat{\mathbf{x}})$. Particularly in the case of simulations, if the measurement vector $\mathbf{y}$ consists of the expectation of a normal distribution (which is equivalent to the noise-free situation), then the inverse problem solution should be given by the expectation of $\hat{\mathbf{x}}$. According to

the moments of the log-normal distribution, we have:

$$\mathrm{E}[\hat{\mathbf{x}}] = e^{\ln \hat{\mathbf{x}} + \hat{\sigma}^{\mathbf{2}}/\mathbf{2}}, \tag{B.3}$$

$$\mathrm{Var}[\hat{\mathbf{x}}] = \left(e^{\hat{\sigma}^2} - 1\right) e^{2\ln \hat{\mathbf{x}} + \hat{\sigma}^{\mathbf{2}}}. \tag{B.4}$$

The difference with $\exp(\ln \hat{\mathbf{x}})$ can be non-negligible if the retrieval uncertainties are large, which is the case when the averaging kernels get smaller. As an example, both quantities are compared in the case of a NO_2 retrieval (fig. B.1). Clearly, the retrieval accuracy gets worsened at lower and higher altitudes, but it is the only statistically consistent solution.

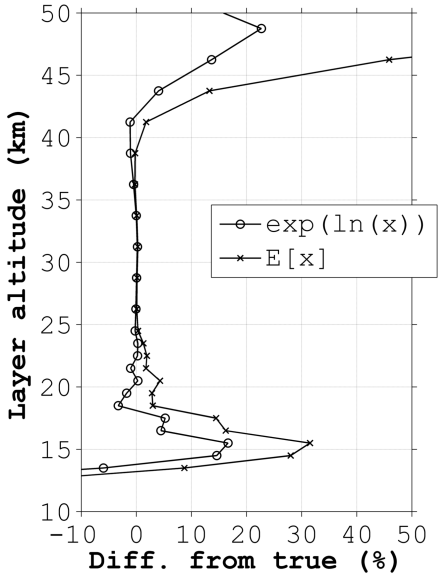

Figure B.1 – Comparison of the NO_2 retrieved profile accuracy when simple exponential is used (circles) and when the log-normal nature of the solution is considered (crosses).

Appendix C

Acousto-optic interaction and polarization aspects

One of the very interesting features of an AOTF is the way the incident and diffracted beams can be distinguished. Not only do they exit the acousto-optic (AO) interaction region in slightly different directions, but if the incident one is linearly polarized, the diffracted one has its own polarization state rotated by 90°. Unwanted order suppression strategies relying on both aspects potentially yield very efficient straylight reduction and high image quality.

Here we provide additional details related to polarization aspects in AOTFs, starting from concepts introduced in section 4.2.3. We will also work out the case of the KDP AOTF as it offers a more simple geometry of interaction compared to TeO_2. In support to the discussion, we reprint an adapted version of the momentum matching diagram in Fig. C.1. The oscillation direction of the light beams electric field has been added.

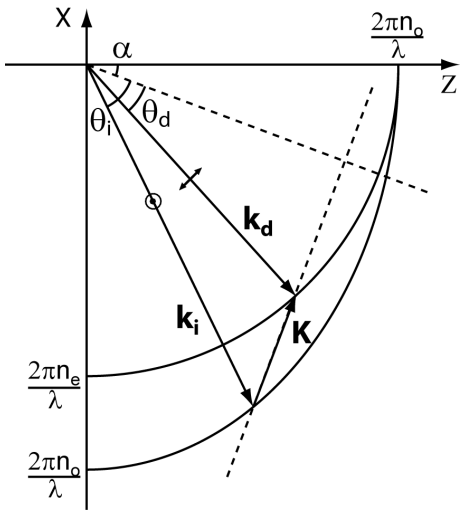

Figure C.1 – Momentum matching diagram in a KDP AOTF.

As KDP is optically and acoustically anisotropic, its properties depend on the propagation direction, be it a light beam or an acoustic wave. Most parameters take therefore the form of rank-2 (relative dielectric permittivity $\boldsymbol{\epsilon}$, impermeability $\boldsymbol{\eta}$, acoustic strain $\mathbf{S}$,...) or rank-4 tensors

(photo-elastic tensor $\bar{\mathbf{p}}$, susceptibility tensor $\boldsymbol{\chi}$,...) in order to represent the multiple connections existing between those quantities.

As seen in section 4.2.1, a given material can offer up to three eigenmodes of polarization to an incident electric field (light). Each mode is associated to a refractive index which determines the wave velocity. In the frame of the eigenmodes, the relative dielectric permittivity tensor is diagonal:

$$\boldsymbol{\epsilon} = \begin{pmatrix} n_x^2 & 0 & 0 \\ 0 & n_y^2 & 0 \\ 0 & 0 & n_z^2 \end{pmatrix}. \tag{C.1}$$

These three refractive indices associated to three orthogonal polarization modes are the length of the principal axes of the index ellipsoid:

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1.$$

In its most general form, one uses the impermeability tensor to express the index ellipsoid:

$$\eta_{ij}\, x_i\, x_j \;=\; 1. \tag{C.2}$$

It follows that $\boldsymbol{\eta} = \boldsymbol{\epsilon}^{-1}$.

The elasto-optic effect describes how an elastic deformation can change the optical properties of a medium. One way of expressing this is to suppose that the acoustically-induced strain affects the impermeability tensor:

$$\Delta\eta_{ij} = p_{ijkl}\, S_{kl}. \tag{C.3}$$

The coupling strength is dictated by the interaction medium properties and captured in the elements of the photo-elastic tensor $\bar{\mathbf{p}}$. Obviously, a perturbation of $\boldsymbol{\eta}$ is reflected in a change in the dielectric permittivity. As $\boldsymbol{\eta} = \boldsymbol{\epsilon}^{-1}$, we can write

$$\epsilon_{im}\, \eta_{mn} = \delta_{in}. \tag{C.4}$$

Differentiating this expression, we find

$$\Delta\epsilon_{im}\eta_{mn} + \epsilon_{im}\Delta\eta_{mn} = 0.$$

Multiplying by ϵ_{nj} to the right and using eq. (C.4), we obtain the associated perturbation for the permittivity:

$$\Delta\epsilon_{ij} = -\epsilon_{im}\Delta\eta_{mn}\epsilon_{nj}. \tag{C.5}$$

Recalling the definition of the dielectric susceptibility, $\boldsymbol{\epsilon} = 1 + \boldsymbol{\chi}$, we can write equivalently:

$$\Delta\chi_{ij} = -\epsilon_{im}\Delta\eta_{mn}\epsilon_{nj}. \tag{C.6}$$

From the point of view of the optical properties of the crystal, the elastic wave is responsible for a deformation of the index ellipsoid (via $\boldsymbol{\Delta\eta}$). But considering electric field aspects, it corresponds to the appearance of some additional dielectric susceptibility, such that an electric field will induce an additional polarization wave:

$$\mathbf{P}^{\mathrm{AO}} = \epsilon_{\mathbf{0}} \boldsymbol{\Delta\chi} \mathbf{E}. \tag{C.7}$$

Using eq. (C.6) and (C.3), we can write explicitly

$$P_i^{\mathrm{AO}} = -\epsilon_0 \epsilon_{im} p_{mnkl}\, S_{kl} \epsilon_{nj} E_j. \tag{C.8}$$

To see how the incident electric field is handled by the adjacent tensor products in eq. (C.8), let us consider the geometry depicted by Fig. C.1. In the frame defined by the index ellipsoid axes, the incident polarization sees the ordinary index n_o as it is aligned with the Y-axis[1]. The incident electric field is described by

$$\mathbf{E_i}(\mathbf{r}, t) = \frac{1}{2}\, \hat{\mathbf{e}}^{\mathbf{i}} \left(E_i(\mathbf{r}). \exp\left(i\omega_i t - i\mathbf{k_i}.\mathbf{r} \right) + \text{c.c.} \right), \tag{C.9}$$

with in this case,

$$\hat{\mathbf{e}}^{\mathbf{i}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{C.10}$$

The strain tensor is determined from the particle displacement vector $\mathbf{u}$ (see eq. (4.11)-(4.14)). The shear acoustic wave is responsible for the following displacement along the Y-axis:

$$\mathbf{u}(\mathbf{r}, t) = \hat{\mathbf{y}} U \sin\left(\Omega t - \mathbf{K}.\mathbf{r} \right). \tag{C.11}$$

As the acoustic wave propagates at an angle α from the X-axis in the XZ-plane, displacing particles along the Y-axis, the only non-zero strain tensor elements are the $S_{yx} = S_{xy}$ and $S_{yz} = S_{zy}$. Using eq. (4.12), we find:

$$\begin{aligned} S_{yx} = S_{xy} &= -\frac{1}{2} UK \cos\alpha \cos\left(\Omega t - \mathbf{K}.\mathbf{r} \right), \\ S_{yz} = S_{zy} &= -\frac{1}{2} UK \sin\alpha \cos\left(\Omega t - \mathbf{K}.\mathbf{r} \right). \end{aligned}$$

[1] KDP is a uniaxial crystal, i.e. its index ellipsoid has a circular cross-section (defining the XY-plane) and one shorter axis (Z). Any polarization departing from the XY-plane will experience a refractive index belonging to the $[n_e, n_o]$ range.

One can therefore express the strain wave as follows:

$$\mathbf{S}(\mathbf{r},t) = \frac{1}{2}\hat{\mathbf{s}}S\cos\left(\Omega t - \mathbf{K}.\mathbf{r}\right), \tag{C.12}$$

with $\hat{\mathbf{s}}$ playing the role of the unit strain tensor, also represented as a unit vector using the abbreviated subscript notation:

$$\hat{\mathbf{s}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sin\alpha \\ 0 \\ \cos\alpha \end{pmatrix}. \tag{C.13}$$

With the same notation, the photo-elastic tensor for KDP is constituted as follows:

$$\bar{\mathbf{p}} = \left(\begin{array}{ccc|ccc} p_{11} & p_{12} & p_{13} & 0 & 0 & 0 \\ p_{12} & p_{11} & p_{13} & 0 & 0 & 0 \\ p_{31} & p_{31} & p_{33} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & p_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{66} \end{array}\right). \tag{C.14}$$

Putting aside the amplitude of the waves or their space and time varying dependence, we can go back to the acousto-optically induced polarization wave given in eq. (C.8). Using eq. (C.1), (C.13) and (C.14), we find:

$$\frac{-1}{\epsilon_0}\mathbf{P}^{\mathrm{AO}} \propto \begin{pmatrix} 0 & n_x^2 n_y^2 p_{66}\cos\alpha & 0 \\ n_x^2 n_y^2 p_{66}\cos\alpha & 0 & n_y^2 n_z^2 p_{44}\sin\alpha \\ 0 & n_y^2 n_z^2 p_{44}\sin\alpha & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{C.15}$$

Finally, the vector-matrix product yields the following result:

$$\mathbf{P}^{\mathrm{AO}} \propto -\epsilon_0 \begin{pmatrix} n_x^2 n_y^2 p_{66}\cos\alpha \\ 0 \\ n_y^2 n_z^2 p_{44}\sin\alpha \end{pmatrix}. \tag{C.16}$$

We see that the polarization direction of the diffracted beam indeed lies in the XZ-plane, as represented in Fig. C.1.

To summarize, we have seen that the local elastic deformations caused by the shear acoustic wave in the KDP crystal are responsible for some additional dielectric susceptibility. In the presence of an electric field, it yields an extra polarization and this is the elasto-optic effect. As KDP is

optically and acoustically anisotropic, tensors are used to represent this interaction. The combined action of a shear strain and the tetragonal symmetry of the medium makes the elasto-optic effect take the form of a rotation matrix. Hence, an incident linearly polarized light coupling with the local shear strain induces an electric displacement field whose polarization is rotated by 90°.

Appendix C: AO interaction and polarization

Bibliography

[1] O. Aharon and I. Abdulhalim. Liquid crystal Lyot tunable filter with extended free spectral range. *Opt. Exp.*, 17:11426–11433, 2009.

[2] J. H. Allen. Orthogonality and convergence of discrete Zernike polynomials. Master's thesis, University of New Mexico, New Mexico, U.S.A., 2010.

[3] J. Antila, A. Miranto, J. Mäkynen, M. Laamanen, A. Rissanen, M. Blomberg, H. Saari, and J. Malinen. MEMS and piezo actuator-based Fabry-Perot interferometer technologies and applications at VTT. *Proc. SPIE*, 7680:1–12, 2010.

[4] G. Arlt and H. Schweppe. Paratellurite, a new piezoelectric material. *Solid State Communications*, 6(11):783–784, 1968.

[5] B. A. Auld. *Acoustic Fields and Waves in Solids.* Krieger Publishing Company, 1990.

[6] N. P. Barnes and D. J. Gettemy. Variation of the refractive index with temperature and the tuning rate for KDP isomorphs. *J. Opt. Soc. Am.*, 72:895–898, 1982.

[7] A. Berk, G.P. Anderson, P.K. Acharya, L.S. Bernstein, L. Muratov, J. Lee, M. Fox, S.M. Adler-Golden, J.H. Chetwynd, M.L. Hoke, R.B Lockwood, J.A. Gardner, T.W. Cooley, C.C. Borel, P.E. Lewis, and E.P. Shettle. MODTRAN5: 2006 Update. *Proc. SPIE*, 6233, 2006.

[8] P. F. Bernath, C. T. McElroy, M. C. Abrams, C. D. Boone, M. Butler, C. Camy-Peyret, M. Carleer, C. Clerbaux, P.-F. Coheur, R. Colin, and et al. Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Let.*, 32:L15S01, 2005.

[9] J. Bertaux, G. Megie, T. Widemann, E. Chassefiere, R. Pellinen, E. Kyrola, S. Korpela, and P. Simon. Monitoring of ozone trend by stellar occultations: The GOMOS instrument. *Adv. Space Res.*, 11:3237–3242, 1991.

[10] C. D. Boone, R. Nassar, K. A. Walker, Y. Rochon, S. D. McLeod, C. P. Rinsland, and P. F. Bernath. Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer. *Appl. Opt.*, 44(33):7218–7231, 2005.

[11] A. E. Bourassa, D. A. Degenstein, and E. J. Llewellyn. SASKTRAN: A spherical geometry radiative transfer code for efficient estimation of limb scattered sunlight. *J. Quant. Spectrosc. Ra.*, 109:52–73, 2008.

[12] A. E. Bourassa, C. A. McLinden, C. E. Sioris, S. Brohede, A. F. Bathgate, E. J. Llewellyn, and D. A. Degenstein. Fast NO2 retrievals from Odin-OSIRIS limb scatter measurements. *Atmos. Meas. Tech.*, 4:965–972, 2011.

[13] A. E. Bourassa, A. F. Bathgate, B. J. Elash, and D. A. Degenstein. Precision estimate for Odin-OSIRIS limb scatter retrievals. *J. Geophys. Res.*, 117:D04303, 2012.

[14] A. E. Bourassa, D. A. Degenstein, W. J. Randel, J. M. Zawodny, E. Kyrölä, C. A. McLinden, C. E. Sioris, and C. Z. Roth. Trends in stratospheric ozone derived from merged SAGE II and Odin-OSIRIS satellite observations. *Atmos. Chem. Phys.*, 14:6983–6994, 2014.

[15] H. Bovensmann, J. P. Burrows, M. Buchwitz, J. Frerick, S. Noël, V. V. Rozanov, K. V. Chance, and A. P. H. Goede. SCIAMACHY: Mission Objectives and Measurement Modes. *J. Atmos. Sci.*, 56:127–150, 1999.

[16] G. P. Brasseur and S. Solomon. *Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere*. Springer, 2005.

[17] W. Bray, editor. *The diary of John Evelyn*, volume 2. W. Dunne, 1901. In open access at `https://archive.org`.

[18] L. Brillouin. Diffusion de la lumière et des rayons X par un corps transparent homogène: influence de l'agitation thermique. *Ann. Phys.*, 17:88–122, 1922.

[19] S. Brohede, C. A. McLinden, J. Urban, C. S. Haley, A. I. Jonsson, and D. Murtagh. Odin stratospheric proxy NOy measurements and climatology. *Atmos. Chem. Phys.*, 8:5731–5754, 2008.

[20] A. Bucholtz. Rayleigh-scattering calculations for the terrestrial atmosphere. *Appl. Opt.*, 34:2765–2773, 1995.

[21] J. P. Burrows, A. Dehn, B. Deters, S. Himmelmann, A. Richter, S. Voigt, and J. Orphal. Atmospheric remote-sensing reference data from GOME: Part 1. Temperature-dependent absorption cross-sections of NO_2 in the 231-794 nm range. *J. Quant. Spectrosc. Rad. Transfer*, 60:1025–l031, 1998.

[22] J. P. Burrows, A. Richter, A. Dehn, B. Deters, S. Himmelmann, S. Voigt, and J. Orphal. Atmospheric remote-sensing reference data from GOME-2. Temperature-dependent absorption cross sections of O3 in the 231-794 nm range. *J. Quant. Spectrosc. Ra.*, 61:509–517, 1999.

[23] S. Chandrasekhar. *Radiative Transfer*. Dover Publications, 1960.

[24] I. C. Chang. Noncollinear acousto-optic filter with large angular aperture. *Appl. Phys. Lett.*, 25:370–372, 1974.

[25] W. Chehade, M. Weber, and J. P. Burrows. Total ozone trends and variability during 1979–2012 from merged data sets of various satellites. *Atmos. Chem. Phys.*, 14:7059–7074, 2014.

[26] S. I. Chizhikov, N. G. Sorokin, I. Ledovskaya, and E. V. Makarevskaya. Elastic properties of KDP and DKDP crystals of high temperatures. *Sov. Phys. Crystallogr.*, 18:544–545, 1974.

[27] D. A. Chu, Y. J. Kaufman, G. Zibordi, J. D. Chern, J. Mao, C. Li, and B. N. Holben. Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). *J. Geophys. Res.*, 108, 2003.

[28] W. P. Chu, M. P. McCormick, J. Lenoble, C. Brogniez, and P. Pruvost. SAGE II inversion algorithm. *J. Geophys. Res.*, 94:8339–8351, 1989.

[29] P. E. Ciddor. Refractive index of air: new equations for the visible and near infrared. *Appl. Opt.*, 35:1566–1573, 1996.

[30] P. J. Crutzen. Estimates of possible future ozone reductions from continued use of fluoro-chloro-methanes (CF_2Cl_2, $CFCl_3$). *Geophys. Res. Lett.*, 1:205–208, 1974.

[31] J. Côté, A. Kron, J. de Lafontaine, J. Naudet, and S. Santandrea. PROBA-2 Attitude and Orbit Control System: In-Flight Results of Innovative GNC Functions. *Proc. of the 18th IFAC World Congress*, 18:721–726, 2011.

[32] D. M. Cunnold, J. M. Zawodny, W. P. Chu, J. P. Pommereau, F. Goutail, J. Lenoble, M. P. McCormick, R. E. Veiga, D. Murcray, N. Iwagami, K. Shibasaki, P. C. Simon, and W. Peetermans. Validation of SAGE II NO2 measurements. *J. Geophys. Res. Atm.*, 96:12913–12925, 1991.

[33] F. Dalaudier, V. Kan, and A. S. Gurvich. Chromatic Refraction with Global Ozone Monitoring by Occultation of Stars. I. Description and Scintillation Correction. *Appl. Opt.*, 40:866–877, 2001.

[34] P. Debye and F. W. Sears. On the scattering of light by supersonic waves. *Proc. Natl. Acad. Sci. USA*, 18:409–414, 1932.

[35] D. A. Degenstein, A. E. Bourassa, C. Z. Roth, and E. J. Llewellyn. Limb scatter ozone retrieval from 10 to 60 km using a multiplicative algebraic reconstruction technique. *Atmos. Chem. Phys.*, 9:6521–6529, 2009.

[36] E. Dekemper, N. Loodts, B. Van Opstal, J. Maes, F. Vanhellemont, N. Mateshvili, G. Franssens, D. Pieroux, C. Bingen, C. Robert, L. De Vos, L. Aballea, and D. Fussen. Tunable acousto-optic spectral imager for atmospheric composition measurements in the visible spectral domain. *Appl. Opt.*, 51:6259–6267, 2012.

[37] E. Dekemper, F. Vanhellemont, N. Mateshvili, G. Franssens, D. Pieroux, C. Bingen, C. Robert, and D. Fussen. Zernike polynomials applied to apparent solar disk flattening for pressure profile retrievals. *Atmos. Meas. Tech.*, 6:823–835, 2013.

[38] T. Deshler, J. L. Mercer, H. G. J. Smit, R. Stubi, G. Levrat, B. J. Johnson, S. J. Oltmans, R. Kivi, A. M. Thompson, J. Witte, J. Davies, F. J. Schmidlin, G. Brothers, and T. Sasaki. Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes. *J. Geophys. Res.*, 113:D04307, 2008.

[39] G. M. B. Dobson. Forty Years Research on Atmospheric Ozone at Oxford: a History. *Appl. Opt.*, 7:387–405, 1968.

[40] J. Dodion, D. Fussen, F. Vanhellemont, C. Bingen, N. Mateshvili, K. Gilbert, R. Skelton, D. Turnbull, S. D. McLeod, C. D. Boone, K. A. Walker, and P. F. Bernath. Cloud detection in the upper troposphere-lower stratosphere region via ACE imagers: A qualitative study. *J. Geophys. Res.*, 112:D03208, 2007.

[41] R. Dragani, S. Abdalla, R. J. Engelen, A. Inness, and J.-N. Thépaut. Ten years of ENVISAT observations at ECMWF: a review of activities and lessons learnt. *Q. J. R. Meteorol. Soc.*, 2014.

[42] Q. Errera, F. Daerden, S. Chabrillat, J. C. Lambert, W. A. Lahoz, S. Viscardy, S. Bonjean, and D. Fonteyn. 4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses. *Atmos. Chem. Phys.*, 8:6169–6187, 2008.

[43] C. Fabry and H. Buisson. L'absorption de l'ultra-violet par l'ozone et la limite du spectre solaire. *J. Phys. Theor. Appl.*, 3:196–206, 1913.

[44] J. C. Farman, B. G. Gardiner, and J. D. Shanklin. Large losses of total ozone in Antarctica reveal seasonal ClOx/NOx interaction. *Nature*, 315:207–210, 1985.

[45] D. E. Flittner, P. K. Bhartia, and B. M. Herman. O3 profiles retrieved from limb scatter measurements: Theory. *Geophys. Res. Lett.*, 27:2601–2604, 2000.

[46] G. Georgiev, D. A. Glenar, and J. J. Hillman. Spectral Characterization of Acousto-Optic Filters Used in Imaging Spectroscopy. *Appl. Opt.*, 41:209–217, 2002.

[47] G. C. Ghosh and G. C. Bhar. Temperature dispersion in ADP, KDP and KD*P for nonlinear devices. *IEEE J. Quantum Electron.*, 18:143–145, 1982.

[48] K. L. Gilbert, D. N. Turnbull, K. A. Walker, C. D. Boone, S. D. McLeod, M. Butler, R. Skelton, P. F. Bernath, F. Chateauneuf, and M.-A. Soucy. The onboard imagers for the Canadian ACE SCISAT-1 mission. *J. Geophys. Res.*, 112:D12207, 2007.

[49] L. L. Gordley, J. Burton, B. T. Marshall, M. McHugh, L. Deaver, J. Nelsen, J. M. Russell, and S. Bailey. High precision refraction

measurements by solar imaging during occultation: results from SOFIE. *Appl. Opt.*, 48:4814–4825, 2009.

[50] L. L. Gordley, M. Hervig, C. Fish, J. M. Russell III, S. Bailey, J. Cook, S. Hansen, A. Shumway, G. Paxton, L. Deaver, T. Marshall, J. Burton, B. Magill, C. Brown, E. Thompson, and J. Kemp. The Solar Occultation For Ice Experiment (SOFIE). *J. Atmos. Solar-Terr. Phys.*, 2009.

[51] M. Gottwald, E. Krieg, S. Slijkhuis, G. Lichtenberg, C. von Savigny, S. Noël, H. Bovensmann, K. Bramstedt, and A. Rozanov. The status of the SCIAMACHY line-of-sight pointing knowledge. *Proc. ESA Living Planet Symp.*, 2010.

[52] W. Gunning, J. Pasko, and J. Tracy. A Liquid Crystal Tunable Spectral Filter: Visible And Infrared Operation. *Proc. SPIE*, 0268:190–194, 1981.

[53] N. Gupta and V. Voloshinov. Hyperspectral Imager, from Ultraviolet to Visible, with a KDP Acousto-Optic Tunable Filter. *Appl. Opt.*, 43:2752–2759, 2004.

[54] C. S. Haley and S. Brohede. Status of the Odin/OSIRIS stratospheric O3 and NO2 data products. *Can. J. Phys.*, 85(11):1177–1194, 2007.

[55] J. W. Harder, J. W. Brault, P. V. Johnston, and G. H. Mount. Temperature dependent NO_2 cross sections at high spectral resolution. *J. Geophys. Res.*, 102:3861–3879, 1997.

[56] S. E. Harris and R. W. Wallace. Acousto-Optic Tunable Filter. *J. Opt. Soc. Am.*, 59:744–747, 1969.

[57] B. Hassler, I. Petropavlovskikh, J. Staehelin, T. August, P. K. Bhartia, C. Clerbaux, D. Degenstein, M. De Mazière, and et al. Past changes in the vertical distribution of ozone–Part 1: Measurement techniques, uncertainties and availability. *Atmos. Meas. Tech.*, 7(5):1395–1427, 2014.

[58] B. N. Holben, T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov. AERONET: A federated instrument network and Data archive for aerosol characterization. *Remote Sens. Environ.*, 66:1–16, 1998.

[59] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory*, 8:179–187, 1962.

[60] S. J. Janz, E. Hilsenrath, D. E. Flittner, and D. F. Heath. Rayleigh scattering attitude sensor. *SPIE Proc.*, 2831, 1996.

[61] I. T. Jolliffe. *Principal Component Analysis*. Springer, New York, 2nd edition, 2002.

[62] A. Jones, J. Urban, D. P. Murtagh, P. Eriksson, S. Brohede, C. Haley, D. Degenstein, A. Bourassa, C. von Savigny, T. Sonkaew, A. Rozanov, H. Bovensmann, and J. Burrows. Evolution of stratospheric ozone and water vapour time series studied with satellite measurements. *Atmos. Chem. Phys.*, 9:6055–6075, 2009.

[63] P. Katzka and I. C. Chang. Noncollinear Acousto-Optic Filter For The Ultraviolet. *Proc. SPIE*, 0202:26–32, 1980.

[64] G. M. Keating, editor. *COSPAR International Reference Atmosphere (CIRA), Part III: Trace Constituent Reference Models*. Advances in space research. Committee on Space Research, 1996.

[65] A. Khotanzad and Y. H. Hong. Invariant Image Recognition by Zernike Moments. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(5):489–497, 1990.

[66] K. W. Kirby and L. G. DeShazer. Refractive indices of 14 nonlinear crystals isomorphic to KH2PO4. *J. Opt. Soc. Am.*, 4:1072–1078, 1987.

[67] S. Kirkwood, P. Dalin, and A. Réchou. Noctilucent clouds observed from the UK and Denmark–trends and variations over 43 years. *Ann. Geophys.*, 26:1243–1254, 2008.

[68] H. Koschmieder. Theorie der horizontale Sichtweite. *Beitr. Physik fr. Atmos.*, 12:33–55, 1924.

[69] K. Lefever, R. van der A, F. Baier, Y. Christophe, Q. Errera, H. Eskes, J. Flemming, A. Inness, L. Jones, J.-C. Lambert, B. Langerock, M. G. Schultz, O. Stein, A. Wagner, and S. Chabrillat. Copernicus atmospheric service for stratospheric ozone: validation and intercomparison of four near real-time analyses, 2009–2012. *Atmos. Chem. Phys. Discuss.*, 14:12461–12523, 2014.

[70] J. Lenoble. *Atmospheric radiative transfer*. A. Deepak Pub., 1993.

[71] F. A. Lindemann and G. M. B. Dobson. A Theory of Meteors, and the Density and Temperature of the Outer Atmosphere to which it Leads. *Proc. R. Soc. Lond. A*, 102:411–437, 1923.

[72] K.-N. Liou. *An Introduction to Atmospheric Radiation.* Academic Press, 2002.

[73] E. J. Llewellyn and et al. The OSIRIS instrument on the Odin spacecraft. *Can. J. Phys.*, 82:411–422, 2004.

[74] F. Lohberger, G. Hönninger, and U. Platt. Ground-Based Imaging Differential Optical Absorption Spectroscopy of Atmospheric Gases. *Appl. Opt.*, 43:4711–4717, 2004.

[75] E. N. Lorenz. Empirical orthogonal functions and statistical weather prediction. Technical report, Department of Meteorology, MIT, Cambridge, Massachusetts, U.S.A., 1956.

[76] R. P. Loughman, D. E. Flittner, B. M. Herman, P. K. Barthia, E. Hilsenrath, and R. D. McPeters. Description and sensitivity analysis of a limb scattering ozone retrieval algorithm. *J. Geophys. Res.*, 110:D19301, 2005.

[77] R. Lucas and P. Biquard. Propriétés optiques des milieux solides et liquides soumis aux vibrations élastiques ultra sonores. *J. Phys. Radium*, 3:464–477, 1932.

[78] J. D. Lumpe, R. M. Bevilacqua, K. W. Hoppel, and C. E. Randall. POAM III retrieval algorithm and error analysis. *J. Geophys. Res.*, 107:4575–4580, 2002.

[79] W. P. Mason. The Elastic, Piezoelectric, and Dielectric Constants of Potassium Dihydrogen Phosphate and Ammonium Dihydrogen Phosphate. *Phys. Rev.*, 69:173–194, 1946.

[80] Y. Matsumi and M. Kawasaki. Photolysis of atmospheric ozone in the ultraviolet region. *Chem. Rev.*, 103:4767–4781, 2003.

[81] C. A. McLinden, V. E. Fioletov, C. S. Haley, N. Lloyd, C. Roth, D. Degenstein, A. Bourassa, C. T. McElroy, and E. J. Llewellyn. An evaluation of Odin/OSIRIS limb pointing and stratospheric ozone through comparisons with ozonesondes. *Can. J. Phys.*, 85:1125–1141, 2007.

[82] R. D. McPeters, S. J. Janz, E. Hilsenrath, T. L. Brown, D. E. Flittner, and D. F. Heath. The retrieval of O3 profiles from limb scatter measurements: Results from the Shuttle Ozone Limb Sounding Experiment. *Geophys. Res. Lett.*, 27:2597–2600, 2000.

[83] R. D. McPeters and G. J. Labow. Climatology 2011: An MLS and sonde derived ozone climatology for satellite retrieval algorithms. *J. Geophys. Res.*, 117:D10303, 2012.

[84] F. Mélin and G. Zibordi. Aerosol variability in the Po Valley analyzed from automated optical measurements. *Geophys. Res. Lett.*, 32:L03810, 2005.

[85] D. Mihalas. *Stellar Atmospheres*. W. H. Freeman, 1978.

[86] M. J. Molina and F. S. Rowland. Stratospheric sink for chlorofluoromethanes: chlorine atom-catalysed destruction of ozone. *Nature*, 249:810–812, 1974.

[87] H. Neckel. Analytical Reference Functions $F(\lambda)$ for the Sun's Limb Darkening and Its Absolute Continuum Intensities ($\lambda\lambda$ 300 to 1100 nm). *Solar Physics*, 229:13–33, 2005.

[88] L. Oikarinen. Polarization of light in UV-visible limb radiance measurements. *J. Geophys. Res.*, 106:1533–1544, 2001.

[89] R. A. Phillips. Temperature variation of the index of refraction of ADP, KDP and deuterated KDP. *J. Opt. Soc. Am.*, 56:629–632, 1966.

[90] U. Platt. Differential optical absorption spectroscopy (DOAS). *Chem. Anal. Series*, 127:27–83, 1994.

[91] A. N. Ponomarev, I. D. Rodionov, and G. E. Teterin. Wide-aperture acousto-optic tunable filters for visible and UV lights. *Proc. SPIE*, 2449:200–207, 1995.

[92] N. Rahpoe, C. von Savigny, M. Weber, A. V. Rozanov, H. Bovensmann, and J. P. Burrows. Error budget analysis of SCIAMACHY limb ozone profile retrievals using the SCIATRAN model. *Atmos. Meas. Tech.*, 6:2825–2837, 2013.

[93] D. F. Rault. Ozone profile retrieval from Stratospheric Aerosol and Gas Experiment (SAGE III) limb scatter measurements. *J. Geophys. Res.*, 110:D09309, 2005.

[94] C. D. Rodgers. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing Company, Incorporated, 2000.

[95] L. S. Rothman and et al. The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Ra.*, 110:533–572, 2009.

[96] S. Sabbah, R. Harig, P. Rusch, J. Eichmann, A. Keens, and J.-H. Gerhard. Remote sensing of gases by hyperspectral imaging: system performance and measurements. *Opt. Eng.*, 51(11):111717, 2012.

[97] G. Seibert. *The history of sounding rockets and their contribution to European space research*. ESA Publications Division, 2006.

[98] K. Stamnes. The theory of multiple scattering of radiation in plane parallel atmospheres. *Rev. Geophys.*, 24:299–310, 1986.

[99] C. Stedham, M. Draper, J. Ward, E. Wachman, and C. Pannell. A novel acousto-optic tunable filter for use in hyperspectral imaging systems. *Proc. SPIE*, 6889, 2008.

[100] S. Tegtmeier and et al. SPARC Data Initiative: A comparison of ozone climatologies from international satellite limb sounders. *J. Geophys. Res.*, 118:D019877, 2013.

[101] B. Thurairajah, S. M. Bailey, K. Nielsen, C. E. Randall, J. D. Lumpe, M. J. Taylor, and J. M. Russell III. Morphology of polar mesospheric clouds as seen from space. *J. Atmos. Solar-Terr. Phys.*, 104:234–243, 2013.

[102] M. Toohey and et al. Characterizing sampling biases in the trace gas climatologies of the sparc data initiative. *J. Geophys. Res.*, 118:D020298, 2013.

[103] N. Uchida. Optical Properties of Single-Crystal Paratellurite (TeO_2). *Phys. Rev. B*, 4:3736–3745, 1971.

[104] H. C. van de Hulst. *Multiple light scattering: tables, formulas, and applications*. Academic Press, 1980.

[105] H. C. van de Hulst. *Light Scattering by Small Particles*. Dover Books on Physics. Dover Publications, 2012.

[106] A. C. Vandaele, C. Hermans, S. Fally, M. Carleer, R. Colin, M.-F. Mérienne, A. Jenouvrier, and B. Coquart. High-resolution Fourier

transform measurement of the NO2 visible and near-infrared absorption cross sections: Temperature and pressure effects. *J. Geophys. Res.*, 107:4348–4361, 2002.

[107] A. C. Vandaele, C. Hermans, P. C. Simon, M. Roozendael, J. M. Guilmot, M. Carleer, and R. Colin. Fourier transform measurement of NO2 absorption cross-section in the visible range at room temperature. *J. Atmos. Chem.*, 25:289–305, 1996.

[108] V. N. Vishnevskii and I. V. Stefanskii. Temperature dependence of the dispersion of the refractivity of ADP and KDP single crystals. *Opt. Spectrosc.*, 20:195–196, 1966.

[109] D. Vokrouhlicky, P. Farinella, and F. Mignard. Solar radiation pressure perturbations for Earth satellites. 1: A complete theory including penumbra transitions. *Astron. Astrophys.*, 280:295–312, 1993.

[110] V. Voloshinov and N. Gupta. Ultraviolet-Visible Imaging Acousto-Optic Tunable Filters in KDP. *Appl. Opt.*, 43:3901–3909, 2004.

[111] V. B. Voloshinov and N. Gupta. Tunable acousto-optic filters for monitoring of atmospheric ozone. *Proc. SPIE*, 4574:162–173, 2002.

[112] V. B. Voloshinov, K. B. Yushkov, and B. B. J. Linde. Improvement in performance of a TeO_2 acousto-optic imaging spectrometer. *J. Opt. A: Pure Appl. Op.*, 9:341–347, 2007.

[113] C. von Savigny, C. S. Haley, C. E. Sioris, I. C. McDade, E. J. Llewellyn, D. Degenstein, W. F. J. Evans, R. L. Gattinger, E. Griffioen, E. Kyrölä, N. D. Lloyd, J. C. McConnell, C. A. McLinden, G. Mégie, D. P. Murtagh, B. Solheim, and K. Strong. Stratospheric ozone profiles retrieved from limb scattered sunlight radiance spectra measured by the OSIRIS instrument on the Odin satellite. *Geophys. Res. Lett.*, 30:1755–1763, 2003.

[114] C. von Savigny, J. W. Kaiser, H. Bovensmann, J. P. Burrows, I. S. McDermid, and T. Leblanc. Spatial and temporal characterization of SCIAMACHY limb pointing errors during the first three years of the mission. *Atmos. Chem. Phys.*, 5:2593–2602, 2005.

[115] E. S. Wachman and C. N. Pannell. High-performance hyperspectral imager using a novel acousto-optic tuneable filter. *Proc. SPIE*, 6966, 2008.

[116] J. Y. Wang and D. E. Silva. Wave-front interpretation with Zernike polynomials. *Appl. Opt.*, 19:1510–1518, 1980.

[117] J.W. Waters, L. Froidevaux, R.S. Harwood, R.F. Jarnot, H.M. Pickett, W.G. Read, P.H. Siegel, R.E. Cofield, M.J. Filipiak, D.A. Flower, and et al. The Earth observing system microwave limb sounder (EOS MLS) on the aura satellite. *IEEE T. Geosci. Remote*, 44:1075–1092, 2006.

[118] F. J. W. Whipple. The temperature of the atmosphere at levels accessible to air-waves. *Terr. Magn. Atmos. Electr.*, 38:13–16, 1933.

[119] WMO. *Scientific assessment of ozone depletion: 2010*, 2010.

[120] J. Xu and R. Stroud. *Acousto-optic devices: principles, design, and applications*. Wiley series in pure and applied optics. Wiley, 1992.

[121] T. Yano and A. Watanabe. Acoustooptic TeO_2 tunable filter using far-off-axis anisotropic Bragg diffraction. *Appl. Opt.*, 15(9):2250–2258, 1976.

[122] A. Yariv and P. Yeh. *Optical Waves in Crystals: Propagation and Control of Laser Radiation*. Wiley Series in Pure and Applied Optics. Wiley, 2002.

[123] C. Zehner, editor. *ATMOS 2012 Conference – Advances in Atmospheric Science and Applications: Session Summaries and Recommendations*. ESA, 2012.

[124] F. Zernike. Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. *Physica*, 1:689–704, 1934.

[125] F. Zernike. Refractive Indices of Ammonium Dihydrogen Phosphate and Potassium Dihydrogen Phosphate between 2000 ä and 1.5 μm. *J. Opt. Soc. Am.*, 54:1215–1219, 1964.