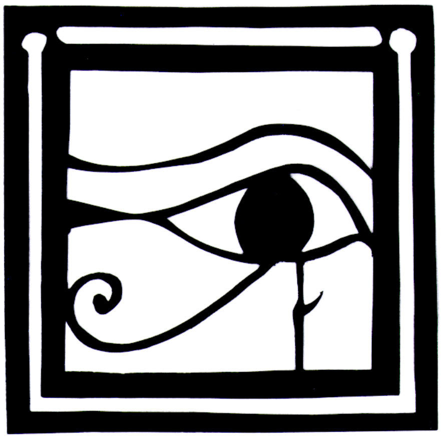
Atmospheric Water Vapour and Astronomical Millimetre Interferometry



Martina Corinna Wiedner

Mullard Radio Astronomy Observatory and
Gonville & Caius College, Cambridge

A dissertation submitted for the degree of
Doctor of Philosophy in the University of Cambridge
June 31, 1998

Summary

Atmospheric water vapour poses one of the major challenges in millimetre and submillimetre interferometry. In a manner similar to seeing at optical wavelengths, phase fluctuations, which are caused by changing amounts of water vapour at millimeter wavelengths, render deep, high resolution observations impossible unless atmospheric corrections are applied.

I have built the first radiometers which measure the strength of the water vapour transition line at 183 GHz. The difference in the measured amount of water vapour at the antennas of an interferometer directly translates into a phase shift introduced by the atmosphere. Thus, by monitoring the water vapour we can correct for phase shifts caused by the atmosphere. The instruments are installed at the interferometer consisting of the James Clark Maxwell Telescope (JCMT) and the Caltech Submillimeter Observatory (CSO), on Mauna Kea, Hawaii.

This thesis gives a brief overview of the distribution of water vapour as well as possible phase correction techniques for millimetre and submillimetre interferometers. Simulations of the 183 GHz line are presented emphasising the effect of temperature and pressure on the emission line profile and on the optical path. Optics, electronics and the performance of the 183 GHz water vapour monitors are described in detail. The monitors work well reducing the phase fluctuations from 55° to 23° root mean square (rms) under good weather conditions and from 127° to 48° rms in fair weather.

In addition, I have employed the JCMT-CSO interferometer to observe the ultraluminous infrared galaxy Arp 220 in the CO 3→2 transition and in adjacent continuum. Since single baseline interferometry does not allow full two-dimensional imaging, models have been fitted to the data. The best fitting model contains three sources of emission : a blue-shifted western

nucleus, a red-shifted north eastern nucleus separated by $1\farcs1$ from the western nucleus, and a south eastern nucleus separated by $0\farcs9$ which contains slightly blue-shifted gas, consistent with a model suggested by Downes and Solomon (1998).

Preface

This dissertation is the result of work undertaken in the Mullard Radio Astronomy Observatory, Cambridge between October 1994 and June 1998. The work described in this thesis is my own except where stated otherwise. Neither it, nor any similar dissertation has been submitted for a degree, diploma or other qualification at this, or any other university. This dissertation does not exceed 60,000 words in length.

Acknowledgements

First of all I thank my supervisor Richard Hills for suggesting this project and giving me advice and guidance throughout my time as a Ph.D. student. He introduced me to the principles of millimetre and submillimetre receivers and proposed the design of the water vapour monitors (WVMs).

During the planning and building phase of the WVMs I was supported by my colleagues from the MRAO workshop. I especially thank John Jennings, who designed most of the electronic circuits which were not commercially available. He helped order parts, soldered innumerable components onto circuit boards, tailored cables, always found the tools or instruments I needed and many more things. Roger Dace was responsible for transforming my mirror design into mechanical drawings and for arranging the fabricating of the mirrors. He also made the calibration loads and designed all the support structures. Thanks must go to Colin Hall who wrote the interface for the WVM PCs and created the initial design of the flip mirror control. Pierre Martin-Cocher improved the motor control and also soldered many connectors and cables. I thank all the other people from the workshop for their help as well as the friendly atmosphere: Brian Wood, who provided the 1.5 and 5 GHz low pass filters and who tolerated me using his bench and tools, Ken Duggan, who made a lot of the semi-rigid coaxial cables and repaired any voltmeters I had broken, Roy Jilley, who assisted John and made the printed circuit board of the V/F circuit, Hugh Gibson, who adapted the phase look loop design to suit the WVMs and who patiently helped me test and repair the detectors, Mike Carter, who fabricated many of the mechanical components and took most of the pictures of the instrument presented in this thesis, Basil Humphreys, who assisted him, and Bob Barker, who is absolutely perfect in packing WVMs in boxes for shipment to Hawaii.

Thanks also to Dick Dolling, who skillfully administered the group, booked flights, organised the paperwork for the apparatus and much more. Last but not least, I thank Dennis Bly: standing in front of his black board and listening to his excellent explanation I found that electronics suddenly became straight-forward and easy.

During my trips to Hawaii I learned a lot about interferometry from the SBI team consisting of Oliver Lay, John Carlstrom and Richard Hills, and about astronomy from the many observers. Many people gave me a hand testing and installing the WVMs in Hawaii and plotting the first data, in particular Bob Barker, Pierre Martin-Cocher, David Brown, the SBI team, Rachel Akinson, Christine Wilson and the staff of the JCMT and CSO. They also made my stay on the mountain very enjoyable.

On the astronomy side of this project, I thank Oliver Lay for explaining his data reduction software for interferometric observations. John Carlstrom provided the computer program which calculates binary separations from phase information. Christine Wilson helped me get all the software running. I immensely enjoyed the long discussions about the Arp 220 data with Andrew Harrison. John Richer brilliantly explained the principles of line radiation. Mike Jones and Claire Chandler were my main targets for questions about interferometry.

A big thanks to everybody in MRAO! I am sure that there is hardly anybody who has not contributed to this thesis either by answering at least one of my many questions or by making my stay more enjoyable. Thanks to the computer officers. Particular thanks must go to my office mates, especially Aled Jones and Ole Moeller, but also Steven Hancock, Christian Perwass and Sandi Setiawan, who not only tolerated me, but also helped with direct advice or by listening patiently. They gave the office a very special atmosphere and made it a great place.

For proof reading the thesis I would like to thank Richard Hills, John Richer, Hugh Gibson, Keith Baxter, Dennis Bly, Aled Jones, Andrew Harrison and my father. Lots of thanks to David Green, the Latex doctor, who always spotted straight away the faults my latex files were suffering from.

On the financial side, I am very grateful to have received the Sir Isaac Newton Scholarship. From May 1995 till April 1998 the Deutschen Akademischen Austauschdienst (German Academic Exchange Service) supported me with money from the second and third "Hochschulsonderprogramm" and took all financial worries off my shoulders, for which I thank it and the German tax payer.

Outside work I thank my most excellent friends, who were always cheerful and supportive.

Last but not least, I thank my parents, my brother and my relatives for their never-ending support, encouragements and their belief in me.

Contents

Chapter 1

Introduction and the JCMT-CSO Interferometer

1.1 Introduction

For thousands of years people all over the world have been fascinated by stars and planets and have observed and recorded them. Recently, however, a remarkable number of new discoveries have been made and our knowledge of the universe has increased greatly, partly due to new theories such as relativity and quantum mechanics, but perhaps even more due to the advance in instrumentation. Firstly, new instrumentation opened up the electro-magnetic spectrum allowing astronomical observations at wavelengths between 10^2 (radio) and 10^{-24} m (γ-ray). This led to the discovery of many objects such as quasars, young stars and also ultraluminous infrared galaxies, an example of which is presented in this thesis. Secondly, the resolution of observations has been enormously increased by new instrumentation, in particular by interferometers, which were first built for optical observations by Michelson in 1890 and later for radio observations by Ryle and Vonberg in 1946.

When developing instruments new problems are encountered which have to be overcome. The earth's atmosphere poses one of the major challenges as it interacts with electromagnetic radiation and hence limits the resolution as well as the wavelengths available for astronomical observations. One way of solving this problem is to avoid the atmosphere and observe from space. An alternative approach is to study the atmosphere and correct for its effect on electromagnetic radiation where possible. This thesis follows the second approach, investigates the properties of atmospheric water vapour and presents an instrument which can at least partly correct for atmospheric effects at millimetre and submillimetre wavelengths.

My Ph. D. project gave me the great opportunity to be exposed to all stages of developing an instrument, including the design, building, testing, installing the instrument and taking data with it. It also introduced me to a great many techniques relevant in radio astronomy such as Gaussian optics, electronics and data processing.

As an introduction there will be a brief description of the JCMT-CSO interferometer, for which the radiometers were built and which was used for my astronomical observations. Chapter 2 explains why water vapour causes phase fluctuations in millimetre and submillimetre interferometry and presents different techniques to correct them. Of these techniques we have chosen to build radiometers which observe the water transition line at 183 GHz. Chapter 3 shows simulations of the 183 GHz line and investigates which parameters determine the line shape and which determine the optical path length. Chapters 4 to 6 describe the radiometers themselves. Chapter 4 introduces the optics, chapter 5 the electronics, chapter 6 describes the performance of the instruments, gives estimates of their accuracy and shows some data. The final chapter presents some interferometric observations of the ultraluminous galaxy Arp 220.

1.2 James Clark Maxwell Telescope

The James Clark Maxwell Telescope (JCMT) (Fig. 1.1) is a submillimetre telescope [1] located at an elevation of 4090 m (13400 feet) on Mauna Kea in Hawaii, USA, at longitude $155°30'$, latitude $19°49'$. The telescope has an altazimuth mount, with a primary mirror of 15 m diameter and a surface accuracy of 25μm, resulting in signal loss of less than 3%. It is equipped with bolometers operating between 150 and 870 GHz (2000 - 350 μm) and with heterodyne receivers covering a frequency range of 210 to 505 GHz (1400 - 590 μm), soon up to 690 GHz (430μm).

[1] The JCMT is operated by the Joint Astronomy Centre on behalf of the Particle Physics and Astronomy Research Council of the United Kingdom, the Netherlands Organisation for Scientific Research and the National Research Council of Canada.

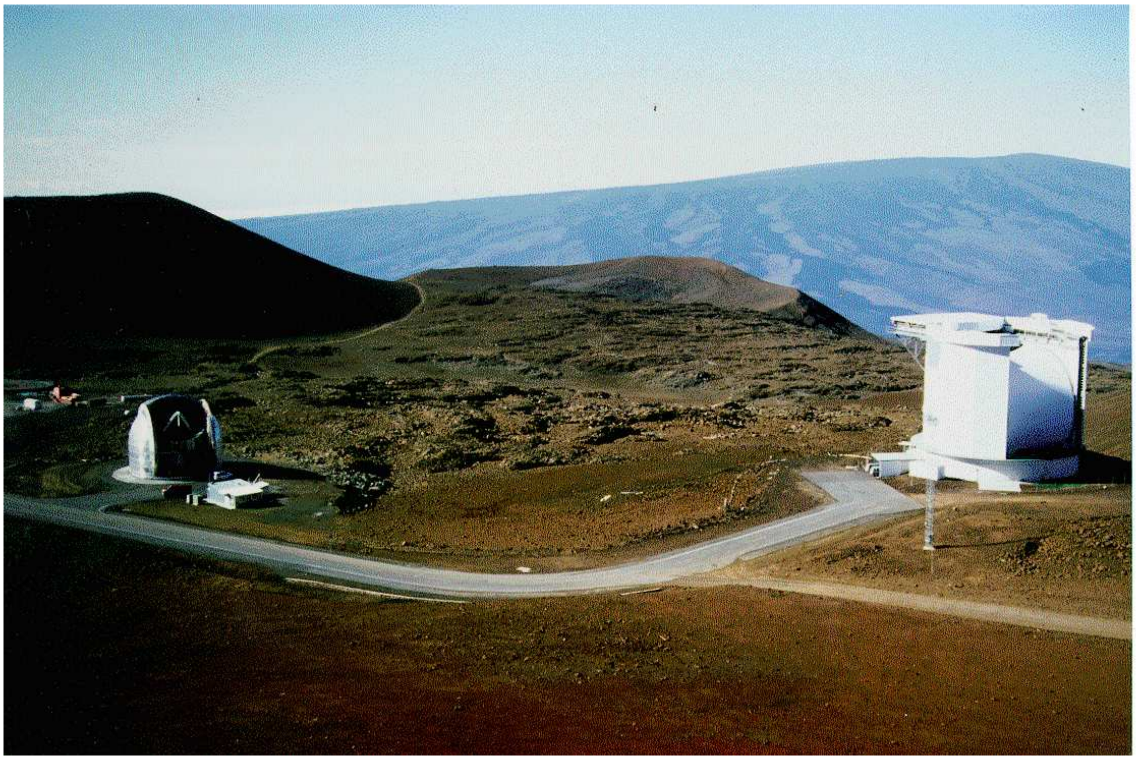


Figure 1.1. The Caltech Submillimeter Observatory (CSO) (left) and the James Clark Maxwell Telescope (JCMT) (right) on Mauna Kea in Hawaii.

1.3 Caltech Submillimeter Observatory

Approximately 160 m east of the JCMT is the Caltech Submillimeter Observatory (CSO) (Fig. 1.1), which is operated by the California Institute of Technology (Caltech) and funded by the National Science Foundation (NSF) and private money. Like the JCMT it has an altazimuth mount, but the primary mirror is only 10.4 m in diameter with a similar surface accuracy as the JCMT. The bolometers at the CSO can observe between 230 and 860 GHz (1300 - 350μm), the heterodyne receivers operate between 180 and 950 GHz (1670 - 315 μm).

1.4 JCMT-CSO Interferometer

In 1993 the two telescopes were connected to work as the first submillimetre interferometer (0.1 mm $\leq \lambda \leq$ 1 mm). The single baseline of the interferometer has an extent of -157.9 m in east-west direction, +43.4 m in polar direction and -1.9 m towards the celestial equator. The minimum fringe spacing at 350 GHz is $1\overset{''}{.}1$. Figure 1.2 shows the elevation and the projected baseline of astronomical sources as a function of their declination and hour angle.

The astronomical signal received by each antenna is mixed with the signal of a local oscillator (LO) to down-convert it to an intermediate frequency (IF) of 1.5 GHz. Since the two LO signals need to be highly coherent for interferometry both LOs are phase locked to a reference signal transmitted from the CSO to the JCMT via a wideband optical fibre. The same fibre also transmits the IF signal from the CSO to the JCMT. Due to the rotation of the earth the astronomical source moves through the fringes at up to $\sim$15 Hz at 350 GHz. Coarse path compensation is accomplished by sending the IF signals through delay lines of optical fibres. For fine path compensation numerically controlled oscillators are used to generate offset frequencies which are introduced into the local oscillator and IF chains in such a way that the fringe rate is removed.

After path compensation the two IF signals are mixed in the Dutch Autocorrelation Spectrometer (DAS) at the JCMT. The 1024 complex frequency channels of the DAS cover between 125 and 920 MHz bandwidth depending on the configuration. Assuming an efficiency η of 50% the conversion factor from the measured effective antenna temperature to Jansky is about 50 Jy/K.

The system temperature (T_{sys}) of the interferometer is the geometrical mean of the system temperature of the JCMT and the CSO and lies between 300 and 800 K mainly depending

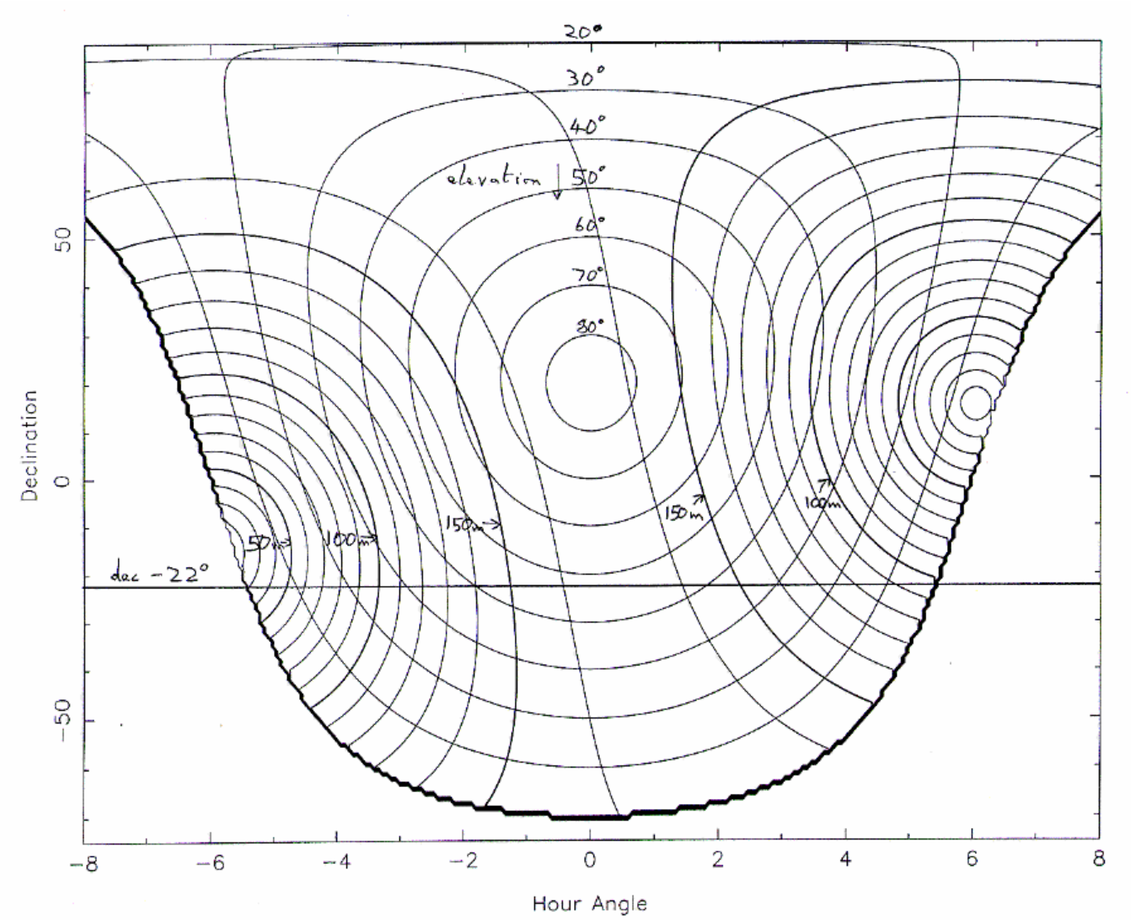


Figure 1.2. Diagram for JCMT-CSO interferometer to determine the elevation and projected baseline of a source given its declination and hour angle. The diagram is not symmetric about HA=0, because the baseline of the JCMT-CSO interferometer has a component in the north-south direction. The plot was created by R. E. Hills.

on the atmospheric emission. Depending on T_{sys} as well as the integration time Δt and the bandwidth $\Delta\nu$ or line width Δv of the observations the noise equivalent flux density σ is

$$\sigma_{cont} = 260\mathrm{mJy}\left(\frac{\eta}{0.5}\right)^{-1}\left(\frac{t_{int}}{10\mathrm{s}}\right)^{-0.5}\left(\frac{\Delta\nu}{1\mathrm{GHz}}\right)^{-0.5}\left(\frac{T_{sys}}{800\mathrm{K}}\right), \quad (1.1)$$

$$\sigma_{line} = 7.7\mathrm{Jy}\left(\frac{\eta}{0.5}\right)^{-1}\left(\frac{t_{int}}{10\mathrm{s}}\right)^{-0.5}\left(\frac{\Delta v}{1\mathrm{km/s}}\right)^{-0.5}\left(\frac{\nu_{sky}}{345\mathrm{GHz}}\right)^{-0.5}\left(\frac{T_{sys}}{800\mathrm{K}}\right), \quad (1.2)$$

where ν_{sky} is the observing frequency (Carlstrom et al., 1995).

The speed of the data-taking computers limits the minimum integration time to 10 s. Usually ten such 10 second observations are averaged together. Sources are usually observed for about 8 h, so that fluxes can be measured over a wide range of projected baselines, thus sampling the source at different resolutions. Images cannot be obtained from the observations with a single baseline interferometer, but spectra can be measured and models can be tested (see chapter 7).

Approximately every 10 minutes a calibrator is observed to monitor gain changes in the system. As the phase varies significantly on shorter time scales (30 s) the calibrator cannot be used to correct for phase fluctuations. Calibrators must have sizes much smaller than the minimum fringe spacing of $\sim 0\farcs5$ otherwise some of the flux will be resolved out by the interferometer. Usually quasars with a flux of a few Jansky are employed.

There are standard data reduction software as well as model fitting routines for the JCMT-CSO interferometer (Lay, 1994).

A much more detailed description of the JCMT-CSO interferometer is given in Oliver Lay's thesis (Lay, 1994). Thompson, Moran and Swenson (1994) give an excellent description of interferometry in general.

Chapter 2

Atmospheric water vapour and interferometric phase fluctuations

In millimetre and submillimetre interferometry, atmospheric water vapour causes phase fluctuations, which are generally the limiting factor for obtaining high angular resolution. Water vapour also attenuates the astronomical signal and increases the system temperature. It is therefore important to understand the properties of water vapour, its spatial and temporal spectra as well as its effects on electro-magnetic radiation. Correcting for phase fluctuations caused by atmospheric water vapour will be one of the main challenges faced by millimetre and submillimetre interferometers, and routine subarcsecond imaging will only be possible once this problem is solved.

This chapter will give some background on atmospheric water vapour. First it will describe its effect on electromagnetic radiation and how it causes phase fluctuations in interferometric observations (§2.1). In §2.2 we will look at the distribution of water vapour: what are typical amounts of water vapour on a dry site, such as Mauna Kea? How does the amount of water vapour change with height in the atmosphere? What are its temporal and spatial power spectra? The last section (§2.3) introduces different schemes for correcting phase fluctuation and contains a table listing their advantages and disadvantages. The section ends in justifying the choice of 183 GHz radiometers for the JCMT-CSO interferometer.

2.1 Effect of water vapour on astronomical observations

2.1.1 Phase fluctuations explained

Water vapour is not well mixed in the atmosphere and fluctuating amounts of water vapour cause phase fluctuations in interferometric observations.

To understand the cause of these fluctuations, consider the propagation of an astronomical signal through an atmosphere containing patches of water vapour, Fig. 2.1. At millimetre and submillimetre wavelengths the refractive index of air rises with the amount of water vapour it contains, so electromagnetic waves propagate more slowly through moist than dry air. In Fig. 2.1 the signal from a distant object has to pass through several patches of moist air to reach antenna 1. It will therefore reach antenna 1 later than antenna 2. An interferometer measures the difference in arrival time Δt between the two antennas as a phase

$$\Delta\phi = \nu\Delta t, \tag{2.1}$$

where ν is the observing frequency.

In general, there will be some wind and the patches of water vapour will be blown across the array, so that at a later time there might be more moist air along the line of sight to antenna 2 and the signal will reach antenna 2 later than antenna 1. As the patches of water vapour move in and out of the line of sight to the antennas the phase fluctuates.

This phenomenon is very similar to optical seeing, but seeing is caused by temperature gradients in the atmosphere and the fluctuations occur on time scales of tens of milliseconds rather than seconds and on spatial scales of tenths of metres rather than tens of meters.

2.1.2 Effects of phase fluctuations

The phase fluctuations basically have two effects on the observations.

1. *Phase fluctuations cause a loss in resolution.*
   Since the phase measures a difference in arrival time at the two antennas it contains information about the difference in optical path length and hence the location of the astronomical source. The phase information determines the location of the source to high accuracy, i. e. to high resolution. If the detected phase varies due to the atmospheric water vapour the resolution is degraded.

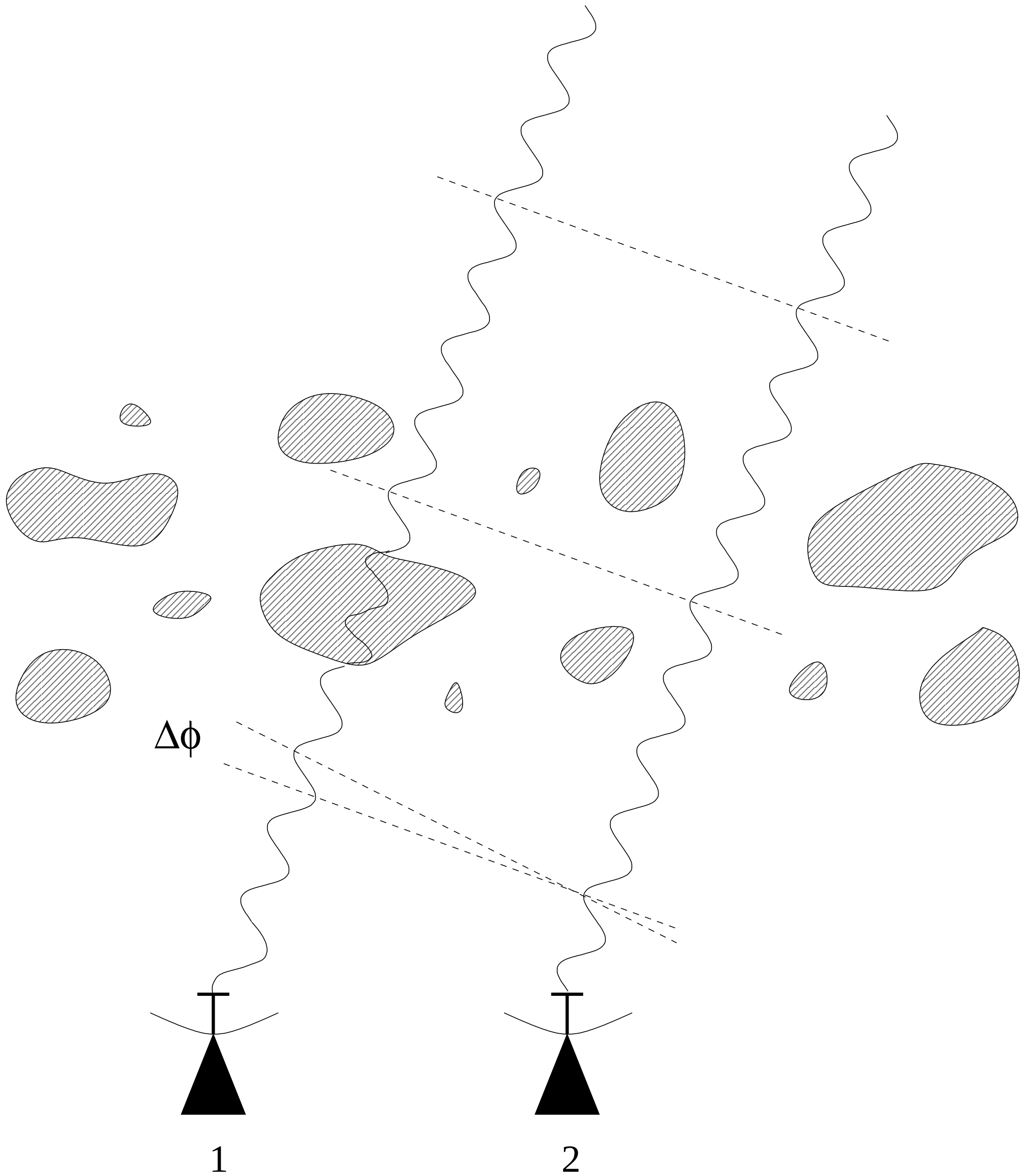


Figure 2.1. Atmospheric water vapour causes phase shifts in millimetre and submillimetre interferometry.

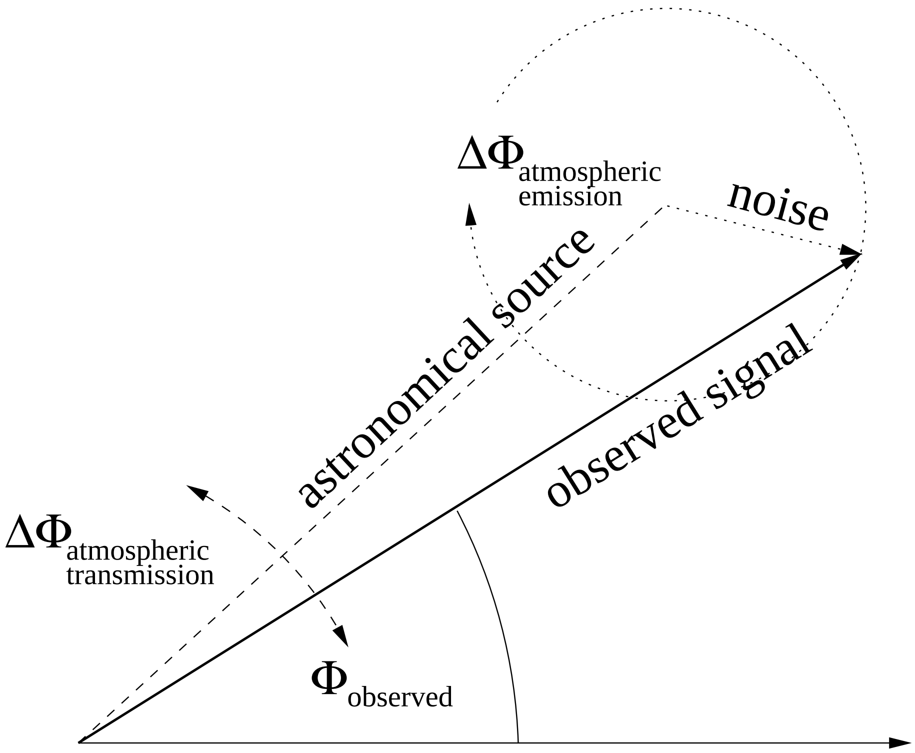


Figure 2.2. The signal observed by an interferometer can be thought of as the vector sum of the astronomical source and noise from atmospheric emission. The noise vectors of consecutive observations have approximately the same length but random phase ($\Delta\Phi_{atmospheric\ emission}$). The visibility vector from the astronomical source usually varies by $\Delta\Phi_{atmospheric\ transmission}$ due to phase fluctuations.

2. *Phase fluctuations decrease the signal strength.* The signal measured by an interferometer can be pictured as a visibility vector in the complex plane, Fig. 2.2. The length of the vector describes the measured intensity, the angle Phi to the real axis its phase. The vector will be composed of a component due to noise from atmospheric emission and a component due to the signal of the astronomical source (Fig. 2.2). An integration corresponds to taking the vector average of these visibility vectors (Fig. 2.3).

   Since the noise vectors have random phase ($\Delta\Phi_{atmospheric\ emission}$) they will sum incoherently and cancel each other out. Ideally the vectors from the astronomical source have a constant phase and so the vector averaged detection corresponds to the flux

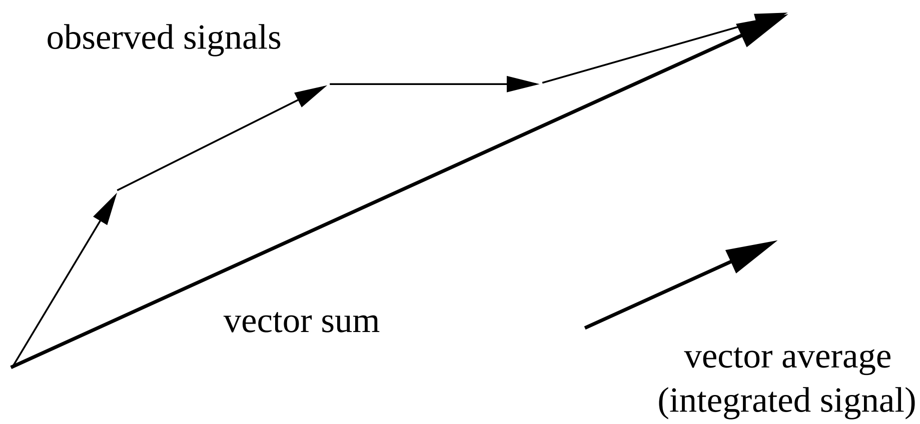


Figure 2.3. An integration corresponds to taking the vector average of many observed signals. If the vectors point in different directions the length of the vector averaged vector is shorter than the length of the individual vectors.

from the astronomical source. However, if water vapour causes the phase to change, the direction of the vectors will vary by $\Delta\Phi_{atmospherictransmission}$ and the vectors will partially cancel each other out. As a result the length of the averaged visibility vector will be shorter than that of an individual vector and the integrated flux will be less than the true flux from the source. The smaller $\Delta\Phi$, the better the agreement between the length of the vector averaged visibilities, and therefore between the measured integrated flux, with the true astronomical signal.

If the phase caused by water vapour can be measured or predicted, this phase can be subtracted from the measured phase. In our picture this corresponds to the voltage vector being turned back to the position it had before it entered the atmosphere. If these phase corrections can be applied successfully high resolution observations as well as long integrations on faint sources will be possible.

2.1.3 Atmospheric transmission

Besides changing the real part of the refractive index of air and causing phase fluctuations, water vapour also attenuates electromagnetic radiation. At submillimetre wavelengths water vapour is the main contributor to the atmospheric opacity. Figure 2.4 shows the transmission on Mauna Kea between 100 and 1000 GHz for different amounts of precipitable water vapour; where 1 mm pwv is defined as the amount of water vapour that would result in a 1 mm

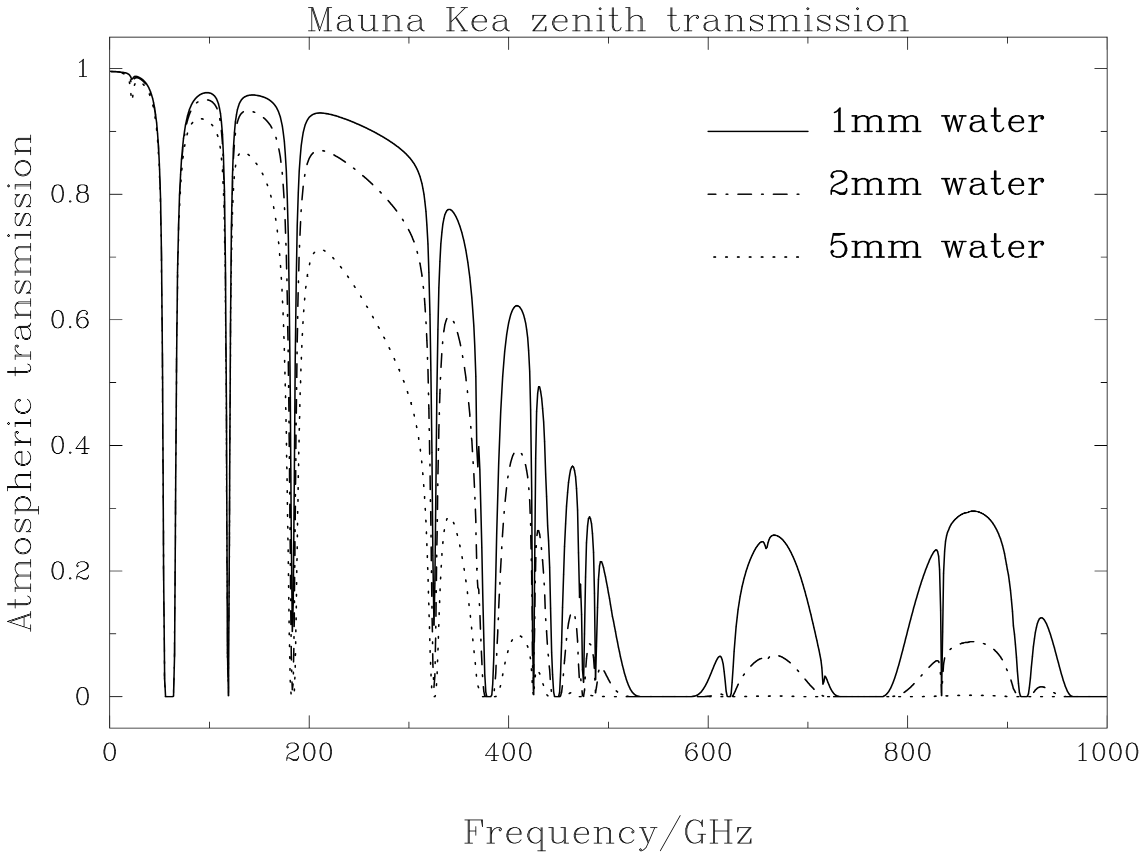


Figure 2.4. Atmospheric transmission for Mauna Kea with 1, 2 and 5 mm of pwv.

deep water layer when condensed on the ground. Unfortunately, in the case of attenuation astronomers cannot retrieve the lost signal.However, the signal of the astronomical source can be deduced by observing a calibrator of known flux, typically a planet, or by computing the attenuation from the amount of water vapour. In addition the attenuation can be reduce by choosing dry sites and by observing in the frequency windows with the best transmission properties.

2.2 Distribution of water vapour

In order to choose the best suited phase correction scheme and to optimise it, it is important to know the distribution of water vapour. In the case of the JCMT-CSO interferometer we would like to know the amount of water vapour typical for Mauna Kea. Knowing the time scale on which variations in water vapour content occur, the optimum rate at which phase corrections need to be applied can be determined. Ideally the phase monitor beam and the astronomical beam should traverse the same atmosphere. Usually, however, they are offset. Using the spatial distribution of water vapour, the expected phase error due to this offset can be determined.

For future interferometers, it is advantageous to know the distribution of water vapour on the earth in order to chose dry observing sites. For interferometers with varying baseline lengths, the baseline at which phase corrections become relevant can be determined from the spatial distribution of water vapour . As for the JCMT-CSO interferometer, optimum calibration time and acceptable beam offsets can be determined from the spatial and temporal distribution of water vapour.

The remainder of this section will give a brief introduction of the distribution of water vapour: its amount (§2.2.1), its distribution in height (§2.2.2), its spatial distribution (§2.2.3) and how it varies with time (§2.2.4).

2.2.1 Amount of water vapour

The less water vapour there is, the better is the atmospheric transmission and the smaller are the phase fluctuations.

The amount of water vapour at different locations on the earth varies by orders of magnitude. Since cold air is less able to hold water vapour than warm air, sites at high altitude or in a colder climate have lower amounts of precipitable water vapour. Particularly dry sites are for example Mauna Kea (4200 m) in Hawaii, the Andes (3000 m to 7000 m) in Chile and Antarctica (3000 m). The atmosphere above these sites contains on average 0.5 to 2 mm pwv in comparison to perhaps 10 to 30 mm in Cambridge. There are large variations on the amounts of water vapour ($\geq$50%) on time scales of weeks or months, but these changes are not easily predictable. However, the diurnal changes of pwv show a very regular pattern. In Fig. 2.5 quartiles of opacity at 225 GHz, which is proportional to pwv, are plotted as a function of local time. As the sun warms up the atmosphere moist air from lower altitudes rises to the summit. The pwv has a maximum at around 2 pm and by 8 pm the moist air will have settled to lower altitudes. (This diurnal change of water vapour and the increase in turbulence at day time are the main reasons for obtaining most submillimetre observations during night time.) Fig. 2.5 also gives a quantitative estimate of the amount of water vapour: there is less than 2 mm pwv during half of the nights and less than $\sim$ 1.2 mm during a quarter of the nights. In subsequent sections 1 and 2 mm pwv will often be used in example.

2.2.2 Distribution of pwv with height

To first approximation precipitable water vapour decreases exponentially with height (see §3.2.1). However, the three measurements of humidity as a function of height for the same location, but on different days plotted in Fig. 2.6 show that the profiles actually vary considerably. Common to all three data sets is that the water vapour decreases with height and that there is little water vapour above 3000 m. This is, of course, the reason why millimetre and submillimetre telescopes are built at high altitude. Above 6000 m there is nearly no water vapour. In the case of the JCMT-CSO interferometer this implies that practically all of the water vapour observed is located within the first 2 km above the telescope.

2.2.3 Spatial power spectrum

In Kolmogorov theory, energy, e.g. from wind and convection, is injected into the atmosphere at an outer scale (Lay, 1997a) . If the air flow is not smooth, turbulence occurs. This turbulence is transfered down to smaller scales conserving the kinetic energy until it dissipates at an inner scale. This turbulence determines the distribution of water vapour, which is the cause for changes in the refractive index.

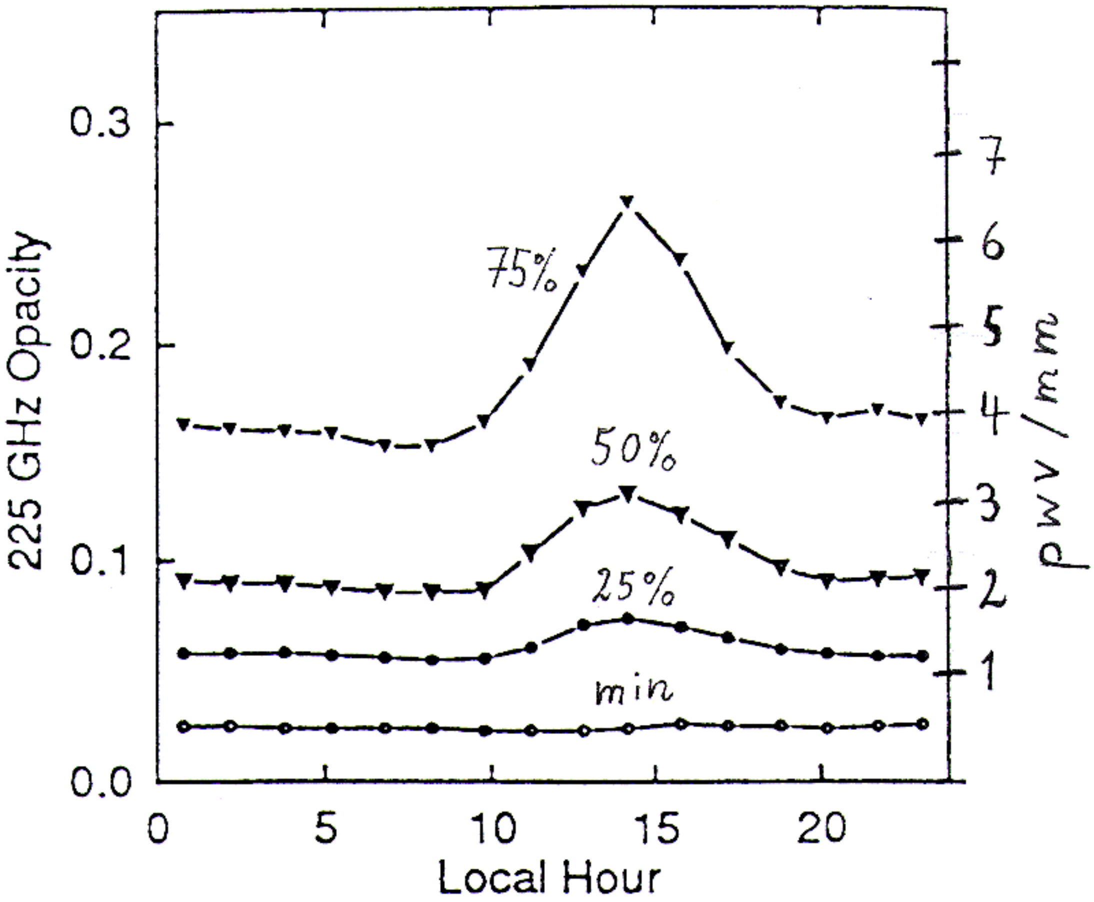


Figure 2.5. Diurnal changes of pwv on Mauna Kea, Hawaii. The 4 curve represents three quartiles of the amount of pwv as well as the minimum amount. The opacity at 225 GHz is related to the amount of pwv by $\tau_{225\ GHz} \approx 0.01 + 0.04\mathrm{pwv/mm}$ (JCMT, 1998). Figure taken from Masson (1992).

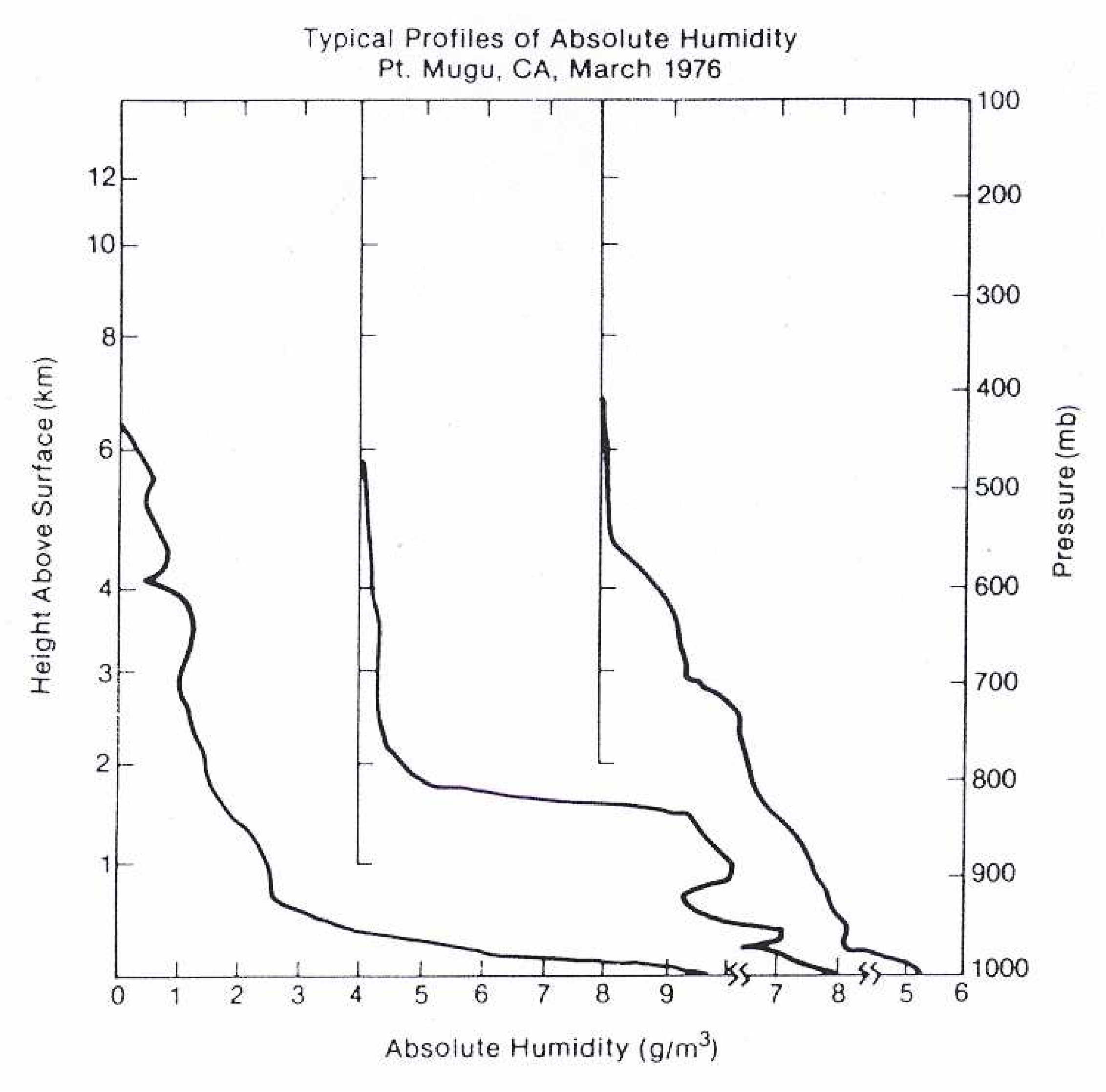


Figure 2.6. Distribution of pwv with height measured on three different days. Figure taken from Westwater (1993).

Radiation from a distant point source results in a plane wave front which has a constant phase in (x,y), a plane perpendicular to the direction of propagation. The uneven distribution of water vapour in the atmosphere distorts the plane wavefront and the phase in (x,y) varies. The phase structure function $D(\boldsymbol{b})$ is defined as the average of the squared difference of the phases at location $\boldsymbol{x}$ and at location $\boldsymbol{x}+\boldsymbol{b}$ (Thompson et al., 1994):

$$D(\boldsymbol{b}) \equiv \langle(\phi(\boldsymbol{x}) - \phi(\boldsymbol{x}+\boldsymbol{b}))^2\rangle. \tag{2.2}$$

Assuming that the turbulence is isotropic, it can be shown that the phase structure function is a power law function of the baseline length $|\boldsymbol{b}|$ (Tatarskii, 1961):

$$D(\boldsymbol{b}) = D(|b|) = \frac{K^2}{\lambda^2} b^{2\alpha}, \tag{2.3}$$

where K is a constant depending on the site. For baselines b shorter than the typical height of the turbulent layer the turbulence must be treated as three-dimensional (thick screen) and α=1/3 is predicted. For baselines longer than the typical height the atmosphere can be approximated to be two-dimensional (thin screen) and α becomes 5/6 (Coulman, 1989). Beyond an outer scale the structure function is expected to be flat (α=0) (Carilli and Holdaway, 1997).

The spatial power spectrum $\Phi(\boldsymbol{f})$, which is often mentioned, relates to the phase structure function in the following way (Conan et al., 1995):

$$\begin{aligned} D(\boldsymbol{b}) &\equiv \langle(\phi(\boldsymbol{x}) - \phi(\boldsymbol{x}+\boldsymbol{b})^2\rangle \\ &= 2\langle(\phi(\boldsymbol{x}))^2\rangle - 2\langle\phi(\boldsymbol{x})\phi(\boldsymbol{x}+\boldsymbol{b})\rangle \qquad (2.4) \\ &\equiv 2\left(B(0) - B(\boldsymbol{b})\right), \qquad (2.5) \end{aligned}$$

where $B(\boldsymbol{b})$ is the correlation function. The Fourier transform of the correlation function $B(\boldsymbol{b})$ is the spatial power spectrum $\Phi(\boldsymbol{f})$, which is proportional to the spatial frequency vector $\boldsymbol{f}$ to the power of -11/3 for $\alpha = 1/3$ (St-Jacques, 1998)

$$\Phi(\boldsymbol{f}) \propto \boldsymbol{f}^{-11/3}. \tag{2.6}$$

For $\alpha = 5/6$ the exponent becomes -8/3. In both cases the power in the phase fluctuations increases with baseline.

The theoretical predictions agree reasonably well with observations. Carilli and Holdaway (1997) measured the root mean square (rms) phase variations, which is equal to the square root of the structure function, with the VLA (Very Large Array in New Mexico, USA) on baselines between 200 m and 33 km. The data are plotted in Fig. 2.7. The data show a transition from the thick screen model to the thin screen model at $\sim$ 1200 m and an outer scale of $\sim$6 km. The measured slopes agree well with theoretical phase structure function.

2.2.4 Temporal power spectrum

The temporal power spectrum can be derived from the spatial power spectrum using the Taylor hypothesis, which states that the turbulence is frozen. If one imagines the water vapour to be concentrated in "clouds" of moist air (as in Fig. 2.1) frozen turbulence means that this screen of clouds moves without breaking up. This assumption appears to be a reasonable approximation when modeling existing interferometers, because the time for breaking up such a "cloud" of high humidity is longer than the time it spends traversing the array (Lay, 1997a) . Using the frozen screen assumption the temporal power spectrum P can be shown to be proportional to frequency (i.e. inverse of time between to measurements) to the power of -2/3 for small frequencies and -8/3 for larger frequencies (Conan et al., 1995).

Measurements taken at OVRO (Owens Valley Radio Observatory, California, USA) with a phase monitor of 100 m baseline are in good agreement with theory (Lay, 1997a). Figure 2.8 shows power spectra of data taken on different days. In (c), (d), (g) and (h) $\log(P)$ is plotted against $\log(\nu)$ showing that the slopes of -2/3 and -8/3 fit the data reasonably well. Plots (a), (b), (e) and (f) show the same data, but here $P \times \nu$ is plotted against $\log(\nu)$. In these plots, the area under any section of the plot is proportional to the phase power in that frequency section. The plots indicate that most of the phase power lies between 0.001 and 0.1 Hz, which corresponds to time scales between 1000 and 10 seconds. If one could for example correct the phase on time scales longer than 10 seconds the phase power would be reduced to less than 10%. Lay (1997b) discusses in much more detail which part of the temporal power spectrum is corrected by the different phase correction techniques and by how much the data quality is improved.

2.2.5 Summary

1. The amount of water vapour varies significantly with location on the earth and with

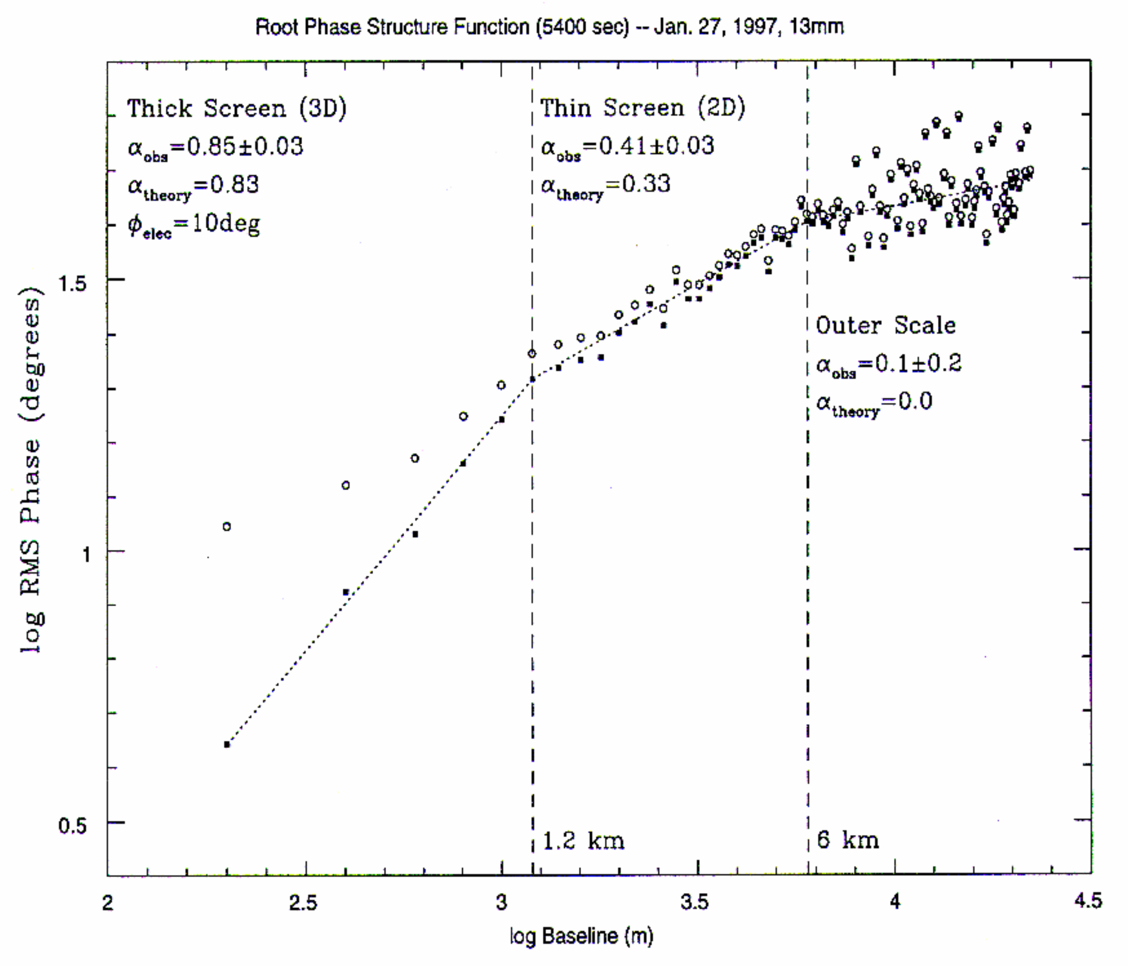


Figure 2.7. The root phase structure function from observations at 13 mm in the BnA array of the VLA on January 27, 1997. The open circles show the rms phase variations versus baseline length measured on the VLA calibrator 07848+240 over a period of 90 min. The filled squares show these same values with a constant electronic noise term of 10° subtracted in quadrature. The measurements agree well with the predictions from Kolmogorov theory. Figure by Carilli and Holdaway (1997).

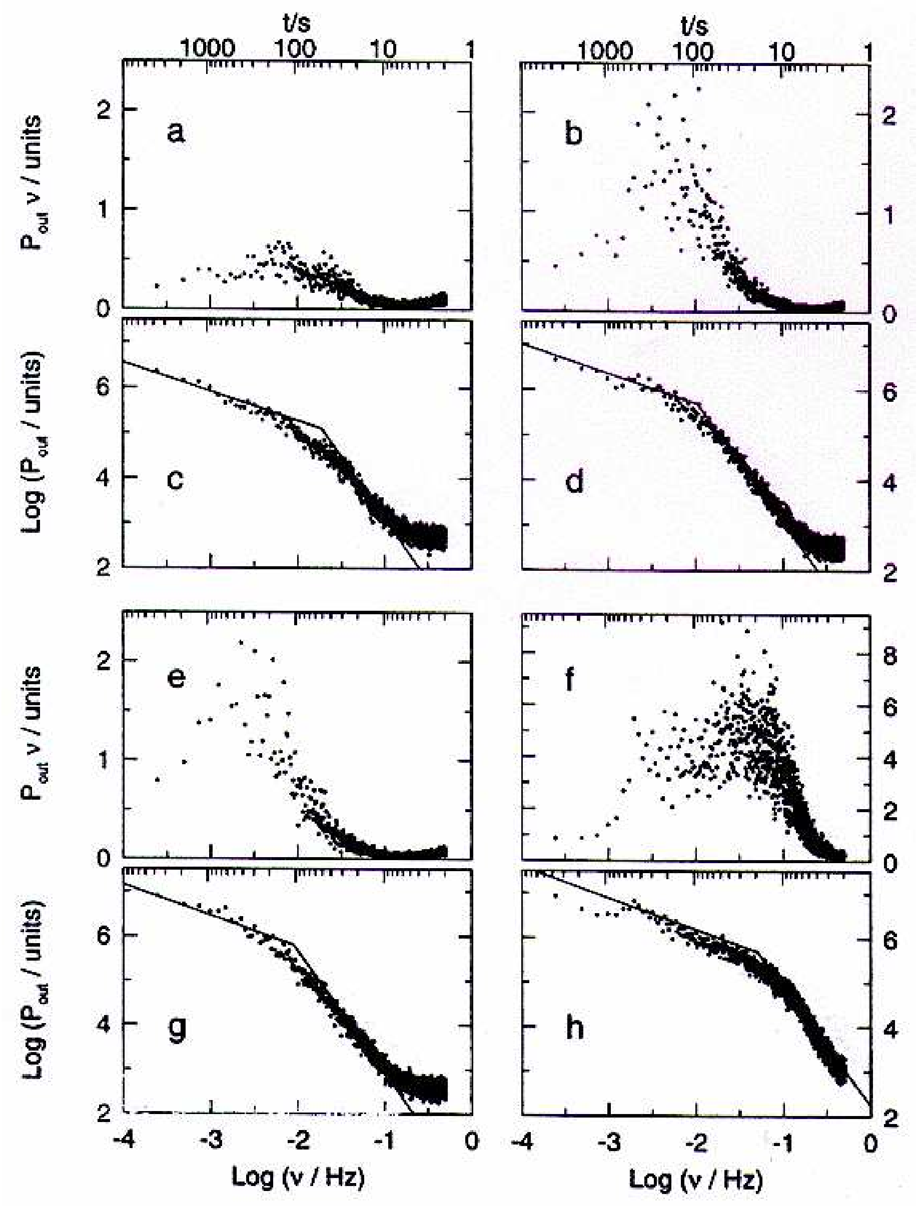


Figure 2.8. Phase power plots and Log-log plots for 4 days on the Owens Valley (Lay, 1997a). (a) and (c) are for Oct. 25 1995, and illustrate a very calm atmosphere, setting limits on the instrumental noise contributions; (b) and (d) are for Feb. 4 1995 showing typical conditions; (e) and (g) are for Feb. 5 1995 which show that there can be substantial power on long time scales; (f) and (h) are for Jan. 12 1995 and show substantial power on short time scales. The lines on the Log-Log plots have gradients of -8/3 and -2/3.

time. There is less water vapour at high altitudes and in cold air.

2. To first approximation all the water vapour is in the troposphere (0-10 km) and decreases exponentially with height.

3. Phase fluctuations increase with baseline, as predicted by Kolmogorov theory, up to an outer scale of about 6 km.

4. Most of the phase power lies on scales of 10 to 1000 s. If phase corrections are performed every 10 s, ~90% of the phase noise can be removed.

5. At Mauna Kea we expect between 0.5 and 4 mm pwv during night time. Corrections should be applied at least every 10 s, and the monitor and astronomical beams should not be further apart than $\sim$ 50 m ($v_{wind} \times 10$s) at the height of the water vapour.

2.3 Phase correction schemes

From the above it is clear that phase correction is essential, especially if observations either with long baselines, with long integration times or at high frequencies are desired. The institutions operating BIMA (Berkeley Illinois Maryland Array in California, USA), IRAM (Institut de Radio Astronomie Millimétrique) and OVRO are all in the process of extending their current baselines. The JCMT-CSO interferometer will soon be operating at 690 GHz. In addition new interferometers such as the SMA (Smithsonian Millimeter Array on Mauna Kea, Hawaii) and the MMA (Millimeter Array) will observe at high frequencies (100 - 1000 GHz) over long baselines ($\sim$ 10 km).

All of the interferometers have tried to reduce the phase fluctuations by choosing relatively dry sites usually at high elevation. In addition, all have tried to correct the phase, though, depending on the interferometer, there have been different approaches. In addition detailed research about phase calibration techniques has been carried out in preparation for the MMA/LSA array, in particular by Holdaway (1992), Woody at al. (1995) and Schilke et al. (1998) .

The phase correction schemes can be divided into two major groups:

1. The phase fluctuations are *measured* directly using a calibrator and subtracted when observing a faint astronomical source.

2. The water vapour content is measured and used to *calculate* the phase correction.

2.3.1 Phase correction by measuring the phase

A bright calibrator is used whose phase is easily measurable. This phase is then subtracted from the observations of a fainter object.

1. Calibrator and astronomical source are observed *alternately.*

   (a) Fast switching phase calibration (FS)

   In standard radio astronomy observations are calibrated against the amplitude and phase of a quasar, which is observed in time intervals T whose durations range between a few minutes to a few hours depending on the stability of the atmosphere and instrument. At millimetre and submillimetre wavelengths the phase fluctuations over short times (minutes to seconds) significantly degrade the signal, therefore the quasar has to be observed more frequently (fast switching phase calibrations, T $<$ 5 minutes). Holdaway and Owen (1995) have performed tests at 22 GHz with the VLA. They used a switching cycle of 80 s. Figure 2.9 shows their results. They can considerably reduce the phase errors on baselines longer than 500 m. At short baselines the phase errors are smaller and the phase needs to be determined quite accurately to further reduce the errors. The method is not very successful on short baselines for two reasons: firstly the quasar is observed at a different time than the astronomical source and over this time the phase introduced by the atmosphere will have changed. Secondly the quasar is at a different position than the astronomical source, therefore different parts of the atmosphere are probed. Therefore phase correction could be improved in two ways:

   i. A shorter cycle time will allow to correct for fast fluctuations in the atmospheric phase. However, for short cycle times the antennas have to be very agile and slew quickly. The set up time also needs to be short. The combination of system temperature and quasar flux give a lower limit of the cycle time because the quasar phase must be measured with high signal to noise.

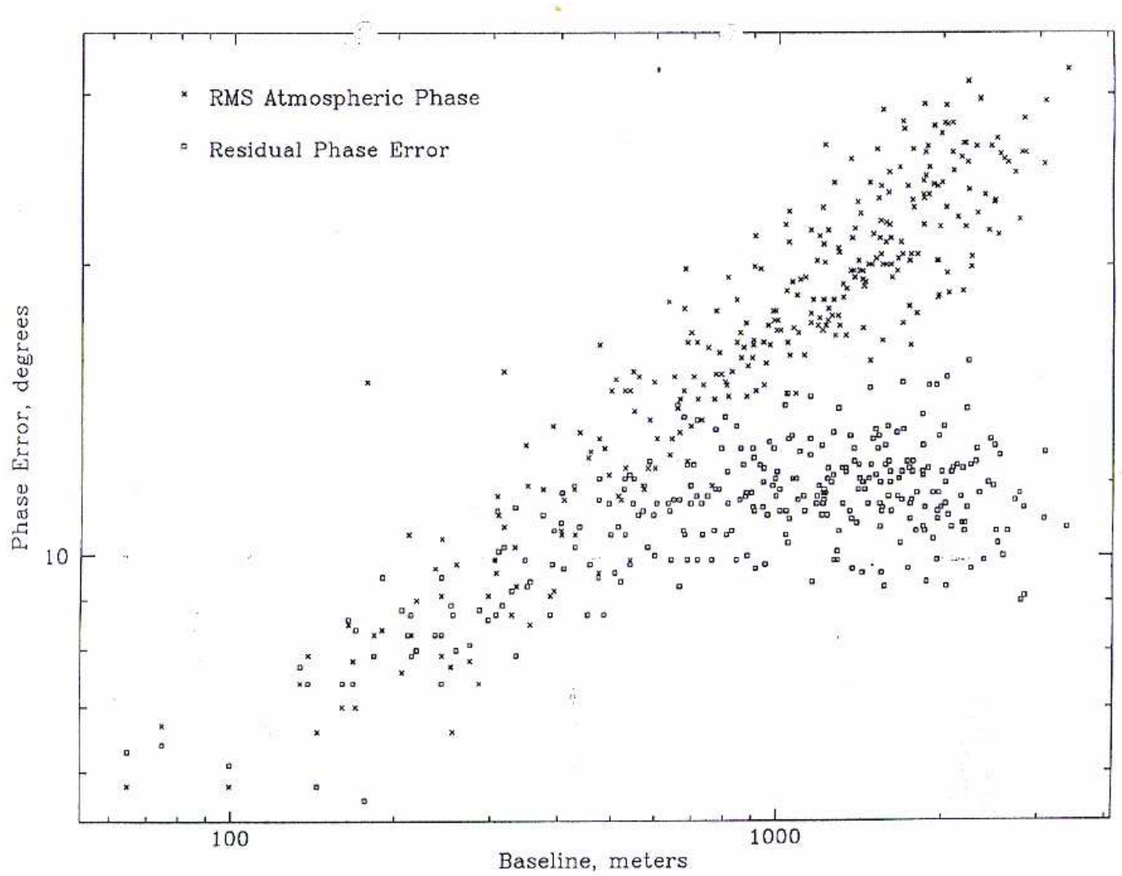


Figure 2.9. The crosses represent the uncorrected phase during observations of the source 2131-021 with the VLA. The open squares show the residual phase after calibrating every 80 s on the nearby source 2134+004. Fast switching phase calibration reduces the rms phase on baselines longer than 500 m. Figure by (Holdaway and Owen, 1995).

ii. According to the phase structure function (§2.2.3) the smaller the angle between the quasar and the astronomical source of interest, the smaller is the phase difference. However, since there are not many bright quasars at millimetre wavelengths, often only faint quasars are in the vicinity of the astronomical source. To detect faint quasars the integration time and therefore the cycle time has to be increased. Alternatively, one can observe the quasar at a different frequency, where there are more bright calibrators. This however might require extra time in setting up the system. It also requires extrapolating the phase from the frequency the quasar is observed at to the frequency the astronomical source is observed at.

(b) Beam switching (BS)

Beam switching is nearly identical to fast phase calibration, except that only the secondary mirror is moved when the calibrator is observed rather than the whole telescope. The secondary mirror can be moved much faster than the primary but it cannot be moved very far. Since the density of submillimetre calibrators is lower than 1 per 2° (Holdaway, 1992), the calibrators have to be observed at lower frequency (e.g. 30 GHz).

2. Astronomical source and calibrator are observed *simultaneously.*

(a) Subarray for monitoring the phase (Sub A)

Rather than observing source and calibrator at different times, some antennas of the array could monitor the calibrator and the others observe the astronomical source. [1] Then there will be no error due to extrapolating the phase in time, but there will be a spatial and directional offset between the line of sight of the monitoring and the observing antennas, such that each antenna probes a slightly different part of the atmosphere.

Asaki et al. (1997) have tested this correction scheme. Instead of observing a bright quasar to obtain the phase measurement they observed a geostationary Japanese Communication Satellite at 19 GHz. Figure 2.10 shows the measured and the

[1] If the antennas are arranged in pairs (one monitoring, one observing) this scheme is called the paired antenna method (PAM). If the monitoring antennas are distributed in a different way one refers to a subarray for monitoring the phase.

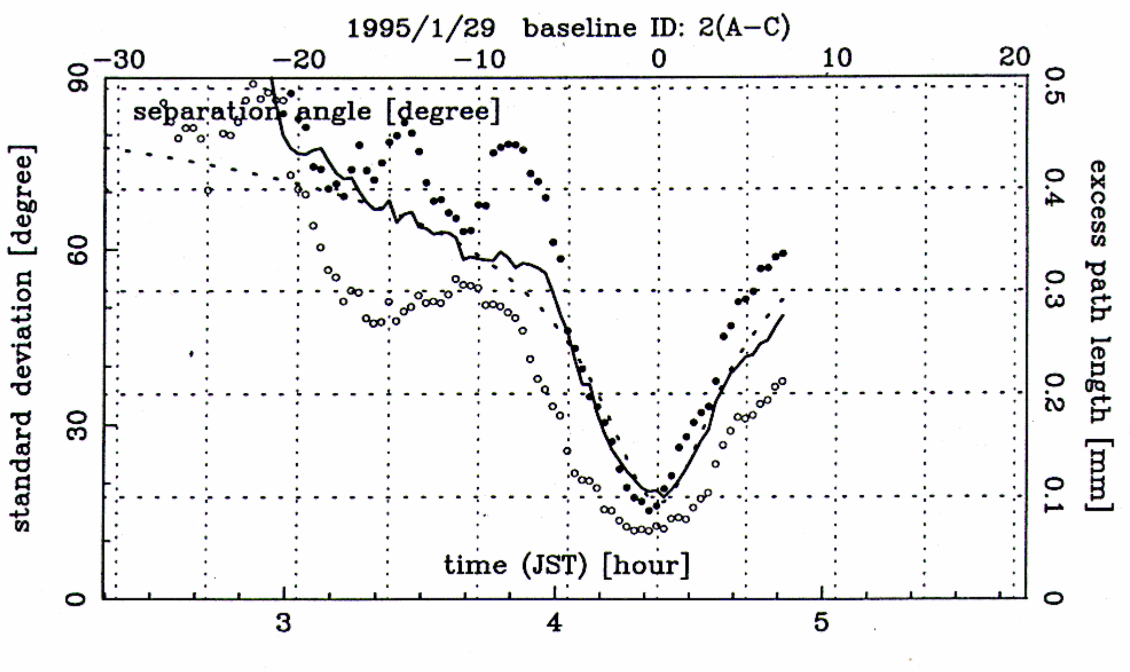


Figure 2.10. The standard deviations of the fringe phase difference is calculated for the time interval of 1000 s for a baseline of 250 m. The rms phase, which is related to the excess path length, is plotted against Japanese standard time, which directly relates to the separation angle of the satellite and the quasar (scale on top of graph). Filled circles show the quasar phase after simply subtracting the satellite phase. If the satellite and the quasar are at different elevations, the phase correction will give lower rms values when the quasar phase is corrected with the satellite phase measured at a slightly different time. The open circles show the best results with time-lag-inserted correction. The lower rms value can be easily explained with the frozen screen model: the water vapour "clouds" will take a certain time to travel from the line of sight to the satellite to the line of sight to the quasar. For both types of correction the gradual increase on the standard deviations with the growing separation angle is evident. The solid line and the dashed line show the result of a simulation based on a statistical model described in Asaki et al. (1997).

corrected phase as a function of the angle between the geostationary satellite and 3C279. Obviously, the smaller the angular separation between the the satellite and the quasar is, the better does the phase correction work. For separation smaller than 5° the rms of the corrected phase is less than 50°. From this experiment it is clear that astronomical source and calibrator or satellite need to be closer than 5° for the phase correction to give good results.

(b) Atmospheric monitoring array (Atm A)

Instead of using the array antennas for measuring the phase, simpler and smaller monitoring antennas could be built. As well as being cheaper, due to their size these antennas could also be closer to the observing antennas and hence the spatial offset can be reduced.

(c) Dual beam (DualB)

Alternatively, the antenna optics could be arranged in a way so that the astronomical source and a calibrator 1° to 2° away could be observed simultaneously (Holdaway, 1992).

2.3.2 Phase correction by measuring the water vapour

Instead of measuring the phase directly it is also possible to measure the amount of water vapour and deduce the phase shift it introduces. The water vapour has transition lines, but there is also an apparent underlying continuum emission (which probably is due to line wings of saturated transition lines in the infrared) so that it can be detected at any millimetre or submillimetre wavelengths, see Fig. 3.1.

1. Total power method (TP)

   In the total power method astronomical receivers are used to measure the water vapour. The variations in the total power signal will be dominated by sky emission, from which the amount of water vapour and hence the phase can be deduced. In the case of dual frequency receivers the signal at one frequency is correlated with the signals from all the other antennas and used for observing the astronomical source. The signal at the other frequency is measured at each antenna without correlating it (total power). This method has been applied successfully at the IRAM Plateau de Bure interferometer near Grenoble, France (Bremer, 1995). Figure 2.11 shows from top to bottom the phase of the astronomical source observed at 86 GHz, the sky at 230 GHz and the corrected phase of the astronomical source. Usually the raw rms phase lies between 30° and 50° and the rms of the corrected phase between 15° and 35°. There have been very successful corrections where an rms phase of 50° was reduced to 12° over a 1.5 hour period. BIMA has run similar tests, but found that the stability of the astronomical receivers needs to be extremely good to get satisfactory results (Plambeck et al., 1996). Tests have also been carried out by the NRAO.

2. Radiometric phase correction (RPC)

   Water vapour radiometers are independent instruments, which measure the water vapour at the frequency of one of its transition lines. They have the advantage that they can be added to any telescope and that they do not take up observing time or observing frequencies. So far two radiometers have been built one at 22 GHz for OVRO (Woody, 1996) and the other at 183 GHz, which is described in this thesis.

Table 2.1 summarises advantages and disadvantages of the different phase correction techniques discussed above.

2.3.3 Choice of 183 GHz radiometer for the JCMT-CSO

For phase calibration at the JCMT-CSO interferometer we chose to build 183 GHz radiometers for the following reasons:

1. The JCMT and the CSO were designed independently and are operated as single dish telescopes most of the time. Therefore a phase correction scheme that easily could be added later without disturbing the single dish performance was needed.

2. Since the JCMT-CSO interferometer operates at high frequencies (200 to 460 GHz, soon 690 GHz), short time fluctuations (<10 sec) need to be corrected to reduce the rms phase fluctuations below 30°.

3. The cost of the phase correction scheme had to be low.

4. As will be discussed in the next chapter, the 183 GHz line is more sensitive to small changes in pwv especially at dry sites such as Mauna Kea.

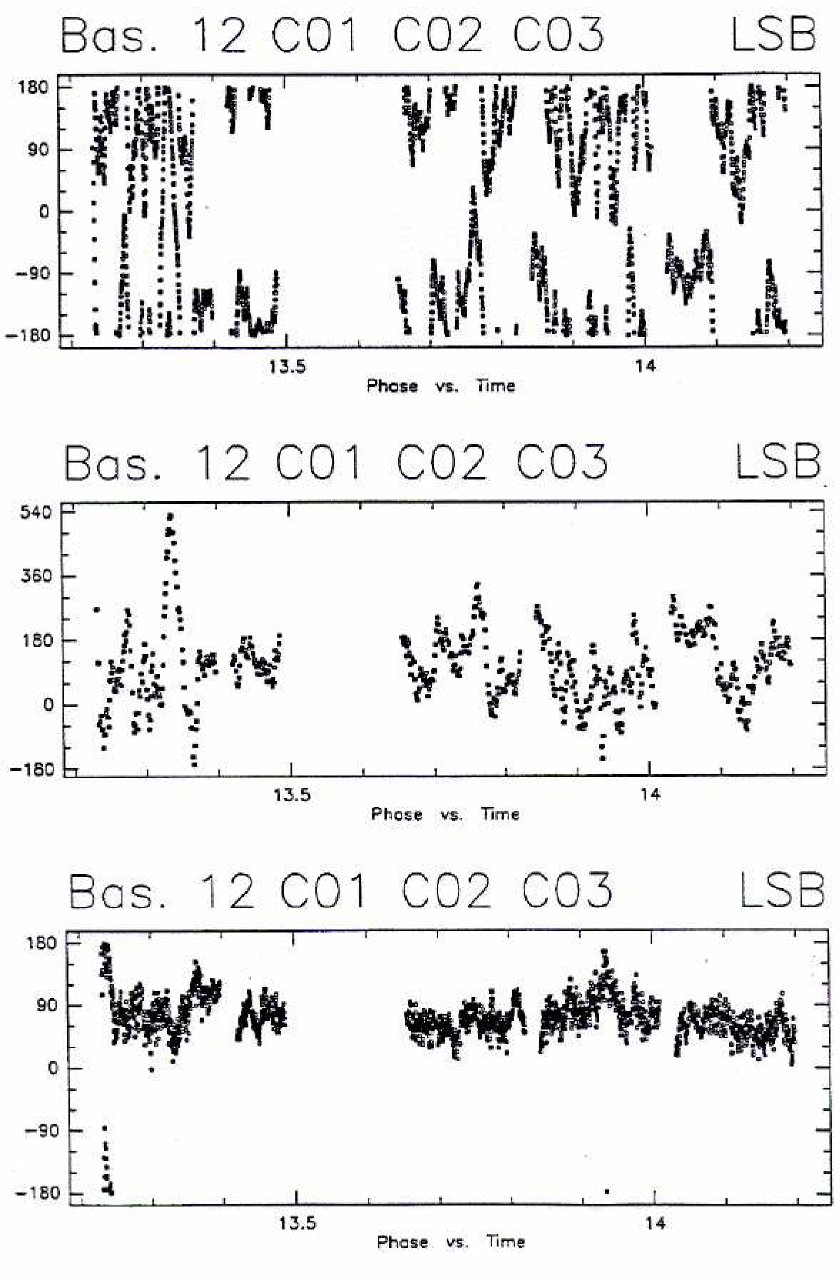


Figure 2.11. Phase correction with the IRAM plateau de Bure interferometer (Bremer, 1995). The top plot show the uncorrected phase of the quasar 0923+392 at 86 GHz; the middle plot show the total power measurement at 230 GHz; the bottom plot depicts the corrected phase.

Technique	FS	BS	Sub A	Atm A	DualB	TP	RPC
Section	1.1(a)	1.1(b)	1.2(a)	1.2(b)	1.2(c)	2.1	2.2
1) Uncertainty due to							
a) differences in							
time	$\geq$10 s	1 s	—	—	—	—	—
elevation	$2-5^\circ$	1°	$2-5^\circ$	$2-5^\circ$	1°	—	$5'$
position	—	—	$\sim$15 m	$\sim$8 m	—	—	—
b) extrapolation							
$\nu_1 \to \nu_2$	—	$\surd$	—	—	$\surd$	—	—
pwv $\to \nu$	—	—	—	—	—	$\surd$	$\surd$
2) Independent instr.	—	—	—	$\surd$	—	—	$\surd$
cost	—	—	—	high	—	—	low
3) Special requ. on							
interferometer's							
mechanics	$\surd$	$\surd$	—	—	—	—	—
optics	—	—	—	—	$\surd$	—	—
Rx backend	—	—	—	—	$\surd$	$\surd$	—
set-up time	$\surd$	$\surd$	—	—	—	—	—
4) Loss in							
obs. time	$> 30\%$	$\sim 30\%$	—	—	—	—	—
obs. frequ.	—	—	—	—	—	50%	—
baselines	—	—	75%*	—	—	—	—
sensitivity	—	—	$\leq 50\%$	—	—	—	—

Table 2.1. Some of the advantages and disadvantages of the different phase correction techniques. The numbers in the second row stand for the section in §2.3 where the technique is described. (1) The first part of the table lists contributions to the phase error, for example that the measurement of the phase and the astronomical source are performed at different times or that an extrapolation from one frequency to another is needed. (2) The second section notes whether the correction technique involves independent instruments. (3) The third section lists special requirements on the interferometer antennas and receiver (Rx) backends for the phase correction to work. (4) The last section shows e.g. how much observing time is needed for the phase calibration. * If an interferometer consists of N antennas, it has 1/2(N)(N-1) baselines. If there are many antennas and half of them are used for phase monitoring, 75% of the baselines are lost.

5. Because of the high sensitivity of the 183 GHz line an uncooled system could be built, which makes maintenance much easier.

6. The optics of the 183 GHz radiometers are more compact than optics at 22 GHz.

Chapter 3

Water vapour line at 183GHz

Chapter 2 explained how fluctuating amounts of water vapour introduce phase shifts in millimetre and submillimetre interferometry and looked at the distribution of water vapour. Several different methods for phase correction were introduced. We have chosen to build dedicated radiometers which measure the intensity of the water vapour emission line at 183 GHz.

This chapter will look at the 183 GHz line and the conversion from measured line emission to interferometric phase shift. Section §3.1 compares the strength of the 183 GHz line with other emission lines. Formulae describing the line profile as well as the excess optical path caused by a given amount of the water vapour are introduced in §3.2. Pressure and temperature of the water vapour change the line profile. Section §3.3 will look at the effect different atmospheric conditions have on the line shape and the optical path length. In order to get a rough idea of the line shape and therefore of the atmospheric conditions, we measure the line strength at 3 different frequencies in the wing of the line. These frequencies and the reason for their choice will be described in section §3.4. Finally in §3.5 the advantages and disadvantages of radiometers measuring the 183 GHz line are listed.

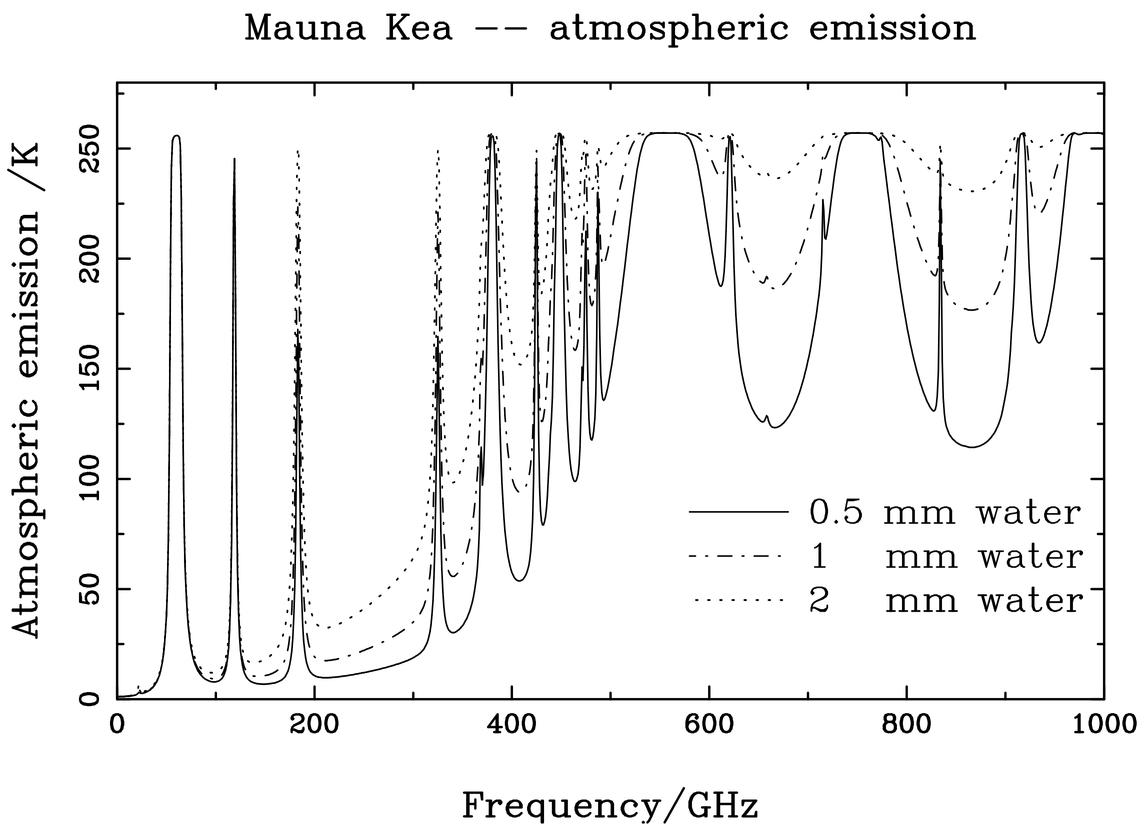


Figure 3.1. Atmospheric emission on Mauna Kea

3.1 Atmospheric emission on Mauna Kea

Fig. 3.1 shows a typical atmospheric emission spectrum for Mauna Kea in Hawaii. The emission is plotted for 0.5 mm, 1 mm and 2 mm of precipitable water vapour (pwv)[1]. The emission lines which change in strength with the amount of water vapour are obviously water emission lines, whereas those which stay constant are emission lines of other atmospheric molecules such as the transitions of oxygen around 60 GHz. The first water line on this plot is at 22 GHz, but it is relatively weak on Mauna Kea, above which there is less than 2.2 mm pwv during half of the nights. Our radiometers measure the water line at 183 GHz, which is much stronger and starts saturating in the line centre at around 2 mm pwv. The next emission lines of water vapour are at 325 GHz, 380 GHz and 425 GHz. Besides these distinct emission lines of water vapour there is also an underlying emission which increases approximately with frequency squared and is proportional to the amount of water vapour. In Fig. 3.1 one can see this square law behaviour between 200 GHz and 380 GHz.

Above 500 GHz the atmosphere emits so strongly that it becomes optically thick and therefore opaque to radiation from astronomical sources. Astronomical observations are only possible in a few atmospheric windows around 700 GHz and 900 GHz.

From Fig.3.1 it is clear that the 183 GHz line is very sensitive to the amount of water vapour, i.e. the brightness temperature varies significantly when the amount of water vapour changes.

3.2 Modelling the 183 GHz line

In order to get a better understanding of the water vapour emission at 183 GHz and its effect on the interferometer phase, two simple computer programs were written to simulate the emission line and to compute the extra optical path length. The models used to calculate the atmospheric parameters, the emission line, and the optical path length are described in this section. The results of the simulations will be discussed in section 3.3.

[1] 1 mm of precipitable water vapour means that if all the water vapour in the atmosphere is condensed on the ground, it would create a 1 mm thick water layer.

3.2.1 Atmospheric model

The strength of the water vapour emission depends on the physical temperature and pressure of the water vapour. A small program has been written to calculate temperature, pressure and density of water vapour as a function of height.

Temperature

According to measurements the temperature typically decreases by 2% for every 1 km of height in the troposphere (Thompson et al., 1994). This temperature profile was used in the simulations.

$$T(h) = T_{ground}\,(0.98)^{h/1km}\,, \tag{3.1}$$

where $T(h)$ is the temperature at height h above the ground,
T_{ground} is the temperature at ground level.

Pressure

The pressure decays exponentially with a scale height of H (Houghton, 1986):

$$p(h) = p_{ground} \times e^{-\int_0^h 1/H(h)dh} \tag{3.2}$$

$$H(h) = \frac{RT}{M_r g} = 7.65\left(\frac{T}{260\text{K}}\right)\ \text{km}. \tag{3.3}$$

where R is the molar gas constant $(8.31J/(mol\ K))$,
M_r the molecular weight of air $(28.8g/mol)$,
g the gravitational constant $(9.80m/s^2)$.

Water vapour

To a first approximation, water vapour decreases exponentially with a scale height of 2 km (Thompson et al., 1994):

$$\rho = \rho_{ground} \times e^{-h/2km}. \tag{3.4}$$

In contrast to most other atmospheric gases, water vapour is not distributed evenly, but occurs in irregular patches, which is of course the reason for the phase fluctuations in interferometry. Warm air can contain a lot of water vapour but when it rises and cools the vapour condenses or freezes out and it often rains or snows. Once precipitation has occurred the air will be much drier, even if it warms up again. The absolute amount of water vapour in an air layer therefore depends on its history and can vary greatly. To take account of this the computer program uses this exponential distribution for the underlying water vapour, but additional water vapour can be added to any height in order to represent moist layers of air.

3.2.2 Line emission and line shape

To describe the line strength the brightness temperature T_{bri} rather than specific intensity (brightness) B_ν is used. The brightness temperature is defined as:

$$T_{bri} \equiv \frac{c^2}{2k\nu^2} B_\nu, \tag{3.5}$$

where ν is the frequency considered.

In the Rayleigh-Jeans regime ($h\nu << kT$) the brightness temperature of a black body is equal to its physical temperature.

To calculate the brightness temperature received by the radiometers, the atmosphere is divided into several layers of height Δs. Each layer is considered uniform with all the gas being at the same physical conditions, but these conditions change from one layer to the next. The program uses the radiative transfer method: radiation is traveling from the higher layers down to the radiometers. In each layer some radiation will be absorbed and the atmosphere in that layer will also emit itself:

$$T_{bri} = T_0 e^{\tau_{-\nu}(\boldsymbol{s})\Delta s} + T_{bb}\left(1 - e^{-\tau_\nu(\boldsymbol{s})\Delta s}\right), \tag{3.6}$$

where T_0 is the brightness temperature of the radiation entering the layer,
$\tau_\nu(s)$ is the optical depth at the frequency ν,
$\boldsymbol{s}$ the position vector of the layer,
T_{bb} the black body temperature of the layer.

The optical depth is the integral of the absorption coefficient over the height of the layer. Using our assumption of constant physical conditions within a layer $\tau_\nu(\boldsymbol{s})$ is just the product of the absorption coefficient and the thickness of the layer Δs. The absorption has a contribution from the line emission $k_{\nu,line}$ as well as from the empirical correction $k_{\nu,cor}$:

$$\begin{aligned} \tau_\nu(\boldsymbol{s}) &= \int_{\boldsymbol{s}-\frac{1}{2}\Delta\boldsymbol{s}}^{\boldsymbol{s}+\frac{1}{2}\Delta\boldsymbol{s}} (k_{\nu,line}(\boldsymbol{s}) + k_{\nu,cor}(\boldsymbol{s}))d\boldsymbol{s} \\ &= (k_{\nu,line}(\boldsymbol{s}) + k_{\nu,cor}(\boldsymbol{s}))\Delta s. \end{aligned} \tag{3.7}$$

As mentioned in §3.1 the correction term describes the underlying radiation which increases approximately with frequency squared (Gaut and Reifenstein, 1971). This term is probably due to the far wings of infrared emission lines of water vapour (Zammit and Ade, 1981).

$$k_{\nu,cor} = 1.08 \times 10^{-11}\rho \left(\frac{300K}{T}\right)^{2.1} \frac{p}{1000mbar}\, \nu^2, \tag{3.8}$$

where ρ is the density of water vapour in grams per meter cubed,
T the temperature of the atmospheric layer in Kelvin,
p its pressure in mbar.

The absorption coefficient for the line emission is:

$$k_{\nu,line} = 1.44\rho\nu T^{-3/2} \left(\sum_{all\ transitions} \left(e^{-E_l/kT} - e^{-E_m/kT}\right) g_l |\phi_{lm}|^2 f\left(\nu, \nu_{lm}\right) \right). \tag{3.9}$$

The sum is taken over all transitions of water vapour. Waters (Waters, 1976) lists the energy levels of the states E_l and E_m as well as the degeneracy g_l and the probability ϕ_{lm} for each transition, see Fig. 3.2. $f(\nu, \nu_{lm})$ describes the line shape. The computer program uses the Ben-Reuven (kinetic) line shape[2]:

[2] The Ben-Reuven line shape agrees better with measurements (Prado-Carrión, 1996) than the Van Vleck and Weisskopf line shape.

$$f(\nu,\nu_{lm}) = \frac{1}{\pi}\frac{4\nu\nu_{lm}\Delta\nu}{\left(\nu_{lm}^2-\nu^2\right)^2+4\nu^2\Delta\nu^2}. \tag{3.10}$$

The centre frequency of the transition ν_{lm} is given in Walters' table (3.2). The line width $\Delta\nu$ is basically proportional to the pressure and inversely proportional to the temperature to the power of x, where x takes values between 0.3 and 0.7 depending on the transition. More precisely it is given by:

$$\Delta\nu = \Delta\nu_{lm}^0\left(\frac{p}{1013\ mbar}\right)\left(\frac{T}{300\ K}\right)^{-x}[1+4.6*10^{-3}\frac{\rho T}{p}\left(\frac{\Delta\nu_{lm}(H_2O)}{\Delta\nu_{lm}^0}-1\right)]. \tag{3.11}$$

The line width parameters $\Delta\nu_{lm}^0$, x and $\Delta\nu_{lm}(H_2O)$ are taken from Waters (1976) and are listed in Fig. 3.2, all the other parameters are as above. The term in the square bracket usually takes values close to 1.

3.2.3 Optical path length

Atmospheric water vapour not only emits, but also increases the refractive index of air and therefore decreases the speed at which electromagnetic waves travel. Radiation needs the extra time Δt to traverse moist air:

$$\Delta t = \frac{1}{c}\int(n_{moist}(\nu,\rho(T,p)) - n_{dry}(\nu))d\boldsymbol{s}, \tag{3.12}$$

where $n_{moist}(\nu,\rho(T,p))$ is the refractive indices of moist air as a function of frequency ν. It also depends on the water vapour density ρ and its temperature T and pressure p.

n_{dry} is the refractive indices of dry air.

The excess path length L is defined as:

$$L = c\Delta t. \tag{3.13}$$

The interferometric phase shift $\Delta\phi$ is due to a difference in optical path lengths ΔL along the line of sight of the two antennas.

TABLE I. Lower-Frequency Spectral Lines of H_2O^{16}

Frequencies[a]	Transitions[b]								Energy levels[c]		Linewidth parameters[d]		
$\nu_{\ell m}$	Upper state			Lower state					E_m	E_ℓ	$\Delta\nu^0_{\ell m}$	$\Delta\nu_{\ell m}(H_2O)$	x
(GHz)	J	K_{-1}	K_1	J	K_{-1}	K_1	g_ℓ	$\lvert\phi_{\ell m}\rvert^2$	(cm^{-1})	(cm^{-1})	(GHz)	(GHz)	
22.23515[e]	6	1	6	5	2	3	3	.0549	447.30	446.56	2.85[f]	13.68[f]	.626
183.31012[g]	3	1	3	2	2	0	1	.1015	142.27	136.16	2.68[g]	14.49[g]	.649
(323.)	10	2	9	9	3	6	3	.0870	1293.80	1283.02	2.30	12.04	.420
325.1538[h]	5	1	5	4	2	2	1	.0891	326.62	315.78	3.03[i]	15.21	.619
380.1968[h]	4	1	4	3	2	1	3	.1224	224.84	212.16	3.19[i]	15.84	.630
(390.)	10	3	7	11	2	10	1	(.0680)	1538.31	1525.31	2.11	11.42	.330
(436.)	7	5	3	6	6	0	1	(.0820)	1059.63	1045.03	1.50	7.94	.290
(438.)	6	4	3	5	5	0	3	.0987	756.76	742.11	1.94	(10.44)	(.360)
(442.)	7	5	2	6	6	1	3	(.0820)	1059.90	1045.11	1.51	8.13	.332
448.0008[h]	4	2	3	3	3	0	3	.1316	300.37	285.42	2.47	14.24	.510

[a] Frequency values in parentheses are calculated by S. N. Ghosh and H. D. Edwards [U.S. Air Force Survey in Geophysics, No. 82, AFCRL, Bedford, Massachusetts (1956)], and should be accurate to ~4 GHz, as determined by comparison of their calculated values with the measured frequencies.

[b] Matrix elements $\lvert\phi_{\ell m}\rvert^2$ are from G. W. King *et al.* [*Phys. Rev.* **71**, 433 (1947)], except for values in parentheses which were estimated by N. E. Gaut [M.I.T. Res. Lab. Electron. Tech. Rep. 467 (1968)] from computations given in *Microwave Spectral Tables* [U.S. Dept. Commerce, NBS Monograph 70].

[c] From R. T. Hall and J. M. Dowling [*J. Chem. Phys.* **47**, 2454 (1967)], and D. M. Dennison [*Rev. Mod. Phys.* **12**, 175 (1940)]. The energy levels in Table I are in wavenumber units of reciprocal centimeters and should be multiplied by Planck's constant for use in (2.3.20).

[d] Except where noted linewidths $\Delta\nu_{\ell m}$ and temperature exponents x are from W. S. Benedict and L. D. Kaplan [*J. Chem. Phys.* **30**, 388 (1959); *J. Quant. Spectrosc. Radiat. Transfer* **4**, 453 (1964)], with values of $\Delta\nu^0_{\ell m}$ computed for nitrogen broadening only. Values in parentheses were estimated by N. E. Gaut [M.I.T. Res. Lab. of Electron. Tech. Rep. 467 (1968)] based on averages of values for similarly arising lines.

[e] H. J. Liebe *et al.*, *J. Quant. Spectrosc. Radiat. Transfer* **9**, 31 (1969). Hyperfine components of this line have been measured by H. Bluyssen *et al.*, *Phys. Lett.* **24A**, 482 (1967).

[f] H. J. Liebe and T. A. Dillon, *J. Chem. Phys.* **50**, 727 (1969).

[g] J. R. Rusk, *J. Chem. Phys.* **42**, 493 (1965).

[h] M. Lichtenstein, V. E. Derr, and J. J. Gallagher, *J. Mol. Spectrosc.* **20**, 391 (1966).

[i] V. Ya. Ryadov., N. I. Furashov, *Izv Vyssh. Uchebn. Zaved. Radiofiz.* [*Sov. Radiophys.*] **9**, 1073 (1966); *Opt. Spektrosk.* **24**, 186 (1968).

Figure 3.2. Line parameters describing transitions of water vapour. The table is taken from Waters (1976).

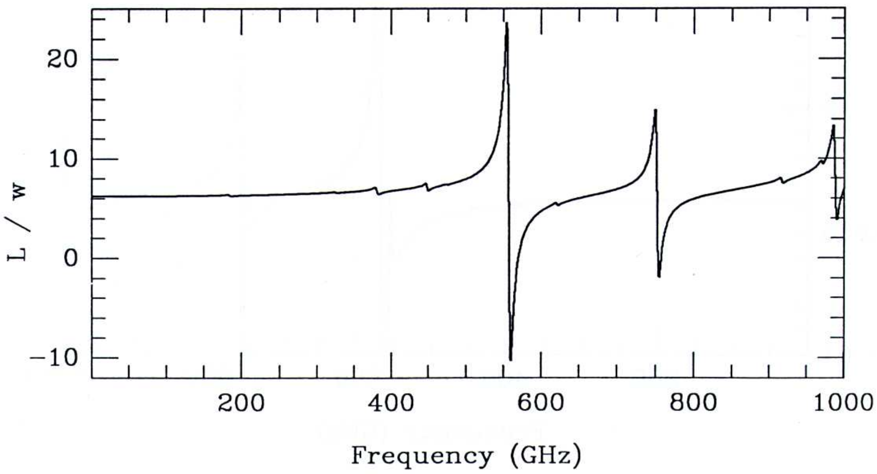


Figure 3.3. Excess optical path divided by precipitable water vapour is plotted versus the observing frequency. Figure taken from Sutton and Hueckstaedt, 1996.

$$\Delta\phi = 360^\circ \times \frac{\Delta L}{\lambda}, \tag{3.14}$$

where λ is the wavelength observed.

To first approximation the excess optical path L is simply proportional to the amount of precipitable water vapour (pwv) (Thompson et al., 1994):

$$L = 6.3pwv, \tag{3.15}$$

where pwv is the precipitable water vapour in the same units as L.

For example 1 mm of pwv, a typical value for Mauna Kea, would give an excess optical path of 6.3 mm.

Equ. 3.15 suggests that the optical path length and therefore the refractive index of moist air increase linearly with the amount of water vapour, but are independent of frequency. According to Equ. 3.15 the proportionality factor L/pwv should take the constant value of 6.3 at all frequencies. However, Fig. 3.3 shows that this is not the case. At the frequencies of the water vapour transitions, in particular around 560 and 760 GHz, the proportionality factor varies significantly from 6.3. Fortunately however, we usually do not need to know the excess path lengths at the frequencies of the water vapour transitions, because at these frequencies the opacity of the atmosphere is so high that ground based observations are nearly impossible. Below 400 GHz the proportionality factor has the constant value of 6.3 to an accuracy of better than 5%. At higher frequencies, in particular between 620 and 720 GHz, the proportionality factor increases with frequency. Since all of our interferometric observations discussed in this thesis were conducted around 350 GHz, we assumed the proportionality constant to be independent of frequency.

According to Thompson et al. (1994) the excess path length L due to water vapour depends on temperature and density as

$$L = 1763 * 10^{-6} \int_0^\infty \frac{\rho(\boldsymbol{s})}{T(\boldsymbol{s})} d\boldsymbol{s}. \tag{3.16}$$

where $\rho(\boldsymbol{s})$ is the water vapour density at position $\boldsymbol{s}$ in g/m^3.

The excess path has the same units as the distance from the telescope s. This formula is used in the computer programs and the following discussion. (It is consistent with Equ. 3.15, because by approximating the atmosphere to be isothermal at 280 K one can analytically solve the integral and Equ. 3.16 becomes 3.15.)

Typically, the amount of water vapour changes by about 0.05 mm over 10 sec, which gives a difference in optical path of 0.31 mm and a phase difference of 137° at 360 GHz (wavelength of 0.83 mm).

3.2.4 Example plot

Fig. 3.4 shows a typical output of the computer program. Brightness temperature (which is proportional to intensity) is plotted against frequency. The emission of water vapour is plotted for 0.5, 1, 2 and 4 mm of pwv distributed exponentially (see Equ. 3.4). The atmospheric parameters chosen are typical for Mauna Kea. At a height of 4200 m, the temperature often lies around 270 K, the pressure is chosen to be 600 mbar. The atmosphere has been divided into 20 layers each 200 m high (see Tab. 3.1).

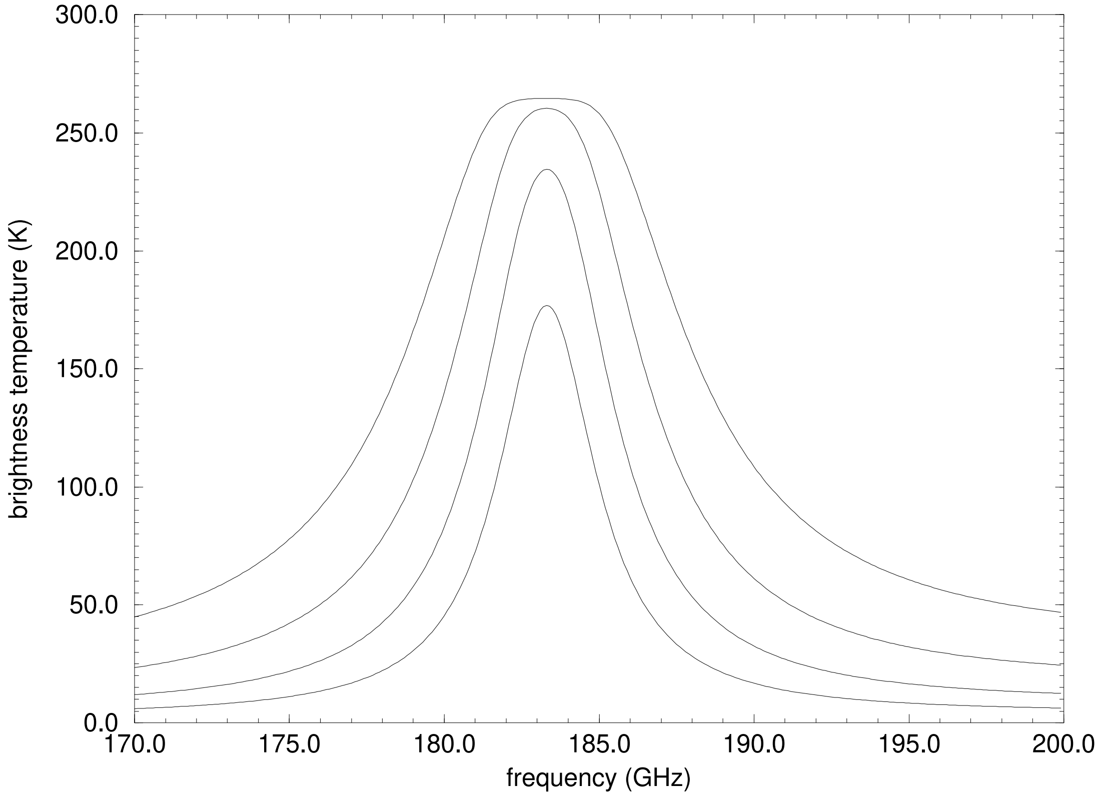


Figure 3.4. Brightness temperature of 0.5, 1, 2 and 4 mm precipitable water vapour.

atmospheric parameter	symbol	value	
pressure	p_{ground}	600	mbar
temperature	T_{ground}	270	K
water vapour	pwv	1	mm
height	–	4200	m
thickness of atm. layer	–	200	m
number of atm. layers	–	20	

Table 3.1. In the computer programs these values for the atmospheric parameters were used unless explicitly stated otherwise.

3.3 Effect of atmospheric parameters on line shape, optical path length and sensitivity

From Fig. 3.4 it is obvious that the emission increases with the amount of water vapour as does the excess path length (Equ. 3.16). Additionally, the physical conditions of the emitting water vapour also influence the line shape and the excess path. This section will look at the effect a change in pressure, temperature and height of the emitting water vapour has on the line shape, on the excess path and on the sensitivity $\Delta T/\Delta L$.

3.3.1 Changes in pressure

1. Line emission
   Fig. 3.5 shows the water vapour line for $p_{ground} = 400$ mbar, 600 mbar and 800 mbar. As before p decays exponentially with height (Equ. 3.2) and all other parameters are as in Tab. 3.1. These are, of course, very extreme changes in pressure, but they demonstrate the effect on the line shape well. Under 800 mbar the line is substantially pressure broadened. At low pressure (400 mbar) the line is narrower but stronger in the line centre. There is, however, a cross over point (pivot point) at 184.5 GHz where the emission is nearly independent of pressure. The first frequency channel of the water vapour monitors measures the intensity at this frequency. In contrast to the 22 GHz line the position of the pivot point shifts slightly away from the line centre when the underlying amount of water vapour increases.

2. Optical path
   The excess optical path is independent of the pressure the water vapour is at (see Equ. 3.16).

3.3.2 Changes in temperature

1. Line emission
   Fig. 3.6 shows the water line for three different ground temperatures, 250, 270 and 290 K. The main difference between the curves is the brightness temperature at the line centre. Line emission of a gas at temperature T cannot exceed the emission of a black body at T. Therefore water vapour at 290 K can radiate more strongly than vapour

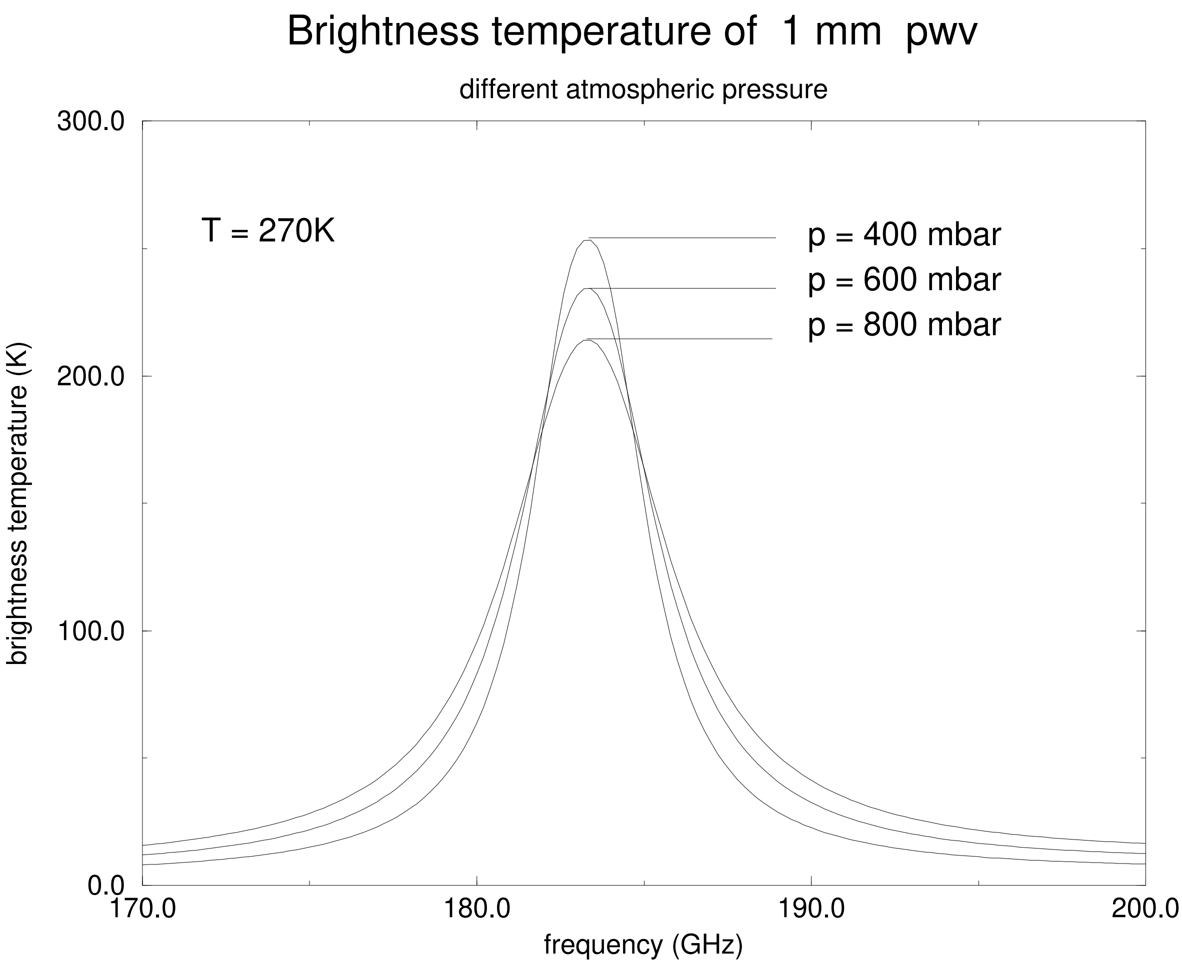


Figure 3.5. Emission of 1 mm of water vapour for p_{ground} = 400, 600 and 800 mbar, T_{ground} = 270 K.

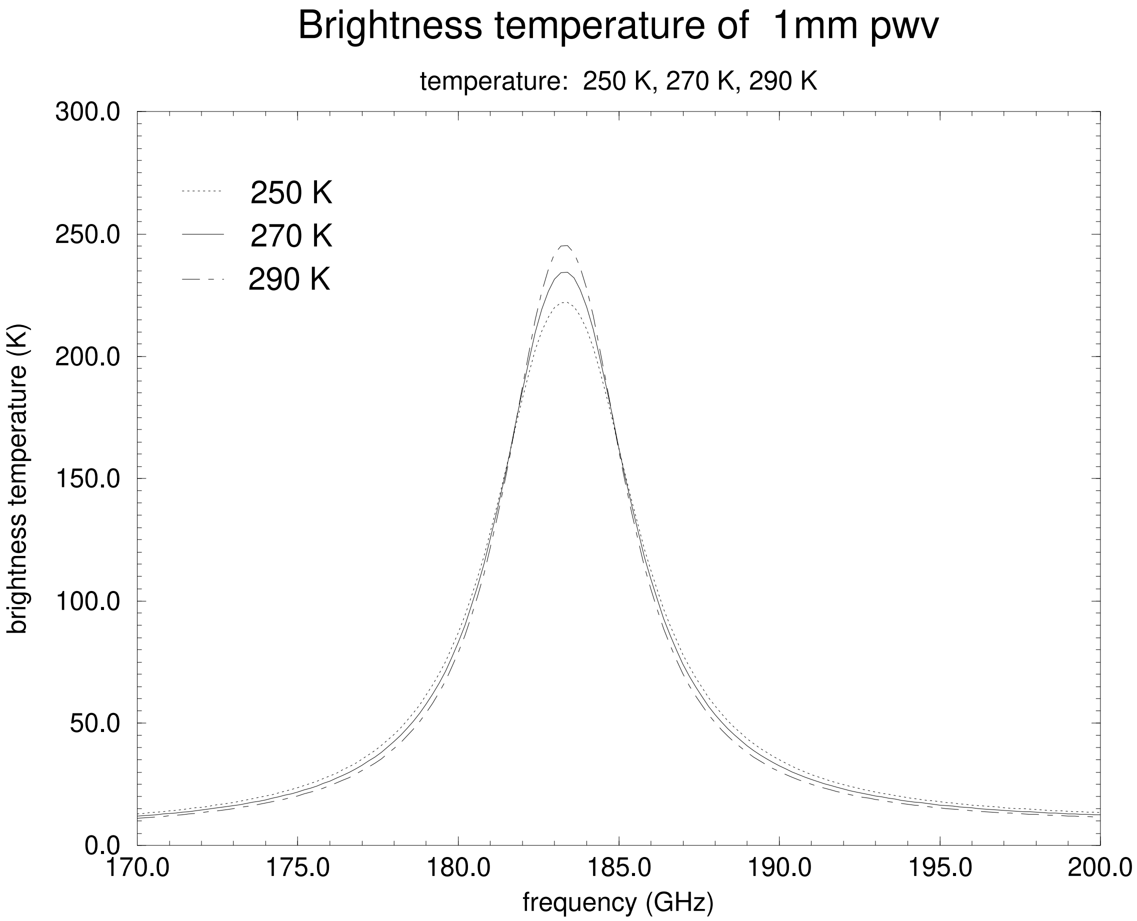


Figure 3.6. Emission of 1 mm of water vapour for T_{ground} = 250 K, 270 K and 290 K, $p_{ground} = 600$ mbar.

at 250 K. When the line is saturated (optically thick) the brightness temperature will correspond to the physical temperature of the water vapour. As in the case of changing pressures there is a pivot point where the brightness temperatures of the three lines are the same. However, this pivot point strongly depends upon the amount of water vapour. As the amount of water vapour is increased, the crossover point shifts away from the line centre.

2. Optical path

   The excess path length depends on temperature as well (Equ. 3.16): it decreases as 1/T. In this example the excess path lengths are 7.61 mm, 7.04 mm and 6.56 mm for ground temperatures of 290 K, 270 K and 250 K.

3.3.3 Water vapour at different heights

The two previous sections illustrate the effect a change in pressure and temperature has on the water line and the excess path. In principle the ground temperature and pressure can easily be determined and subsequently the temperature and pressure at different heights in the atmosphere can be calculated. However, the distribution of water vapour is unknown. In the above examples water vapour was assumed to decrease exponentially with height. This is only an approximation, and it could well be that there are "clouds" of water vapour at certain heights in the atmosphere.

1. Line emission

   Fig. 3.7 shows the emission in the extreme case where all the water vapour is located at 4200 m (just above the telescope), at 6200 m and at 8200 m. Since temperature and pressure decrease with height, the physical conditions of the emitting water are different in the three cases and so are the line shapes. The pressure has the dominant influence on the line shape: the emission from water vapour at 8200 m is concentrated towards the centre due to the low pressure at that altitude, but is not much weaker as would be expected due to its lower temperature.

2. Optical path

   The optical path length increases with the height of the water vapour because the excess optical path length is inversely proportional to the vapour's temperature (Equ. 3.16).

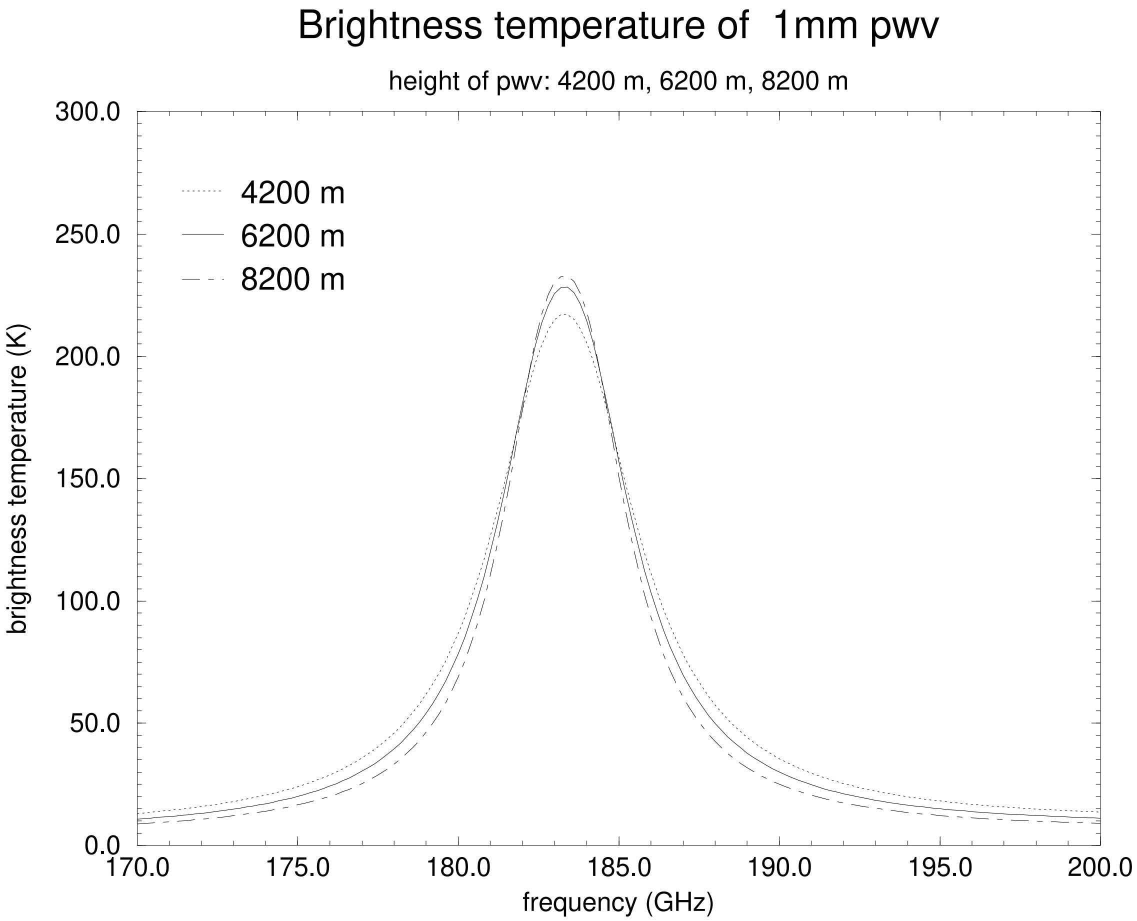


Figure 3.7. Emission of 1 mm of precipitable water vapour concentrated at a height of 4200 m, 6200 mm and 8200 m, $p_{ground} = 600$ mbar, $T_{ground} = 270$ K.

For example 1 mm pwv at 4200 m introduces an excess optical path of 6.53 mm, 1 mm pwv at 6200 m introduces an excess path of 6.77 mm and and for 1 mm pwv at 8200 m the excess optical path becomes 7.05 mm.

3.3.4 Sensitivity

In order to successfully correct the interferometric phase we do not necessarily need to know the absolute amount of water vapour. We do, however, need to know the difference in brightness temperature ΔT_{bri} measured by the radiometers and how it translates to an excess optical path ΔL observed by the interferometer. Therefore the sensitivity $\Delta T/\Delta L$ was calculated by taking the difference in brightness temperature for 1.1 mm pwv and 1.0 mm pwv and dividing it by the difference in excess path for these two values of water vapour. To convert the brightness temperature into phase one needs to divide by this sensitivity.

Sensitivity versus frequency

In Fig. 3.8 the sensitivity $\Delta T/\Delta L$ is plotted against frequency for the three different ground pressures, p_{ground} = 400, 600, 800 mbar. The conversion factor has a double peak structure. The low conversion factor in the line centre at 183.3 GHz is due to the line starting to become saturated. Extra water vapour adds to the optical path length by a fixed amount, but increases the brightness temperature only a little at the line centre. This effect is strongest for p_{ground} = 400 mbar, because the line emission is concentrated in the centre and the line is closer to saturation than more pressure broadened lines. The sensitivity is highest approximately 2 GHz away from the line centre. However, this maximum will shift away from the line centre as the amount of water vapour increases, in the same way as the region of saturation spreads with an increasing amount of water vapour. The sensitivity drops further away from the line centre because there is little emission outside the line. The sensitivity curves cross over at 184.5 GHz because the sensitivity is independent of the underlying pressure at the pivot point.

Fig. 3.9 and 3.10 contrast the sensitivities of the water line for different ground temperatures and for water vapour concentrated at different heights, respectively.

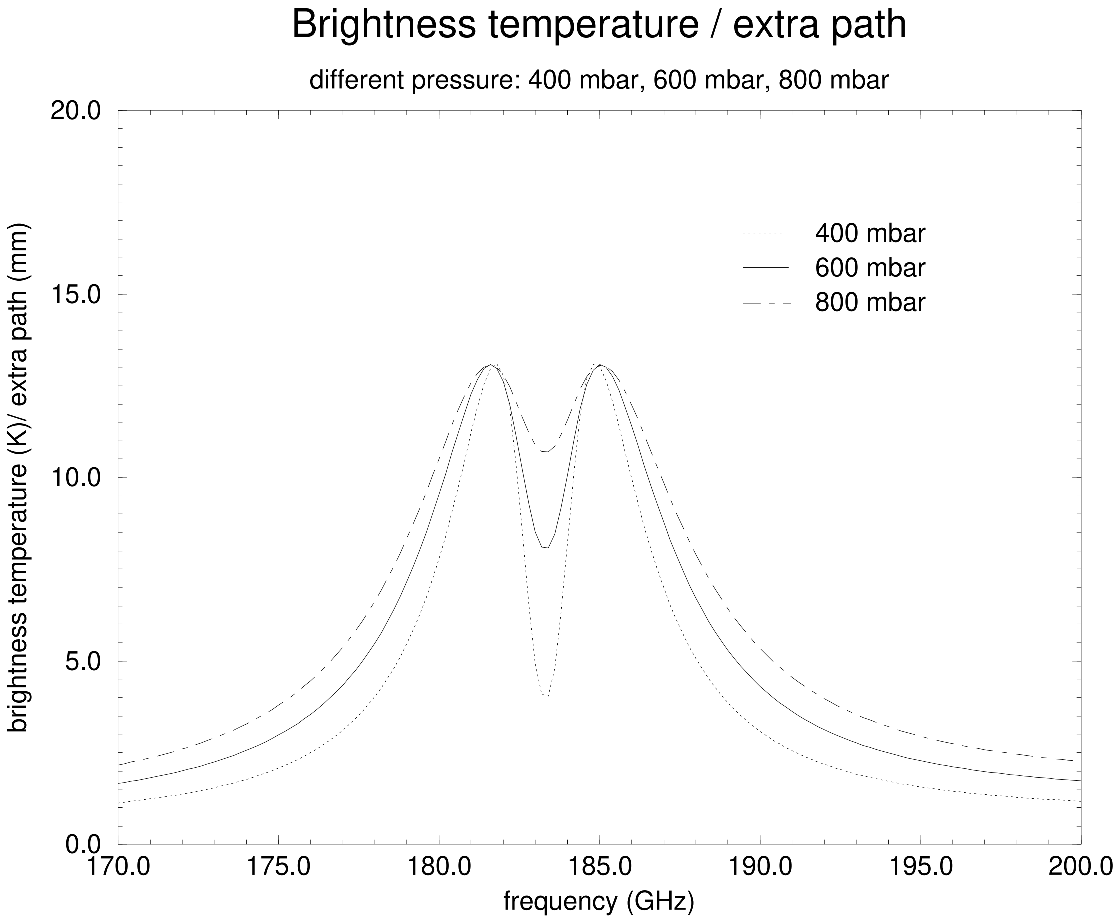


Figure 3.8. Sensitivity $\Delta T/\Delta L$ for 1 mm of precipitable water vapour, $p_{ground} = 400$ mbar, 600 mbar and 800 mbar $T_{ground} = 270$ K.

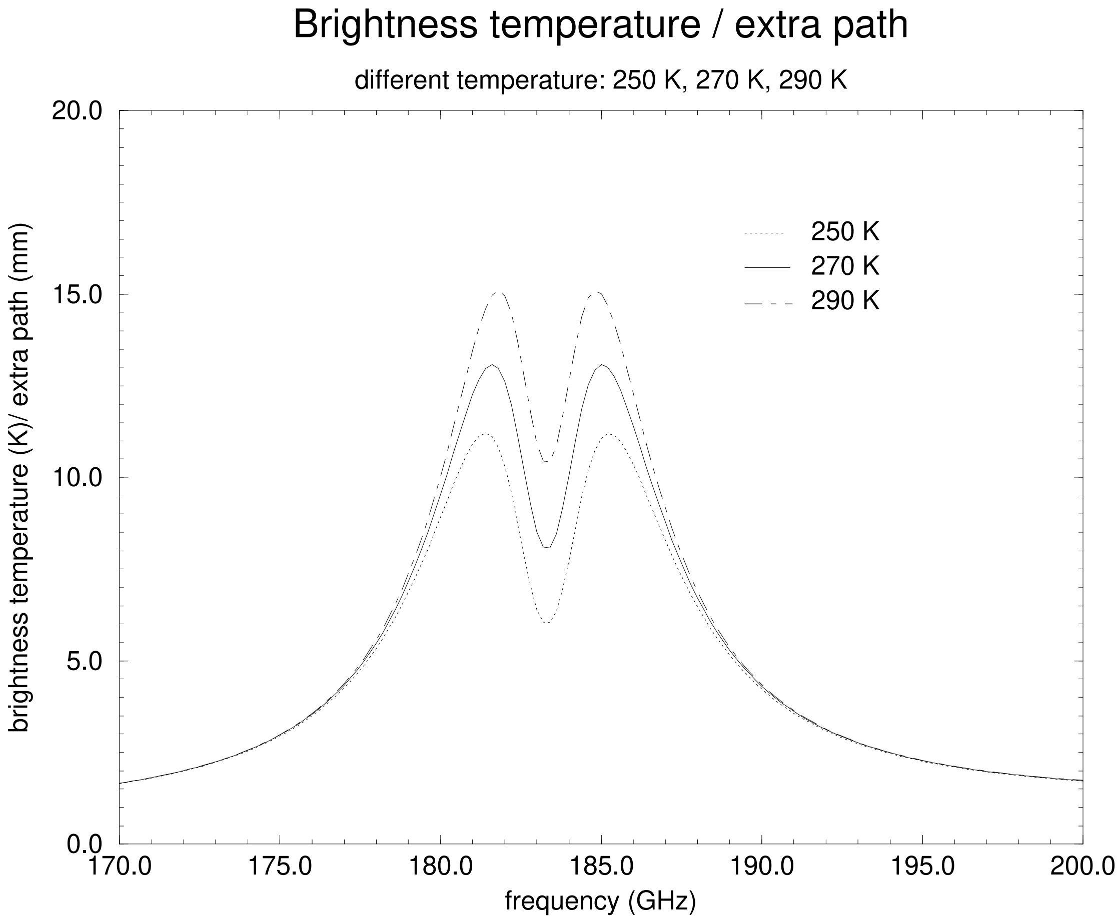


Figure 3.9. Sensitivity $\Delta T/\Delta L$ for 1 mm of precipitable water vapour, $T_{ground} = 250$ K, 270 K and 290 K, $p_{ground} = 600$ mbar.

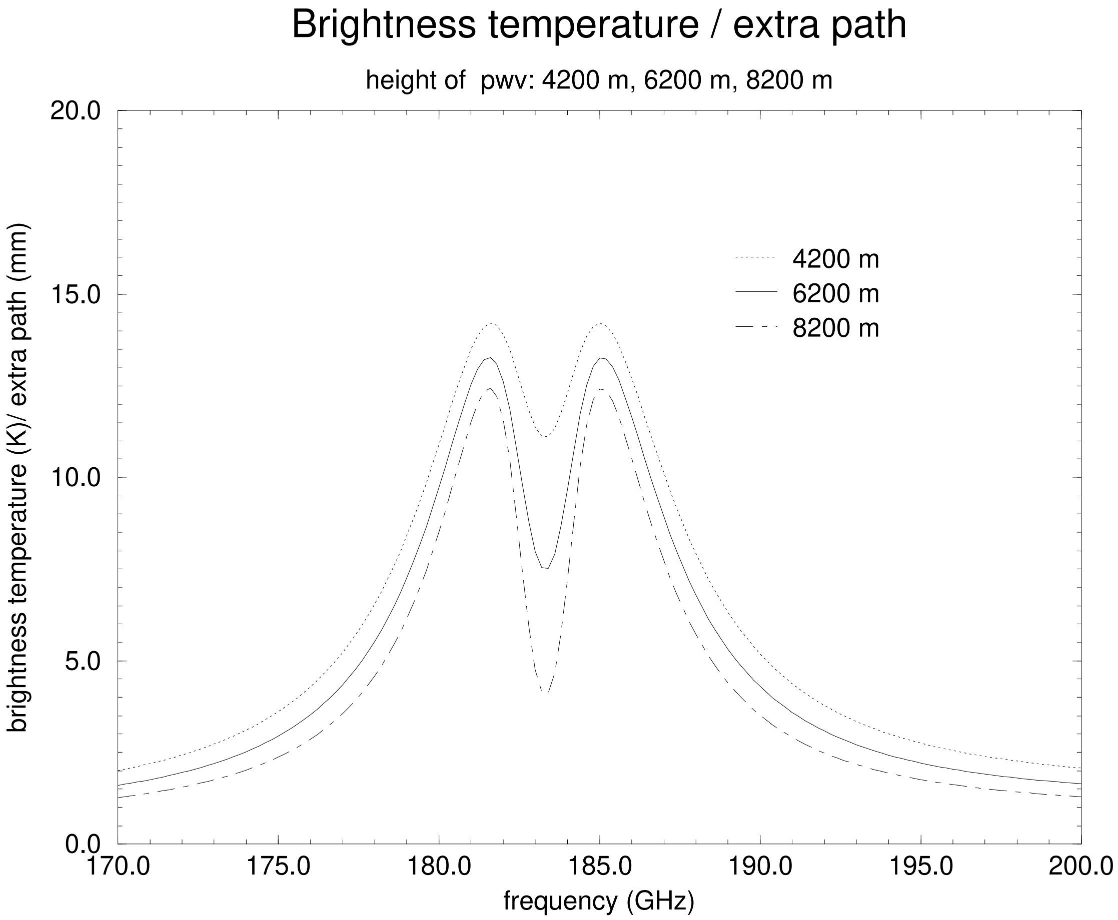


Figure 3.10. Sensitivity $\Delta T/\Delta L$ for 1 mm of precipitable water vapour concentrated at 4200 m, 6200 m and 8200 m, $T_{ground} = 250$ K, $p_{ground} = 600$ mbar.

Sensitivity versus brightness temperature

In practise we have to work with a varying and unknown amount of pwv. We can however estimate the amount of pwv from the observed brightness temperature. In Fig. 3.11 the sensitivity $\Delta T/\Delta L$ times the wavelength λ [3] is plotted against brightness temperature for a wide range of pwv. The sensitivity has been calculated for the three different channels of the radiometers by integrating the brightness temperature over the frequency range of each channel (see section 3.4) and dividing it by its bandwidth.

The sensitivity decreases approximately linearly with increasing brightness temperature. This relationship can be derived by assuming that all the water vapour has the same temperature and pressure. From Equ. 3.7, 3.8 and 3.9 we get:

$$\begin{aligned} \tau_\nu(\boldsymbol{s}) &= (k_{\nu,line} + k_{\nu,cor})\Delta s \\ &= (A_{line}(p,T,\nu)\rho + A_{cor}(p,T,\nu)\rho)\Delta s, \end{aligned} \tag{3.17}$$

where $A_{line}(p,T,\nu)$ and $A_{cor}(p,T,\nu)$ are defined as:

$$\begin{aligned} A_{line}(p,T,\nu) &\equiv 1.44\nu T^{-3/2} \times \\ &\quad \sum_{all\ tran.} \left(e^{-E_l/kT} - e^{-E_m/kT}\right) g_l|\phi_{lm}|^2 f\left(\nu,\nu_{lm}\right), \end{aligned} \tag{3.18}$$

$$A_{cor}(p,T,\nu) \equiv 1.08 \times 10^{-11} \left(\frac{300}{T}\right)^{2.1} \frac{p}{1000}\,\nu^2. \tag{3.19}$$

Remember that the square bracket in Equ. 3.11 is approximately 1 and therefore A_{line} is independent of ρ.

Equ. 3.15 shows that the amount of water vapour is proportional to the optical path.

Simplifying the radiative transfer equation 3.6 by representing the atmosphere as a single emitting layer (i.e. $T_0 = 0$) we get:

$$\begin{aligned} T_{bri} &= T_{bb}\left(1 - e^{-\tau_\nu(\boldsymbol{s})\Delta s}\right) \\ &= T_{bb}\left(1 - e^{-(A_{line}(p,T,\nu)+A_{cor}(p,T,\nu))\frac{L}{6.3}}\right). \end{aligned} \tag{3.20}$$

[3] Multiplication with λ changes the units of the sensitivity from K per mm to K per 360° (or one turn) of phase.

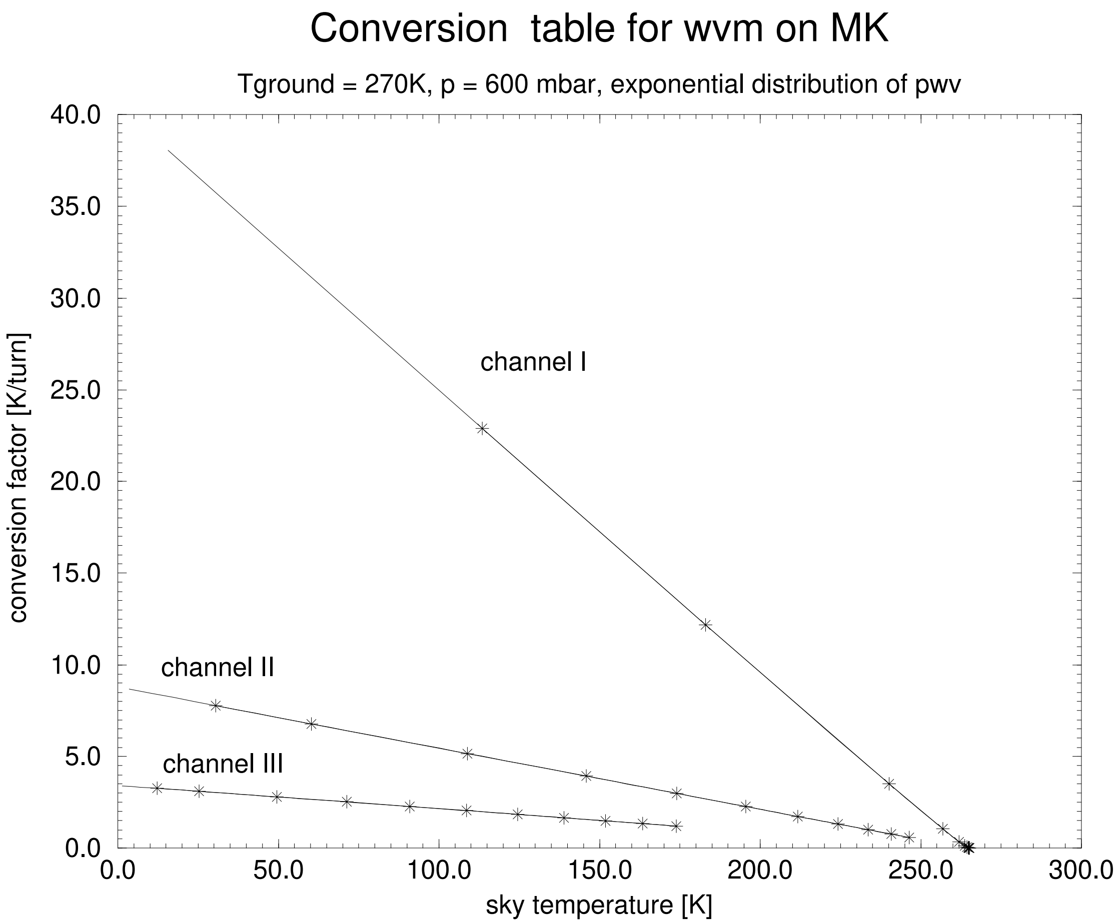


Figure 3.11. Sensitivity $\Delta T/(360^\circ$ of phase) as a function of brightness temperature for $T_{ground} = 270$ K, $p_{ground} = 600$ mbar and observing frequency 350 GHz. The 11 asterisks on each curve indicate brightness temperature and sensitivity for 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 mm pwv from left to right.

The derivative of the brightness temperature T_{bri} with respect to optical path L is

$$\begin{aligned}\frac{d\,T_{bri}}{d\,L} &= T_{bb}e^{-(A_{line}(p,T,\nu)+A_{cor}(p,T,\nu))\frac{L}{6.3}}(A_{line}(p,T,\nu)+A_{cor}(p,T,\nu))\frac{1}{6.3} \\ &= (T_{bb}-T_{bri})(A_{line}(p,T,\nu)+A_{cor}(p,T,\nu))\frac{1}{6.3}. \end{aligned} \tag{3.21}$$

Therefore, in the approximation of constant atmospheric conditions for the emitting water vapour, the sensitivity is proportional to the difference between the physical temperature and the observed brightness temperature.

The slope is steepest for the channel closest to the line centre because its sensitivity is most effected by saturation. On the other hand, this channel is the most sensitive for small amounts of underlying pwv.

Error in determining sensitivity

In Fig. 3.12 precipitable water vapour (pwv) (rising lines) and sensitivity $\Delta T/\Delta L$ times wavelength λ (decreasing, straight lines) are plotted against brightness temperature for the three channels of the WVMs. Since fluctuations in the water vapour might occur at different heights above the telescope or might be due to a change in exponentially distributed pwv, the curves are plotted for 4 different scenarios:

- additional layer of 0.1 mm pwv 200 m above the telescope (long dash),
- additional layer of 0.1 mm pwv 600 m above the telescope (long dash - short dash),
- additional layer of 0.1 mm pwv 1000 m above the telescope (long dash - dot),
- additional layer of 0.1 mm pwv exponentially distributed over height (crosses).

The curves of pwv versus brightness temperature change slope when saturation sets in. This is most evident in channel I: the pwv increases linearly below $\sim$ 150 K, but for pwv $\geq$ 3 mm the temperature stays constant at $\sim$270 K.

The sensitivity decreases linearly with brightness temperature as discussed above. Additional water vapour at high altitude will yield a smaller $\Delta T_{bri}/\Delta$path ratio (sensitivity) than additional water vapour at low altitudes, because the optical path length is inversely proportional to temperature (Equ. 3.16). This behaviour is clearly visible in channels II and III.

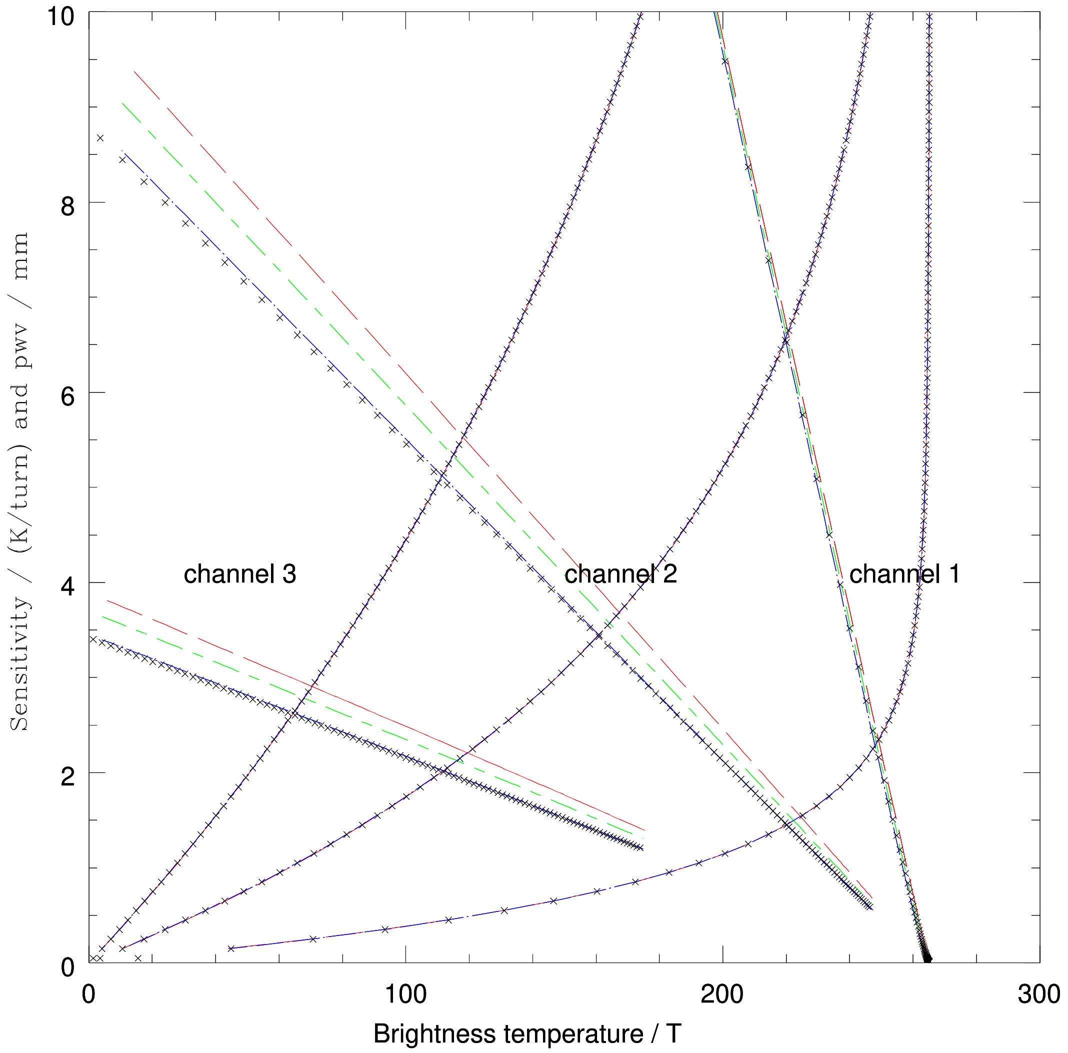


Figure 3.12. Sensitivity (for observations at 350 GHz) and amount of pwv are plotted against brightness temperature for the three channels of the WVMs. The sensitivity is decreasing linearly with increasing brightness temperature. The sensitivity is plotted for an additional layer of 0.1 mm pwv 200 m (long dashes, red), 600 m (long dash - short dash, green) and 1000 m (long dash - dot, blue/purple) above the telescope, as well as for an exponential distribution of pwv (crosses, black). The three rising curves represent the amount of pwv as a function of brightness temperature. Theses curves give approximately the same results for all 4 cases.

In channel I pwv at high altitude will also yield a greater change in brightness temperature than pwv at low altitude, because emission under low pressure is more concentrated on the line centre (Fig. 3.7). The ratio $\Delta T_{bri}/\Delta$path is approximately independent of the height of pwv in channel I.

Fig. 3.12 can be used to determine the brightness temperature and the sensitivity for any amount of water vapour. For 1.5 mm pwv, for example, the expected brightness temperature in channel I is 220 K and the sensitivity is 6.6 K/turn.

Using Fig. 3.12 we determined the uncertainty in the sensitivity and therefore in the conversion factor needed to predict the phase correction. We assumed that:

- there is 1 mm / 4 mm of underlying pwv,
- the additional water vapour might be anywhere between 200 and 1000 m above the telescope or exponentially distributed,
- the measured brightness temperature is determined by comparison with the warm load with an error of 1%, such that the error in brightness temperature is 0.01*(300 K-T_{bri}).

The maximum and the minimum sensitivities as well as the percentage error for each channel are listed in Tab. 3.2.

These error estimates assume that the temperature and the pressure at the height of the telescope are accurately know, and that there are no clouds. Since it should be possible to accurately measure ground temperature and pressure, they will increase the error insignificantly. To better assess the error due to clouds one probably needs to compare theoretical phase predictions with actual measured phases. The errors listed in the table are small, in particular as it is sufficient for phase correction to use the channel or channels with the least uncertainty.

3.3.5 Summary

The examples above show that the line emission as well as the sensitivity not only depend on frequency, but also on the amount of water vapour and its physical conditions. The sensitivity is highest close to the line centre, but there the line is also most sensitive to saturation as well as to changes in temperature and pressure of the emitting water vapour.

1 mm pwv	channel I	channel II	channel III
max. sensitivity [K/mm]	12.2	7.7	3.6
min. sensitivity [K/mm]	11.8	6.6	3.1
percentage error	±2%	±8%	±7%
4 mm pwv	**channel I**	**channel II**	**channel III**
max. sensitivity [K/mm]	0.4	3.4	2.7
min. sensitivity [K/mm]	0.2	2.9	2.2
percentage error	±33%	±8%	±10%

Table 3.2. Error in sensitivities for the three channels of the water vapour monitors for 1 mm and 4 mm of pwv. An uncertainty the extrapolation of the brightness temperature from the warm load temperature was assumed to have an error of 1%. The water vapour is located anywhere between a height of 200 and 1000 m.

3.4 Frequency channels of the water vapour monitors

3.4.1 Number of channels

The water vapour monitors measure the brightness temperature in three different frequency channels in the wing of the line, Fig. 3.13. The more channels the radiometers have the better one can determine the spectrum of the water line, but at the same time the cost goes up and the signal to noise ratio goes down due to necessarily smaller bandwidths in each channel. We decided on three channels as they can cover a wide frequency range but keep the system still relatively simple. The wide frequency range covered allows high sensitivity in at least one channel for all likely amounts of water vapour. In addition, measurements in three independent channels contain some information about the physical conditions of the water vapour, as well as information about other factors which contribute to the measured brightness temperature such as spill-over [4] and clouds. Sources with a flat spectrum such as spill-over contribute equally to each channel. By fitting a theoretical emission spectrum to the three measurements a constant offset can be determined and rejected.

[4]Spill-over means that the water vapour monitors look past the secondary mirror and detect thermal emission radiated by the ground or the telescope dome.

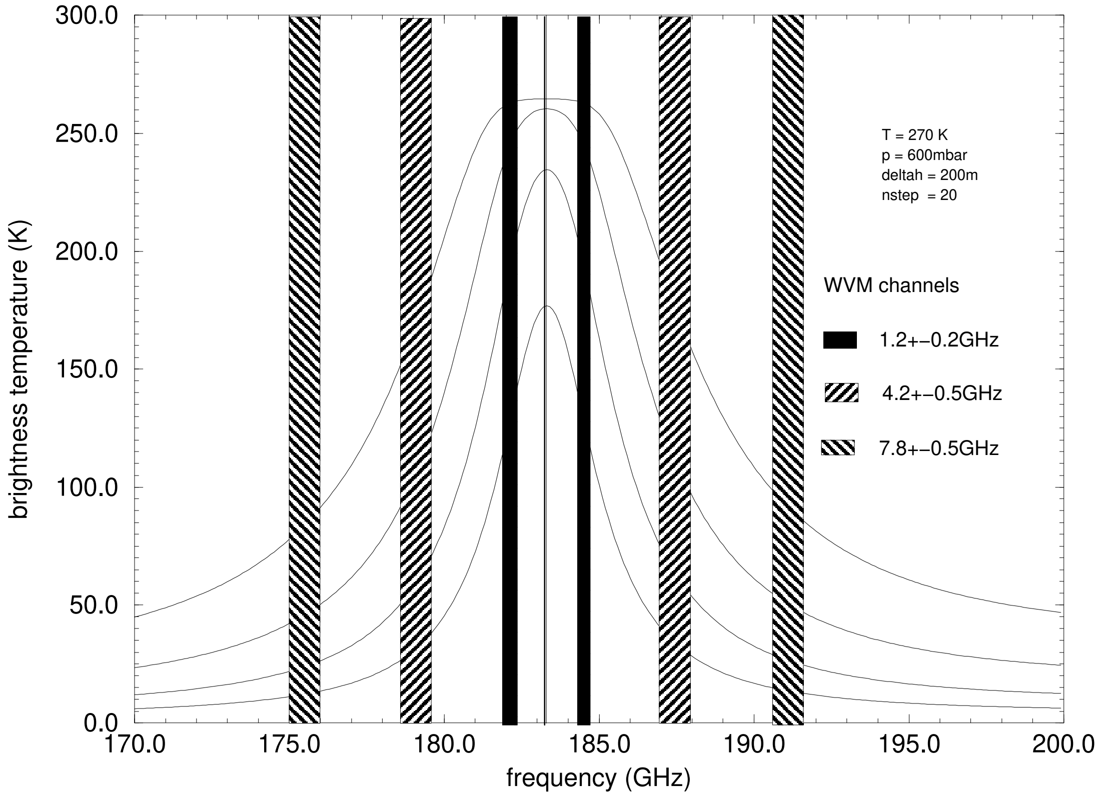


Figure 3.13. The three double-sideband frequency channels of the water vapour monitor superimposed on the spectrum of the water line for 0.5, 1, 2 and 4 mm of pwv.

3.4.2 Frequencies

Fig. 3.13 shows the frequency channels the water vapour monitor is sensitive to superimposed on the emission spectrum of 0.5, 1, 2 and 4 mm of pwv. The first band is 1.2 GHz away from the line centre. It lies on the pivot point at 184.5 GHz, i.e. the frequency which is least sensitive to pressure changes (see §3.3.1). The second frequency band is 4.2 GHz, the third 7.8 GHz away from the line centre. The frequency channels are spread over ~8 GHz, the maximum frequency range at which the mixer and first amplifier operate well. The larger the difference in channel frequencies the easier is it to constrain the physical conditions of the water vapour. The specific frequencies were chosen so that the harmonics of the oscillators of the second down conversion fall outside the frequency bands (see §5.2.2). For example the oscillator of channel I operates at 1.2 GHz, the third (3×1.2 GHz = 3.6 GHz) and fourth harmonic (4×1.2 GHz = 4.8 GHz) both fall outside channel II (3.7 GHz to 4.7 GHz).

Unfortunately, there are ozone lines around 184.4 GHz (Fig. 3.14), which we only learned about after building the radiometers. Fortunately however, the ozone layer is at high altitude and likely to be relatively smooth on the scales of interest to us. Therefore it will add a constant offset to channel I in both the JCMT and the CSO water vapour monitors. As we are interested in the difference in optical path between the JCMT and the CSO we take the difference between the brightness temperatures measured at the two telescopes, and the ozone contribution will cancel itself out. The spectrum of the brightness temperatures is used to determine the physical parameters of the water vapour. But since the ozone lines are very weak (~1 GHz K) in comparison to the water line (for 0.5 mm of pwv, ~120 GHz K) it will barely influence the result. In case a very accurate determination of the physical conditions of the pwv is needed, the ozone lines can easily be taken into account.

Bandwidth

The first band is 400 MHz wide, whereas the second and third frequency bands are 1 GHz wide. The band widths were chosen as wide as possible to increase the signal to noise ratio, but narrow enough for amplifiers to have a flat gain spectrum over each channel and for the theoretical water line to be well approximated by its tangent. [5]

[5] As long as the water line can be approximated by a tangent the band pass of the filter which determines the width of the channel is not of importance, see section 5.2.2.

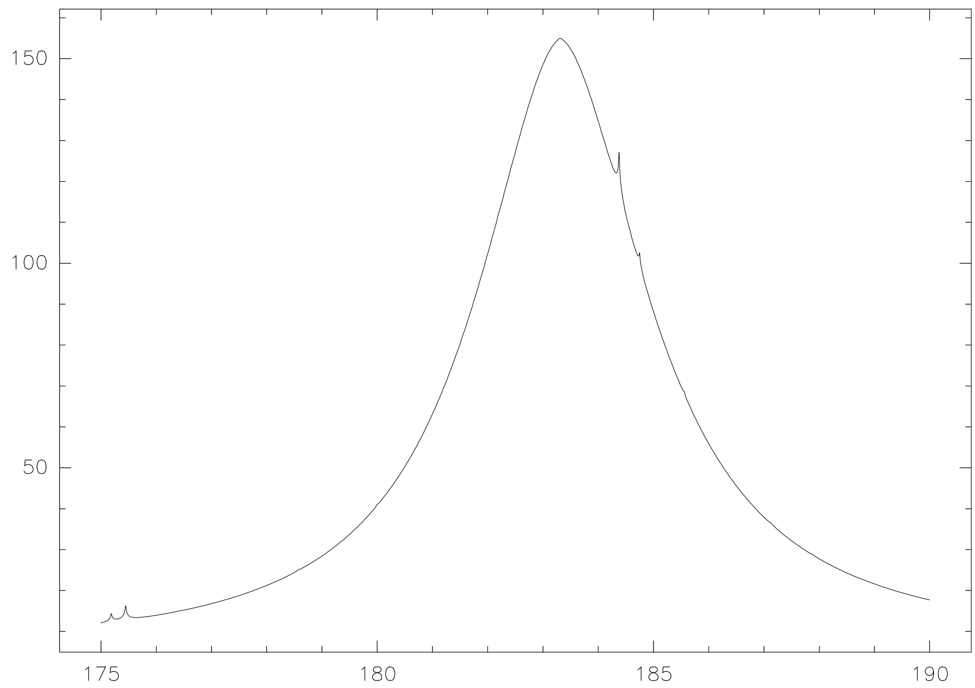


Figure 3.14. Emission from ozone added to the 183 GHz water line for 0.4 mm pwv. The ozone emission is approximately constant over time and space. The strongest line is at ∼184.3 GHz, with a line strength of just under 1 GHz K. The figure was created by Juan Pardo (private communication.

Frequency	$\Delta T/\Delta L$
22 GHz	0.4 K/mm
183 GHz	4 - 20 K/mm
220 GHz	1.7 K/mm

Table 3.3. Sensitivity to water vapour at different frequencies

3.5 Advantages and disadvantages of 183 GHz line

There are several advantages and disadvantages of the 183 GHz line over other monitoring frequencies:

3.5.1 Advantages

1. High sensitivity

   The 183 GHz line is much stronger than the line at 22 GHz or the emission at 220 GHz (see Fig. 3.1). As we have seen it gives a sensitivity of between ∼4 and 20 K/mm. In comparison, the 22 GHz line has a sensitivity of 0.4 K/mm and the continuum emission at 220 GHz will change by 1.7 K/mm (see table 3.3).

   Because of the high sensitivity of the 183 GHz line it is possible to build uncooled radiometers in spite of the higher thermal noise in comparison to cooled systems. An uncooled system is much easier to maintain, cheaper and more compact.

   The higher the observing frequency of the interferometer, the more accurately the excess path length needs to be determined. In contrast to the IRAM and OVRO interferometers, which operate at 3 mm - 1 mm wavelengths, the JCMT-CSO interferometer observes between 1.4 mm - 0.6 mm, and will soon operate at 0.4 mm, therefore the higher sensitivity of the 183 GHz radiometers is an advantage.

2. Optical properties

   In order to measure the water vapour along the same line of sight as that of the astronomical beam the water vapour monitors are mounted on the telescope and use the primary and secondary mirror of the telescope (for more detail see chapter 4 on optics). As the 183 GHz line is relatively close to the observing frequency, beam size, optical path and requirements on mirror accuracy are similar. The water vapour monitors can be installed without altering anything on the telescope optics. In comparison to optics at 22 GHz the feed horn and mirrors are much smaller for 183 GHz and it is much easier to accommodate them in the existing telescopes.

3.5.2 Disadvantages

1. Saturation

   As discussed in §3.3 the 183 GHz line already starts saturating in the line centre above 2 mm pwv (see Fig. 3.4). This is not the case for the much weaker 22 GHz line, nor for the emission at 220 GHz.
   $\rightarrow$ In contrast to the IRAM and OVRO sites, Mauna Kea is very dry and there is less than 2 mm of pwv 50% of the nights.
   $\rightarrow$ Rather than measuring the emission at the line centre, the water vapour monitors measure the brightness temperature in three frequency bands in the wing of the line, spaced 1.2, 4.2 and 7.8 GHz away from the line centre.

2. Conversion factor

   In §3.3 we have seen that the conversion factor $\Delta T/\Delta L$ does not only depend on frequency, but also on the underlying amount of water vapour and its physical conditions. Therefore we have to determine the conversion factor according to the atmospheric conditions.
   $\rightarrow$ Since the water vapour monitors measure the line at 3 different frequencies, the combined data contains information about the physical conditions.

Given the dry site and the short wavelengths the JCMT-CSO interferometer observes at, water vapour monitors operating at 183 GHz seemed best for our application.

Chapter 4

Optics and Calibration Loads

This chapter describes the optics of the water vapour monitors (WVMs). The first section (§4.1) will give a detailed account of the specific optical system: first describing the aims of the optics (§4.1.1), then giving a list of design constraints and their consequences (§4.1.2), briefly mentioning the mirror design and manufacturing (§4.1.3), describing the layout (§4.1.4), discussing the resultant beam (§4.1.5) and ending with the description of the alignment procedure and estimates of its accuracy (§4.1.6). Gaussian optics will describe the system to higher accuracy than simple ray optics. Section §4.2 will give a brief introduction to some principles of Gaussian optics relevant for the WVMs. Several ellipsoidal mirrors are used in the optics of the WVMs. These have some interesting features, which will be described in §4.3. Another mirror which needs a more detailed description is the moving flip mirror, which either directs the radiation from the sky or that from one of the calibration loads to the horn of the WVM (§4.4). The last section (§4.5) introduces the calibration loads.

4.1 Optics of water vapour monitors

4.1.1 Purpose of optics

- In order for the WVM beam to point in the same direction as the astronomical beam and to have the same size, the WVMs share the primary and secondary mirror of the telescope with the astronomical receivers. Since the WVMs are not an integral part of the receiver system optics are needed to direct part of the beam from the secondary mirror to the WVMs.

- The WVM beam and the astronomical beam are to overlap as much as possible, subject to the constraint that the WVM optics must not obstruct the astronomical beam.
- The WVM beams at the JCMT and the CSO telescopes should be similar in spite of the telescopes being different.
- An undistorted beam pattern is desirable.
- The optics must allow for switching between observations of the sky and the hot and warm calibration loads.

4.1.2 Design constraints and their consequences

Fig. 4.1 and 4.2 show the position of the WVM and the mirrors at the JCMT and the CSO telescope, Fig. 4.3 shows a close-up. The optical layout is a result of the following restrictions:

1. Direction of WVM beam

   Since the radiometers and the astronomical receivers have to observe simultaneously and share the primary and secondary mirror of the telescope, the beam needs to be either split in frequency (dichroic splitter) or the radiometers have to be mounted slightly off-axis. A dichroic splitter reflects radiation at certain frequencies and transmits radiation at other frequencies. However there are always some losses in such a splitter.
   → In order not to interfere with the performance of the telescope, the radiometers are installed off-axis of the astronomical beam.

2. Telescope structure

   The existing structure of the telescopes significantly influenced the optics of the WVM. The CSO and the JCMT are different, but we kept the optics of the monitors as similar as possible.
   → Thus the only difference between the WVM optics as installed at the JCMT as opposed to the CSO is the mirror (M3J/M3C) which couples the WVM beam to the

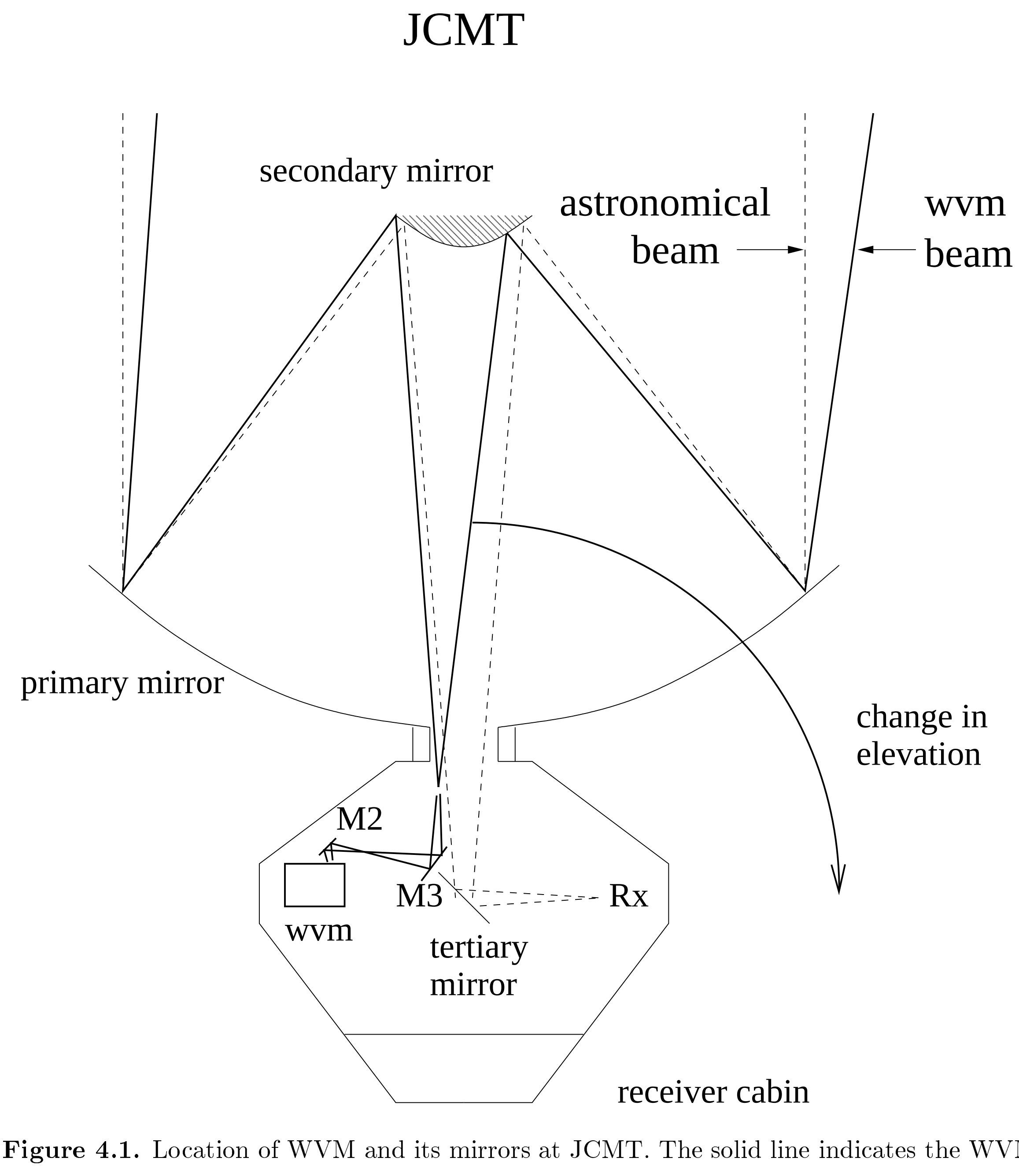


Figure 4.1. Location of WVM and its mirrors at JCMT. The solid line indicates the WVM beam, the dashed line represents the astronomical beam to the receiver (Rx).

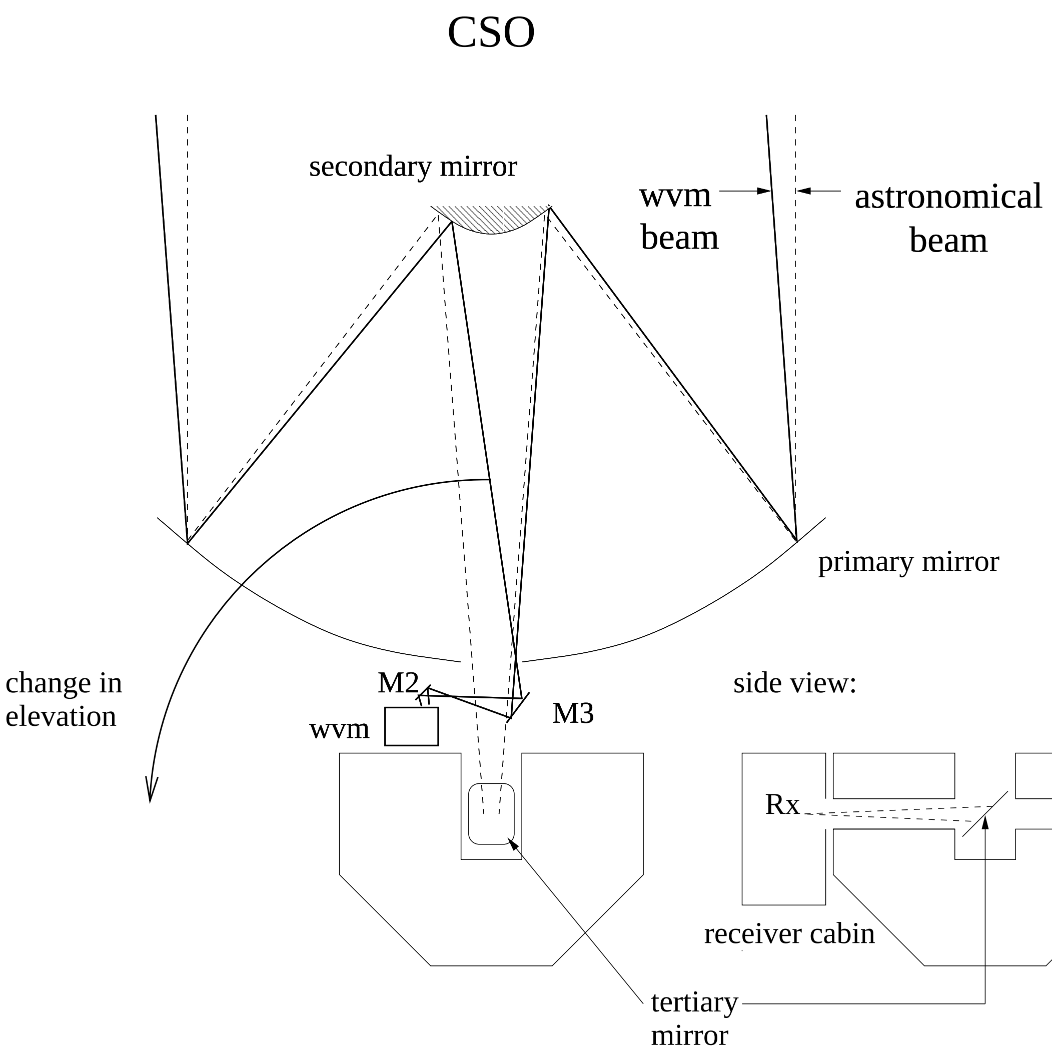


Figure 4.2. The WVM at the CSO is installed below the primary mirror outside the receiver cabin because the Rx cabin is very small. The solid line indicates the WVM beam, the dashed line represents the astronomical beam. The side view shows the location of the receiver cabin.

telescope.

A major constraint on both telescopes was the hole in the primary mirror. Looking out from the radiometers the beam hits the off axis mirror and then has to pass through this hole in order to reach the centre of the secondary mirror. Since the beam started off axis it obviously can't pass through the centre of the hole and still reach the secondary mirror.
The beam traverses the hole far from its centre.
→ The beam needs to be refocused in order to avoid obstructions.

3. Available space

The water vapour monitors need to be installed at a place from which the beam can easily reach the secondary mirror of the telescope. Most of these locations are already taken up by astronomical receivers.

The JCMT has a relatively large receiver cabin directly under the primary mirror. Nearly all the astronomical receivers are in rack spaces along the 12 walls of the receiver cabin. A tertiary mirror at 45° in the middle of the cabin can be rotated to point at any one of these receivers.
→ The WVM is mounted in one of these 19 inch racks (bay 10), as shown in Fig. 4.1. The electronics and the PC for data acquisition are in an adjacent rack, but can easily be moved if the space is needed.

The CSO receiver cabin is at the Nasmyth focus. However, there is no space for extra equipment.
→ Therefore the WVM and its power supply are mounted about 0.5 m below the primary mirror on the hexagonal platform, that supports the primary, see Fig. 4.2, thereby achieving a set-up comparable to that at the JCMT. Since this location is not enclosed during observing nights the data acquisition PC is located in the electronics room and connected to the WVM through a 30 m long cable of shielded twisted pair wires.

4. Size

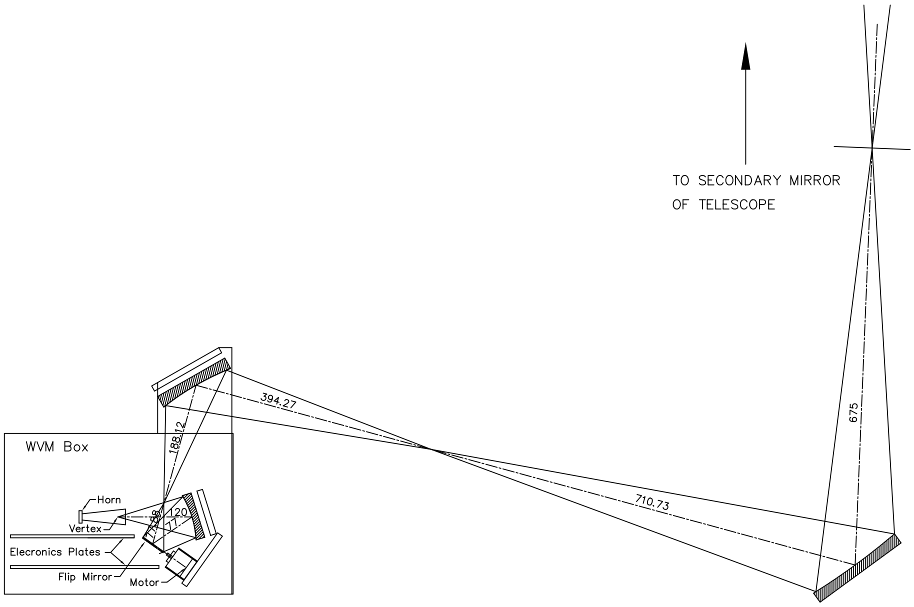


Figure 4.3. Optics of WVM

JCMT receiver cabin space is under considerable demand for new instruments, thus we aimed to keep the WVM as small as possible.
→ As the electronic components are relatively small the main emphasis lay on keeping the mirror system compact.

5. Flip mirror

   In order to calibrate the monitors the system contains a small flip mirror which switches between three positions: the sky, the hot calibration load and the warm calibration load. As the calibration occurs at 1 Hz, it is essential to keep the flip mirror light and small.
   → The beam has a minimum beam waist (focus) close to the flip mirror.

6. Calibration loads

   To reduce the size of the instrument the calibration loads have to be kept small.
   → The beam diameter at the loads has to be small.
   In order to prevent hot air escaping from the calibration loads
   → the loads point downwards (when the telescope is at zenith).

7. Horn

The size and shape of the monitor's horn determines the divergence of the optical beam.

8. Minimum distortion

   As discussed above, the beam needs to be refocused at several points in order to keep the beam size small. This can be achieved by lenses, but lenses have losses $\sim 5\% - 10\%$. Our system uses off-axis ellipsoidal mirrors for refocusing (see discussion in §4.3). However, they distort the beam shape. Using at least one more carefully chosen ellipsoidal mirror can correct for the distortion introduced by the first mirror.
   $\rightarrow$ One has to find an appropriate set of mirrors to focus the beam close to the calibration loads and the hole in the primary mirror and which also keeps the resultant distortion at a minimum.

4.1.3 Mirror design and realisation

As described above the geometry of the telescopes and the desired compactness of the water vapour monitors put quite a few constraints on the optical set up.

We used the ZEMAX ray tracing package (Focus Software Incorporated, 1998) to design the optical layout. First we determined some very different mirror set ups, all of which incorporated the constraints mentioned above. We then chose the most promising arrangement and used the program to optimise the position and ellipticity of the mirrors in order to achieve minimum beam distortion (see §4.3). We also calculated the beam waists using Gaussian beam mode optics (see §4.2). The mirrors have a diameter of 4 times the Gaussian beam waist so that only 0.03% of the power spills over the edge of the mirror. The surface tolerance is less than 10 μm, resulting in a reflectivity of better than 99.4%[1]. The parameters of the mirrors are listed in Tab. 4.1.

4.1.4 Optical layout

The resultant optical set up of the water vapour monitor on the JCMT telescope is shown in Fig. 4.1, Fig. 4.3 is a close up.

The path of light is direction independent, therefore it is just as valid to follow a beam outward from the horn of the radiometer to the sky, as it is to travel with a light ray from the

[1]According to the Ruze formula the power lost due to surface inaccuracies of a mirror σ is $1 - e^{-4\pi\sigma/\lambda}$ (Padman, 1998).

Mirror	ω [mm]	Mirror size [mm×mm]	angle	R_i [mm]	R_r [mm]	f [mm]
M1	17.0	69×75	$-17.2°$	126.2	169.2	72.3
flip	9.5	48×48	$-20.1°$	∞	∞	∞
M2	20.4	82×125	$-45.0°$	199.9	690.1	155.0
M3J	30.8	124×163	$+38.5°$	874.8	807.8	420.0
M3C	33.3	134×173	$+35.0°$	929.8	698.8	399.0

Table 4.1. Parameters of the WVM mirrors. ω is the beam waist radius at the position of the mirror. The angle is measured between the incident beam and the normal of the mirror. The radius of curvatures R_i and R_r were calculated using 4.4, the focal length f using 4.6.

atmosphere to the monitor. Here we will follow the light outward from the radiometer. The beam emerging from the horn first hits the ellipsoidal mirror M1 and is reflected onto the flip mirror. A motor can drive the flip mirror to three different positions. One mirror position directs the beam onto the warm load, one onto the hot load and the third onto the sky. Fig. 4.4 shows a side and front view of the water vapour monitor box indicating the position of the calibration loads. If the flip mirror is in the sky position the beam is directed to the ellipsoidal mirror M2, refocused and directed to mirror M3J. M3J sits slightly to one side of the main axis of the telescope in order keep clear of the astronomical beam (see Fig. 4.1). M3J directs the water vapour monitor beam onto the secondary mirror of the telescope. It then gets reflected back onto the primary mirror and out into the sky.

The water vapour monitor box on the CSO telescope contains exactly the same mirrors as that on the JCMT, except for mirror M3C, which has a slightly different focal length to compensate for the different geometry of the telescopes.

Fig. 4.5 shows a picture of the water vapour monitor mounted in the JCMT receiver cabin. One can see the pick-off mirror (M3J) in the foreground as well as the tertiary mirror of the telescope. Fig. 4.6 shows the radiometer mounted on the CSO. The bars are part of the support structure of the primary mirror. The monitor is seen in the background and the pick-off mirror in the middle ground.

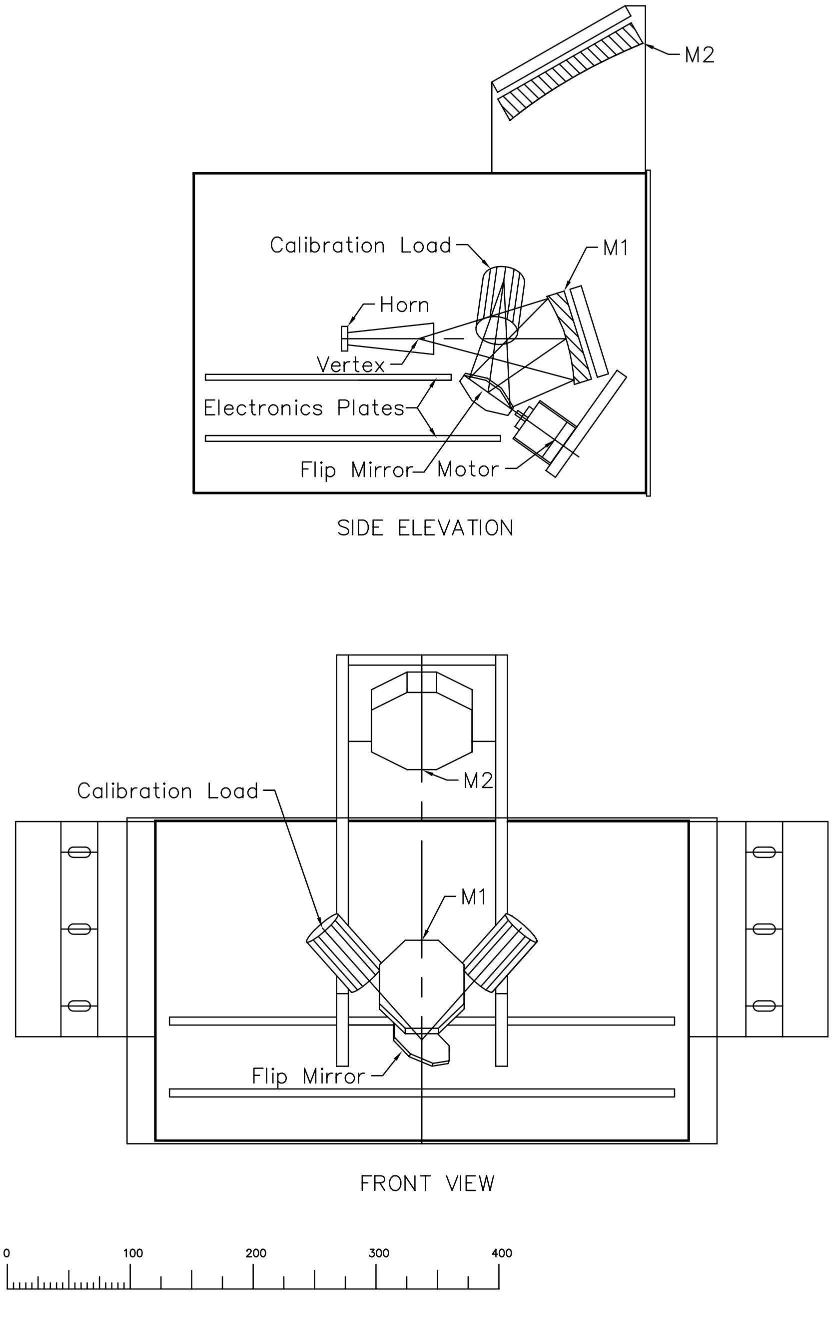


Figure 4.4. Position of the calibration loads. The upper drawing shows a side view, the lower drawing the view from the front. (The scale is in millimetres.)

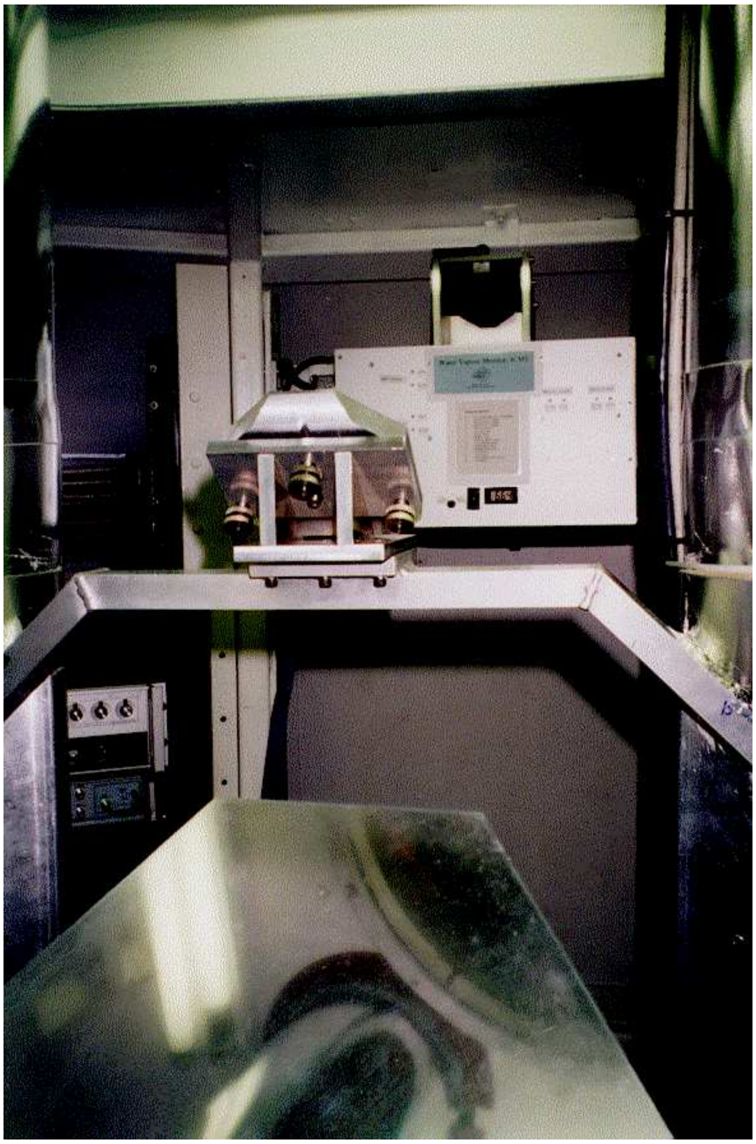


Figure 4.5. WVM mounted in JCMT receiver cabin. The tertiary mirror of the telescope is at the bottom of the foreground. (The reflection of the mirror shows the hole in the primary mirror, through which the WVM beam and the astronomical beam pass.) The pick-off mirror M3J can be seen in the middle ground on the support structure. In the background is the WVM.

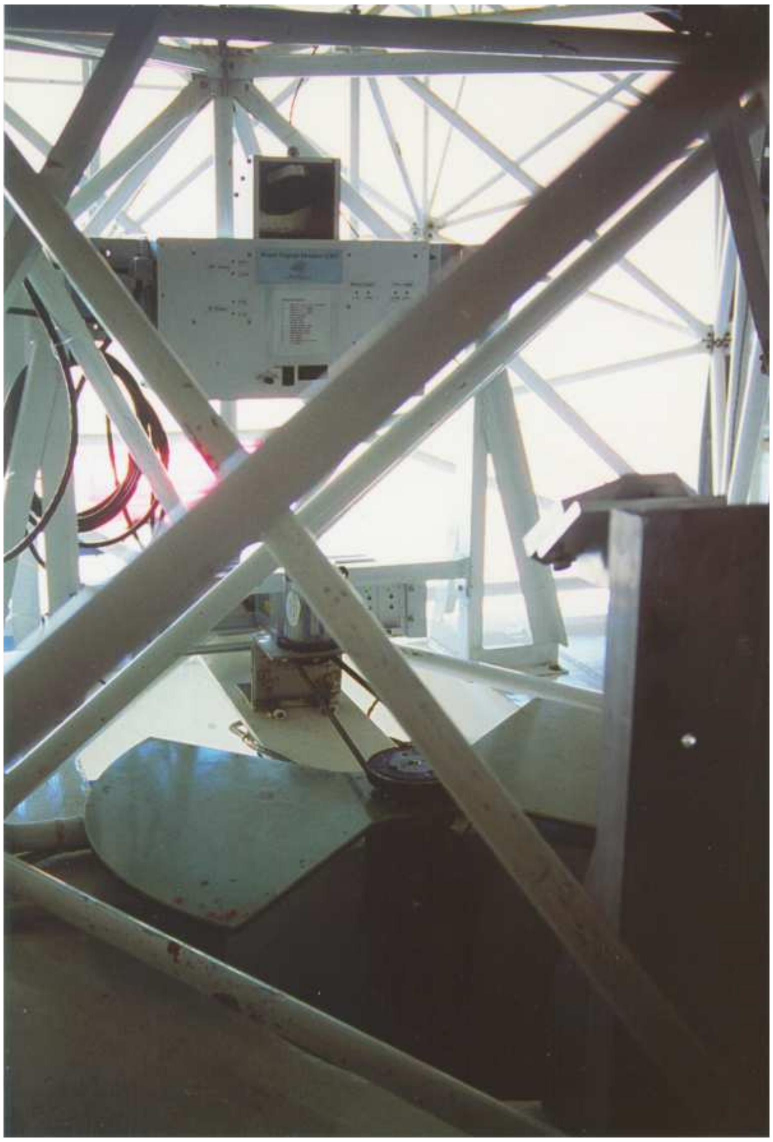


Figure 4.6. WVM mounted on the hexagonal plate of the CSO under the primary mirror. The many bars support the primary mirror. The WVM box is in the background. Underneath it is its power supply. The pick-off mirror M3C is on the right in the foreground. In the middle on the floor is a big chopper blade. It is behind a hole through which the astronomical beam travels to the receivers.

4.1.5 Theoretical beam

Firstly, since the pick-off mirror M3J is mounted slightly to one side of the telescope axis the WVM beam does not point in exactly the same direction as the astronomical beam. In fact at the JCMT the WVM points $7'\!.2''$ lower in elevation than the astronomical beam, at the CSO $11'\!.55''$ lower.

Secondly, we had to introduce a focus at the height of the primary mirror (for the beam to fit through the hole). Since the distance from this point to the secondary is shorter than the focal length of the telescope mirrors the the WVM beam is diverging by $2'\!.79''$ at the JCMT and $5'\!.38''$ at the CSO.

We expect most of the water vapour to be within the first 1000 m above the telescope (see Fig. 2.6), where the overlap of the two beams is very good: at one kilometre height the WVM beam is offset by 2.1 m from the JCMT astronomical beam, and 3.4 m from the CSO astronomical beam. Due to the divergence the beam size of the JCMT WVM increases from 15 m to 16.6 m, and the beam size of the CSO WVM from 10.4 m to 13.5 m. This small difference of the beams should not pose a problem in our applications.

4.1.6 Mirror alignment and accuracy

With the ray tracing computer program we calculated the exact positions and shape of all mirrors. Mounts and mirrors were built to this specification. However, each mirror can be tipped in the x and y plane, i.e. the direction of the optical beam can be changed. A simple experimental procedure was devised to align the mirrors: a small absorber dipped in liquid nitrogen can trace the beam because the detector has a minimum output when the cold absorber is in the middle of the beam. The adjustment is started with mirror M1, the first mirror after the horn, and then proceeds on to M2 and M3J or M3C, always making sure the beam falls into the centre of the next mirror or the calibration load.

This alignment technique allows to centre the beam to a precision of better than $\sim$5 mm. Since the mirrors have a radius of 4 times the beam waist, an offset of 5 mm from the centre would result in a loss in beam power of less than 0.4%.

The adjustment of the pick up mirror (M3J or M3C) is slightly more difficult as it is nearly impossible to hold an absorber in front of the secondary mirror of the telescope, especially at the JCMT. In order to solve this problem, I calculated the coordinates of the desired beam about 1.5 m above the mirror relative to the telescope structure and made sure that the actual beam passes through this position. This alignment was fine tuned at the JCMT by maximising the flux while the telescope was pointing at the sun. As the sun has a diameter of 0.5° the offset of $5\overset{\prime}{.}13''$ between the telescope beam and the radiometer beam is not of importance. If, however, the radiometer beam was not aligned on the secondary mirror the detected signal would be much lower, since the beam would then miss the sun completely. Since the CSO is not supposed to point at the sun, as this will cause distortion of the telescope structure, an alternative method was needed. A circular piece of absorber was placed in the centre of the secondary mirror. Compared to the cold sky in good weather the absorber looks warm to the detector. We adjusted the mirror by maximising the detector output.

The distance between the mirrors is not adjustable. The mirror positions inside the water vapour monitor box should be better than 5 mm, the largest error will be between mirror M2 and M3J (or M3C), maybe up to 20 mm. If the separation of the mirrors is different from the theoretical values the focal points will be shifted. This has three effects: Firstly, it might cause a larger (or smaller) beam size at the subsequent mirror and more (or less) beam power will spill past the mirror. These losses will however be small (see below for total loss). Secondly, the focal point of the WVM beam will be shifted. Assuming a rather large shift of 50 mm the divergence will differ by only $\pm 2''$ in the case of the JCMT, $\pm 4''$ in the case of the CSO. Neither of these errors would significantly change the beam. Thirdly, the arrangement of the ellipsoidal mirrors will not be optimised for minimum distortion (see §4.3 on ellipsoidal mirrors). The water vapour monitors will then weight the intensity of the beam and some parts of the beam will contribute more than others. However, this effect should be very small.

It is difficult to experimentally determine the detailed shape of the water vapour beam leaving the telescope. The losses from imperfectly polished mirrors and misalignments are expected to be less than a few percent. It is however possible to measure the coupling of the radiometers to the sky, which indicates how much power is lost in the optical path between the detector horn and the sky. The measured losses in total were $\sim$5% at the JCMT and $\sim$25% at the CSO. These losses are mostly due to blockages, e.g. the secondary mirror of the CSO is covered with an absorber which on its own introduces a loss of $\sim$22%.

4.2 Gaussian Beams

In general, electromagnetic radiation passing through an aperture undergoes diffraction. A full description can be obtained by evaluating the Fresnel integrals. However, when we are dealing with relatively small angles (much less than one radian) i.e. if the radiation travels along a paraxial beam and if it can be represented by a scalar field distribution, the diffraction pattern can be truncated at the aperture. Just the fundamental Gaussian mode will yield a good description of the radiation and considerably simplify any calculations. (If in addition the wavelength of the radiation is much smaller than the optical elements and focal lengths, ray optics gives an even simpler description.) The propagation of millimetre and submillimetre optics are usually best described by Gaussian beam mode optics.

This section will only introduce the very basic formulae, which were used in the design of the optics for the water vapour monitors. A detailed description can be found in Lesurf (1990) for example.

Paraxial beams can be represented by a scalar field distribution of the intensity of radiation. The scalar wave equation is:

$$\bigtriangledown^2\Psi + k^2\Psi = 0 \tag{4.1}$$

where Ψ is the scalar radiation field,
$k = 2\pi\nu/c$ is the wave number and
ν the signal frequency.

Under the assumption that Ψ varies much more in the x and y direction than in the z direction the beam travels along, an expression for Ψ can be found. Assuming circular symmetry of the beam and that the radiation is only in the fundamental mode the scalar field Ψ can be further simplified and we get three relations:

$$\Psi \propto exp(\frac{-r^2}{\omega^2}), \tag{4.2}$$

$$\omega^2 = \omega_0^2[1 + (\frac{\lambda z}{\pi\omega_o^2})^2], \tag{4.3}$$

$$R = z[1 + (\frac{\pi\omega_0^2}{\lambda z})^2], \tag{4.4}$$

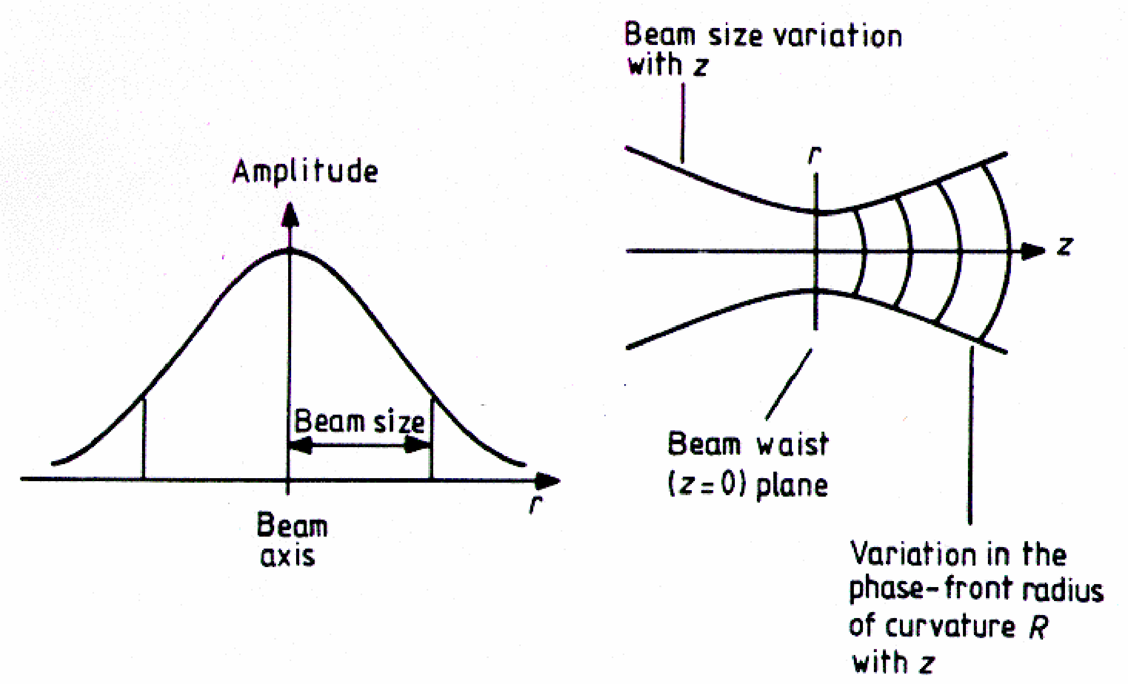


Figure 4.7. Cross section of Gaussian field pattern and variations of size and phase-front curvature along the beam. Figure taken from Lesurf (1990).

where	r	is the distance to the centre of the beam perpendicular to the direction of the beam, i.e. in the x-y plane,
	ω	is the beam (waist) radius,
	R	is the radius of curvature,
	λ	is the signal's wavelength,
	z	is the distance to the minimum beam waist and
	ω_0	is the minimum beam waist radius.

Fig. 4.7 illustrates how the field varies across the beam and how the beam radius changes. According to the first equation the amplitude variation across the beam has a Gaussian shape. The beam (waist) radius ω is the distance at which the amplitude has dropped to 1/e. The beam radius has a hyperbolic shape along the direction of the beam as described by the

second equation. There is a minimum beam waist size ω_0; z is defined to be zero in the plane of this minimum beam waist. The beam is not plane parallel, but has a changing radius of curvature R. Along a spherical surface of radius R the radiation has the same phase.

Mirrors in the optical paths basically change the direction of the beam and the radius of curvature, therefore defining a new minimum beam waist radius ω_0 and a new position for $z = 0$.

The beam radii were calculated to determine the size of the mirrors and to verify that no section of the telescope structure would obstruct any part of the beam.

4.3 Ellipsoidal Mirrors

Each water vapour monitor has three ellipsoidal mirrors (M1, M2 and M3J or M3C). This section will describe the reflection properties of off-axis ellipsoidal mirrors and the distortion of the beam pattern they cause. This distortion can be corrected with one or more ellipsoidal mirrors, if the parameters of these are chosen carefully. Using the example of the water vapour optics we will look at the effect each mirror has on the beam pattern.

4.3.1 Reflection properties and beam distortion

In order to understand the effect of an ellipsoidal mirror on a bundle of light rays, consider for simplification an ellipse (2-dimensional) first, Fig. 4.8. The ellipse has two foci, any ray passing through one of the foci will return through the other focus after reflection on the inside wall of the ellipse. If we start with a bundle of rays spaced at equal angles all passing through focus 1, they reflect on the boundary of the ellipse at different positions, but all pass through focus 2. The outgoing rays, however, are not spaced at equal angles any more. The beam pattern is distorted. One can represent this effect by drawing the beam pattern: Imagining a surface perpendicular to the incoming ray, the position where each ray would pierce this surface is marked by a dot. For the incoming rays we get a straight line of equally spaced dots (for small angles). In the three dimensional case we get a two dimensional array of equally spaced dots, Fig. 4.9 (a). The corresponding diagram of the outgoing beam has unevenly spaced dots, Fig. 4.9 (b), as the rays are not separated by equal angles any more. These diagrams show changes in the beam patterns.

The beam drawn in Fig. 4.9 b) illuminates the atmosphere unevenly. Let's imagine our

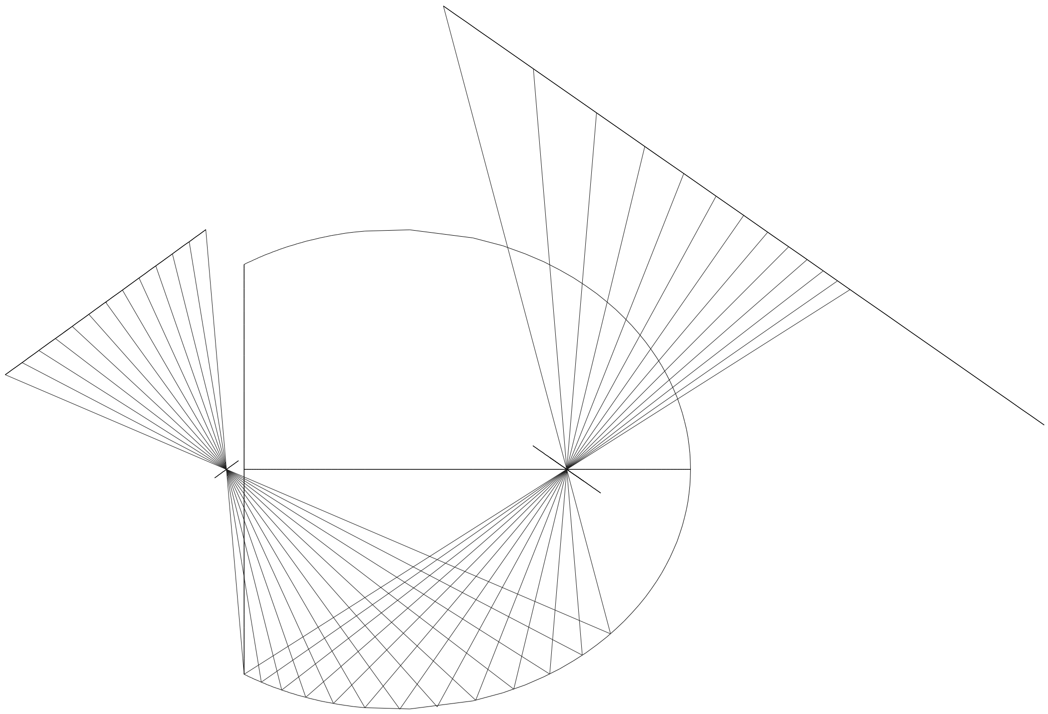
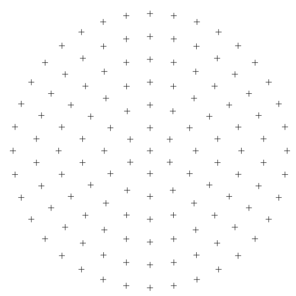
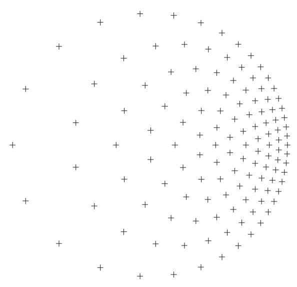


Figure 4.8. Rays reflected on an off-axis elliptical mirror. The incoming rays are evenly spaced in contrast to the reflected rays.



(a)



(b)

Figure 4.9. Beam pattern (a) of undistorted beam (b) of distorted beam after reflection on an off-axis ellipsoidal mirror depicted in Fig. 4.8

receiver is at focus 1, there is an ellipsoidal mirror and the sky is beyond focus 2. According to the beam pattern (Fig. 4.9 (b)) more rays (i.e. light) from the right will be focused on the receiver than from the left. Therefore more emphasis (or weight) will be given to the right than to the left part of the sky. As long as the sky has the same temperature, i.e. brightness, everywhere each beam pattern will give the same integrated intensity. But if the sky is, for example, brighter on the right than on the left, beam (b) will measure a higher intensity than (a). Therefore it is desirable to keep such distortions small.

4.3.2 Correction of distorted beams

The properties of ellipsoidal mirrors have been investigated and correction schemes have been suggested. For example, Dragone (1978) shows that a second ellipsoidal mirror can invert the distortion introduced by the first ellipsoidal mirror. This is obvious if the second mirror is the same as the first. But distortions may also be corrected with a different ellipsoidal mirror or a set of mirrors: Murphy (1986) has defined the distortion ε:

$$\varepsilon = \frac{\omega \tan i}{f}, \tag{4.5}$$

where	i	denotes half the angle between the incident and the reflected beam and
	ω	is the waist radius at the mirror,
	f	is the focal length of the mirror, with

$$\frac{1}{f} = \frac{1}{R_i} + \frac{1}{R_r} \tag{4.6}$$

where	R_i	is the radius of curvature of the incident beam and
	R_r	is the radius of curvature of the reflected beam given by Equ. 4.4.

The distortion ε can be computed for every ellipsoidal mirror. If the sum of the ε for all mirrors along the optical path is zero the beam pattern of the incident and reflected beam are identical. (Note that the angle i can be positive or negative, hence ε can take positive or negative values.) Thus, choosing a set of mirrors giving ε=0, ensures that there are no distortions in the beam pattern and that the detected radiation is unbiased.

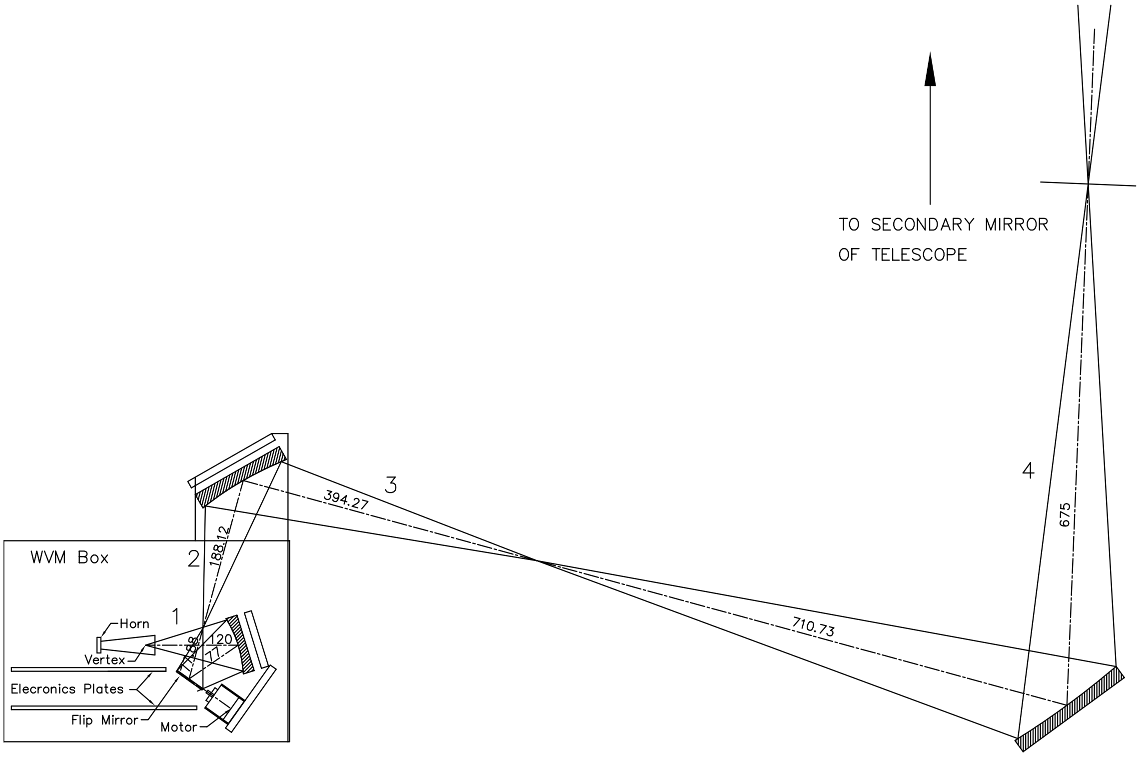


Figure 4.10. Positions 1 to 4 at which beam patterns are drawn.

4.3.3 Beam patterns of water vapour monitors

In order to demonstrate the effect ellipsoidal mirrors have on the beam pattern, the distortion at the different locations of the WVM beam marked in Fig. 4.10 have been computed and are depicted in Fig. 4.11.

Quantitative results are contained in Tab. 4.2, where the properties of the mirrors and the resulting distortions ε are listed. The sum of the distortions is zero for the optics at the JCMT and the CSO. As seen in Fig. 4.11 (5), the beam leaving the telescopes is nearly undistorted.

4.4 Flip Mirror and Motor

Another special mirror in the WVMs is the flip mirror, which is turned by a small motor during calibrations.

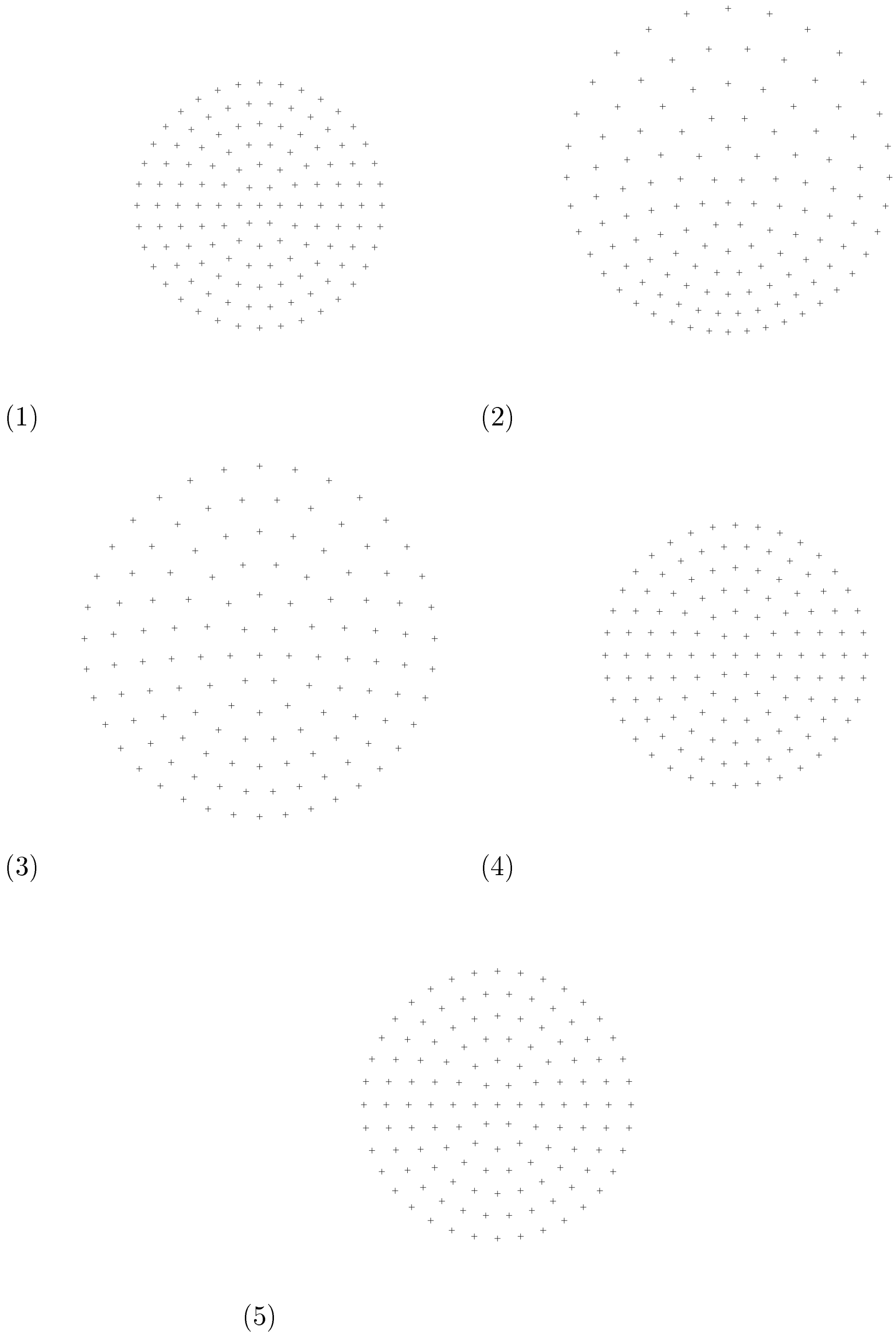


Figure 4.11. Beam pattern (1) emerging from the horn, (2) after reflection on M1, (3) after reflection on M2, (4) after reflection on M3J, (5) on the sky. The positions of the beam patterns are indicated on Fig. 4.10.

Mirror	ω_0(before) [mm]	ω_0(after) [mm]	ω [mm]	angle	f [mm]	ε
M1	3.77	4.96	17.0	$-17.2°$	72.3	-0.073
M2	4.96	13.4	20.4	$-45.0°$	155.0	+0.131
M3J	13.4	12.5	30.8	$+38.5°$	420.0	-0.058
M3C	13.4	10.4	33.3	$+35.0°$	698.8	-0.058

Table 4.2. Mirror parameters and the distortion ε. ω_0(before) is the minimum beam waist of the radiation before it reaches the mirror, ω_0(after) after it has been reflected. ε was calculated using Equ. 4.5. Note that the sign of ε is $(-1)\times$ the sign of the angle for M2 and M3J/M3C because the reflection on the flip mirror inverts the beam.

4.4.1 Flip Mirror

The flip mirror is a small (50 mm x 50 mm), thin (2.5 mm) flat mirror made of aluminum, weighing only 17g. It is mounted on a shaft on the rotary axis of the motor.

4.4.2 Motor

We use a proportional rotary solenoid to drive the mirror to the three positions: sky, hot load and warm load. A proportional rotary solenoid rotates due to a torque proportional to the driving current. The motor's position is monitored by an angle position sensor, whose output voltage is proportional to the position of the mirror.

4.4.3 Motor control

The control circuit, which drives the motor, compares the voltage from the position sensor to an input voltage characteristic of the correct mirror position. As long as there is a difference in the voltages a current is applied to drive the motor. As soon as the correct voltage and therefore position is reached the driving current is turned off.

A logical signal from an industrial PC, which is also used for data acquisition, activates the mirror control. This logical signal is then translated to the appropriate voltage.

The mirror needs about 40 msec to flip from one position to the next and about 60 msec between the loads, see Fig. 4.12. The 40 ms are measured from the time the motor starts

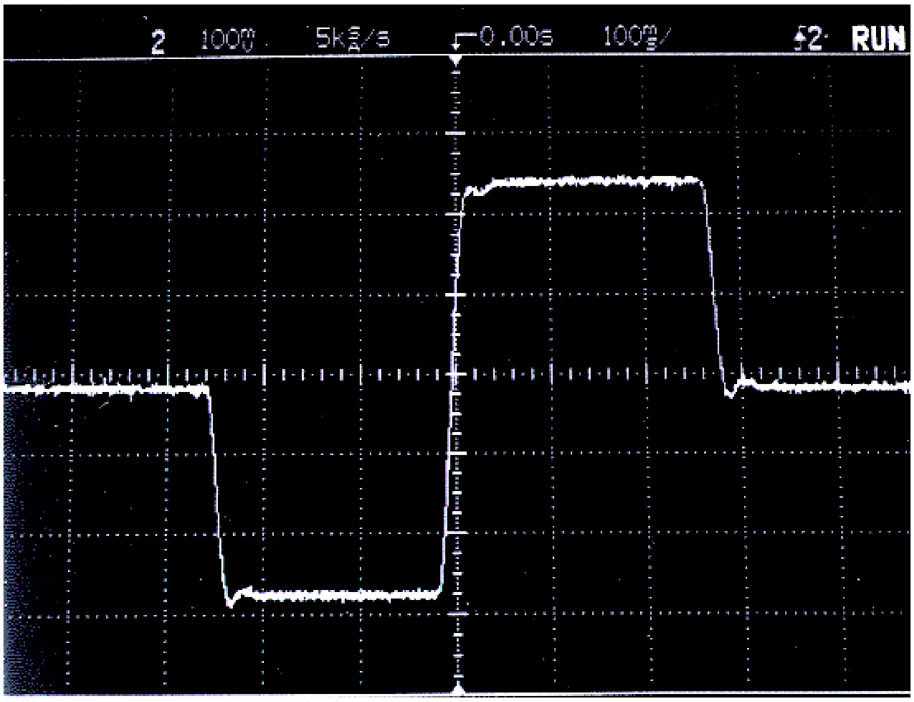


Figure 4.12. Position of the flip mirror when switching from sky to hot load to warm load and back to sky. The switching takes about 20 ms but is followed by a damped oscillation. Therefore no data is taken for 40 ms when the mirror moves between the loads and the sky position and for 80 ms when the mirror moves from the hot to the warm load.

moving to the time the mirror has reached a position where essentially all of the beam power reaches the load or sky. A program run by the PC determines when the mirror is to move from one position to the next. Usually the mirror switches through all 3 positions in 1 second, typically resting 200 msec on each load and 400 msec on the sky. In the remaining time no data is taken as the mirror is moving.

Though the JCMT mirror control was operating as expected, we had unfortunately some initial problems with the mirror control on the CSO. It turned out that the mirror position depends on the temperature of the motor, probably because the friction of the motor is higher at low temperature and the same applied torque does not move the mirror as far. Larger temperature variations are encountered at the CSO WVM than at the JCMT WVM because the CSO WVM is mounted outside. Consequently we thermally insulated the motor. In addition an integrator circuit was added to the motor control, which measures the mean error in each position and corrects for it. The circuit ensures that the motor always moves by the same distance. The motor control now works well.

4.5 Calibration Loads

The water vapour monitors have two calibration loads each, a warm load at $\sim 30^\circ$ C and a hot load at $\sim 100^\circ$ C. We chose not to use a cooled load in order to keep the system compact and the maintenance simple.

The design of the loads is adapted from Keen (1995). The loads are 55 mm long aluminum cylinders with a conical cavity, see Fig. 4.13. The cavity has an opening of 25 mm, about 5 times the radius of the beam waist. The cavity is filled with a thin (2 mm) layer of epoxy, which acts as a black body at millimetre wavelengths, i.e. it radiates at its physical temperature. The aluminum cylinder is surrounded by a heater mat and a subsequent layer of ceramic fibre and aluminum foil for insulation. A temperature sensor at the tip of the cone measures the physical temperature of the load. The sensor signal is used in the heater circuit, which increases the heating when the temperature drops below a certain value and decreases it when the temperature rises too high. The resultant temperature as seen by the sensor is slightly oscillating, but the amplitude of the oscillation is less than 11 mK (Fig. 4.14).

The loads are mounted on the same support structure as the mirrors and the electronic plates (see Figuress 4.4 and 4.15).

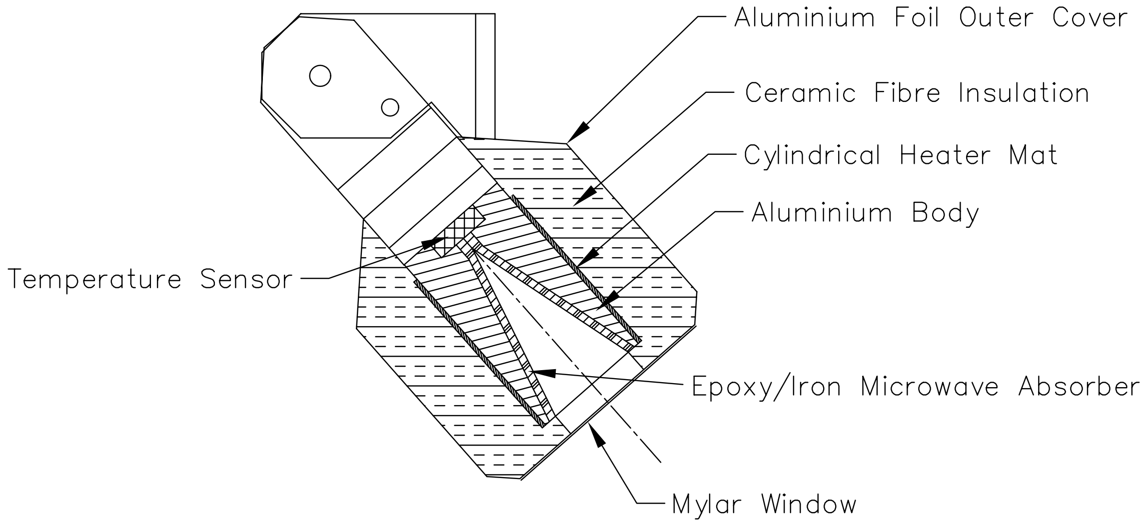


Figure 4.13. Drawing of calibration load.

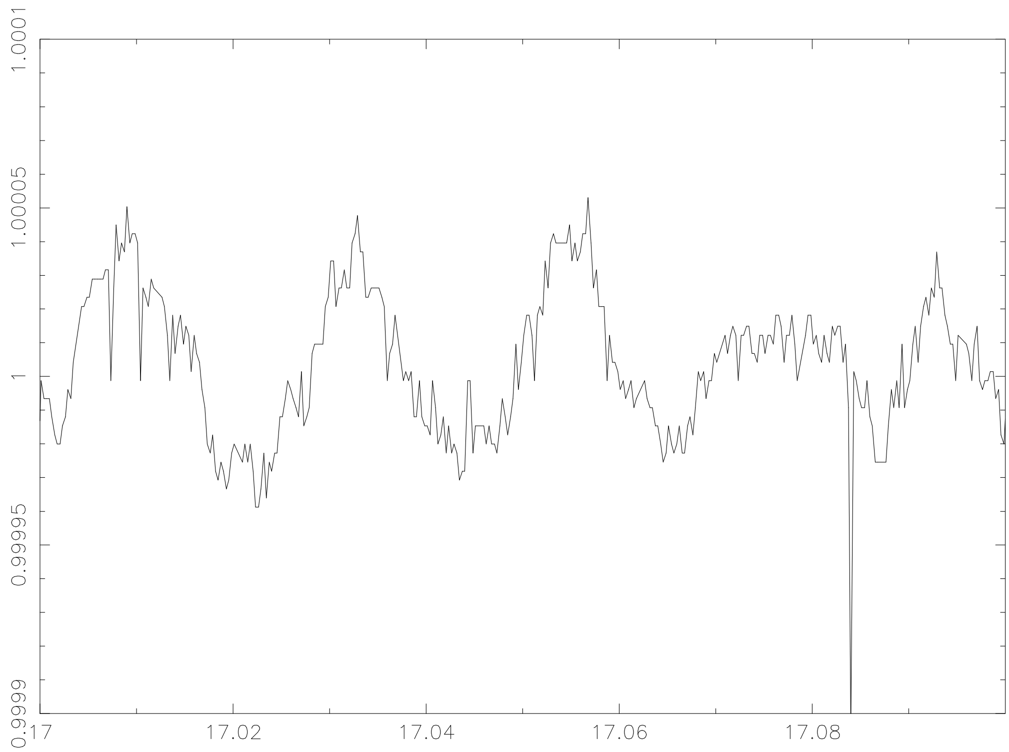


Figure 4.14. Temperature of the hot load. Oscillations on the temperature are caused by the heating circuit. (The spike is an arifact.) The amplitude of the oscillations is < 10 mK. (Full scale of the plot corresponds to 20 mK.)

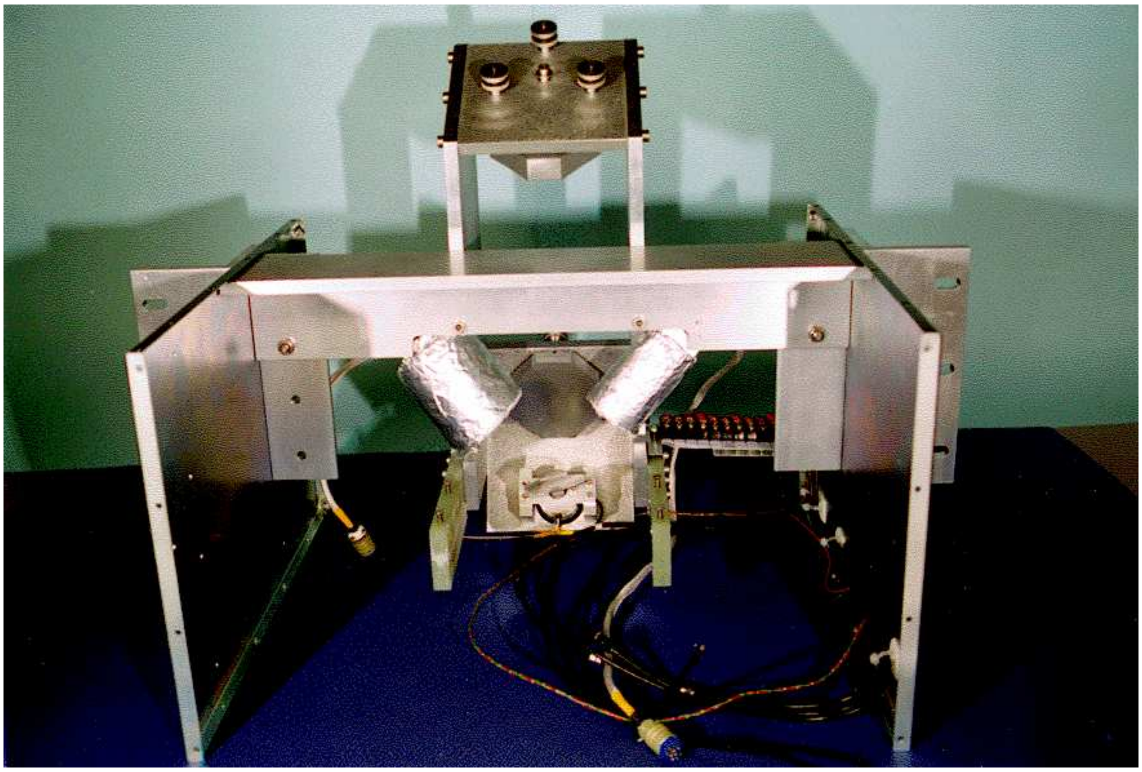


Figure 4.15. View from the back of the WVM: the hot load is on the left, the warm load on the right. Behind the loads is the mirror M1, above the mirror M2.

Chapter 5

Electronics

This chapter describes the electronics of the water vapour monitors. The underlying principle is that of any high frequency receiver: a high frequency signal at a few hundred GHz (receiver frequency RF) is mixed down to an intermediate frequency (IF) signal at a few GHz, amplified and detected. A computer digitises and stores the detected signal. However, the specific design, the components chosen and detailed layout are unique to the water vapour monitors.

In this chapter I will first (§5.1) state the requirements of the radiometers and quickly mention how they influenced our design. Secondly (§5.2), a brief overview of the electronic set-up is given followed by a more detailed discussion of some of the electronic components which were specifically chosen for our application. Section §5.3 outlines the temperature regulation of the electronics. The computer set-up, which runs the water vapour monitors and takes and stores the data, is described in §5.4. At the end of the chapter (§5.5) the performance of the radiometers is summarised.

5.1 Requirements

Here is a list of requirements which influenced the design of the water vapour monitors:

1. Sensitivity

   We aim to reduce the phase fluctuations due to water vapour to 18° at 500 GHz, in order to limit the loss of the astronomical signal due to decorrelation to 10 % ($(\sin 18)^2$). Therefore we need to determine the difference in optical path length between the two telescopes to c/500 GHz×18°/360°=30μm rms, and the optical path for each telescope to 21μm ($= 30\mu m/\sqrt{2}$). As explained in §3.3.4 the sensitivity $\Delta T/\Delta L$, i.e. the change in brightness temperature ΔT given a change in optical path length ΔL is not constant but depends on the saturation of the water line and on the measuring frequency. Assuming an average sensitivity of 10 K/mm for channel 1, a change in optical path of 21μm corresponds to a change in brightness temperature of 0.2 K.

→ The brightness temperature should be measured to an accuracy of 0.2 K.

→ For an IF bandwidth of 200 MHz and an integration time of 1 second a system temperature of 2500 K yields a theoretical sensitivity of 0.18 K, which is sufficient for our application. Therefore we can afford to have an uncooled system with a noise temperature around 2500 K.

2. <u>Stability</u>

   Besides reducing the noise as far as possible, the gain between the input signal and the detector output needs to be constant or at least well known.

   → The electronic components have been selected to be insensitive to temperature changes, especially the detectors.

   → In addition the electronics are kept at a constant temperature, as the gain of amplifiers varies with temperature. Heaters and fans were implemented for this purpose.

   → Frequent calibrations (one per second) are performed to determine the gain.

3. <u>Frequency passband</u>

   The frequency bands must be well determined, i.e. the width of the passband as well as its frequency:

   → A phase lock loop (PLL) insures that the local oscillator (LO) frequency stays constant.

   → The LO operates at the centre frequency of the emission line. Since a double sideband measurement is taken, a slight increase in the LO frequency results in a decrease in brightness on the low frequency side, but an increase in the brightness temperature on

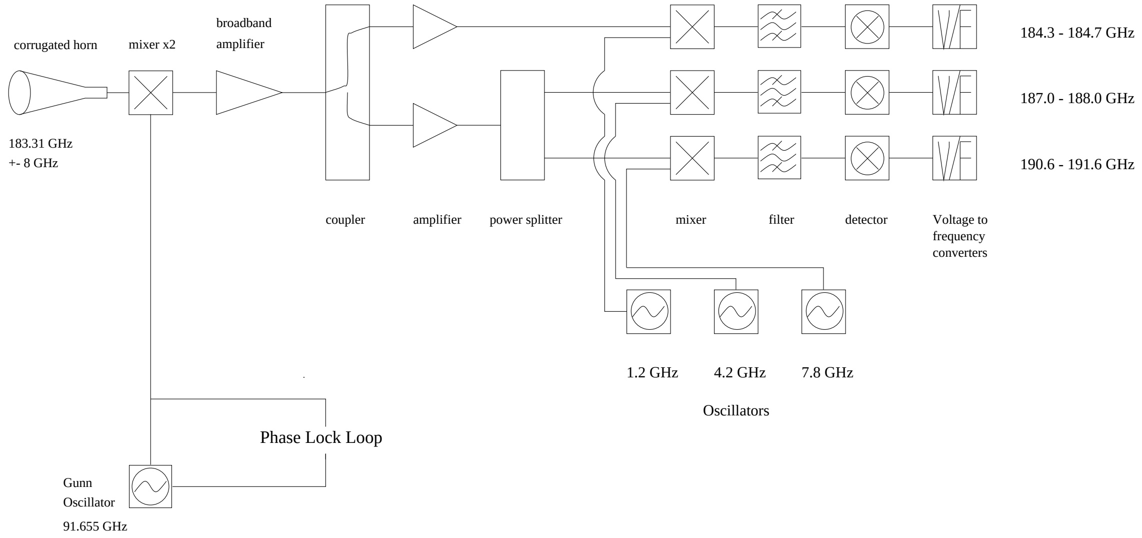


Figure 5.1. Schematics of WVM electronics.

the high frequency side. To first approximation the sum of the signals stays constant.
→ In order to determine the frequency band accurately the signal is mixed down a second time and then passed through low pass filters (see §5.2.2).

4. Easy maintenance and low cost
Last but not least we wanted to keep the construction and the maintenance of the water vapour monitors simple and low cost:
→ Neither the electronics nor the calibration loads are cooled.
→ The Gunn oscillator (Gunn) works at a fixed frequency.

5.2 Description of electronics

According to the requirements stated above we designed and realized the circuit shown in Fig. 5.1. The radiation emitted by the atmospheric water vapour passes through an arrangement of mirrors and is focussed on the horn (see chapter on WVM optics §4). The subharmonic double sideband mixer multiplies the second harmonic of the LO (i.e. a signal

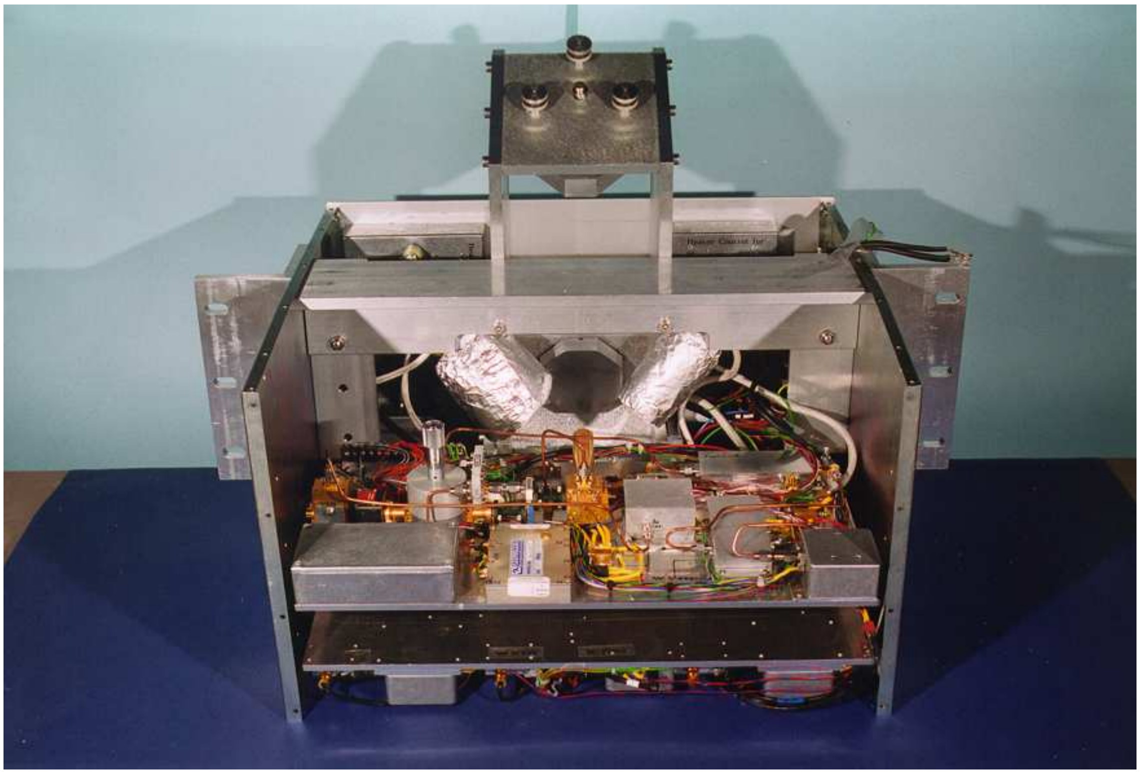


Figure 5.2. The WVM box with the two electronics plates. The upper plate is the RF plate. The horn and mixer are in the centre of the RF plate, the horn is pointing away from the photographer to mirror M1. The lower plate contains the IF electronics.

at twice the LO frequency) with the incoming signal. The signal at the difference frequency, the intermediate frequency (IF), is amplified and then passed to the IF plate, where it is split into three and amplified again. Each branch of the signal is mixed a second time with an oscillator signal at 1.2, 4.2 and 7.8 GHz respectively. Each of the resulting signals passes through a low frequency filter and is measured in a Schottky detector. The voltage to frequency converters (V/F converters) transform the output voltages into frequencies. These frequencies are transmitted to a computer, where they are counted and converted into digital signals, which are further processed.

Physically the electronics is mounted on two temperature regulated aluminum plates. The “RF plate” contains the receiver frequency (RF) components and the first broadband amplifier (Fig. 5.3 and 5.4). The “IF plate” with the intermediate frequency (IF) components

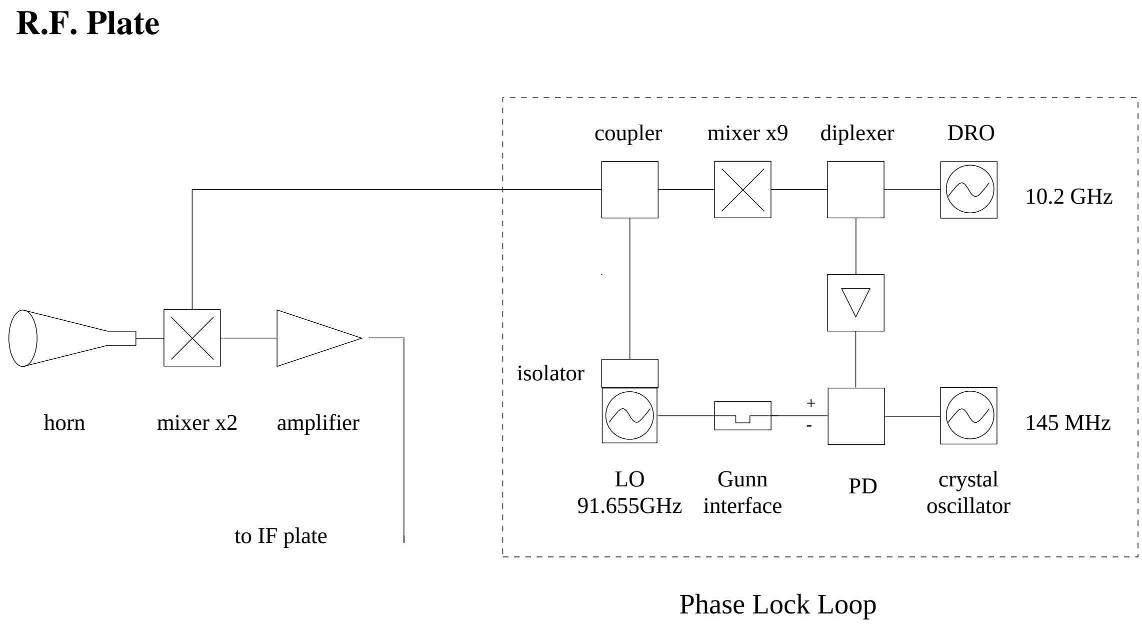


Figure 5.3. Receiver frequency (RF) electronics.

(Fig. 5.5 and 5.6) is located upside down under the "RF plate". Fig. 5.2 shows a picture of a water vapour monitor with the two plates. I will now discuss some of the components in more detail.

5.2.1 RF plate

Horn: The horn focuses the sky signal and passes it through a waveguide to the mixer. To avoid any current flow along its surface the 76 mm long horn is corrugated, i.e. ~0.5 mm wide rings are carved in the inside of the horn. Colleagues from the Rutherford Appleton Laboratory (RAL) built the two corrugated horns.

Mixer: The mixer mixes the signal from the Gunn oscillator and the sky to to an intermediate frequency (IF), which is much easier to work with. One can think of the mixer as lowering the RF signal by subtracting the constant frequency produced by the Gunn. Each water vapour monitor uses a double sideband subharmonically pumped mixer built by RAL. Double sideband means that both the upper and lower sidebands (e.g. 183+5 GHz and 183-5 GHz) are mixed down and contribute to the IF (5 GHz). A subharmonically pumped mixer doubles the Gunn frequency before subtracting the

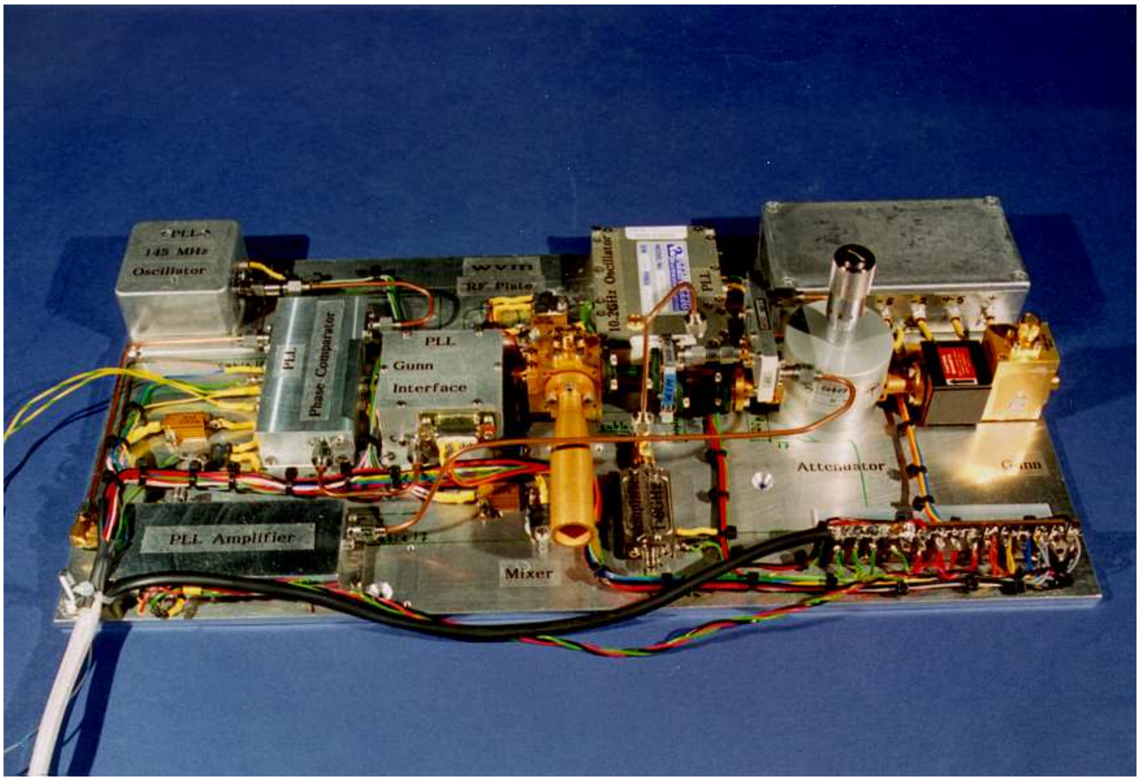


Figure 5.4. Picture of RF electronics.The horn is in the middle pointing towards the photographer. Just behind the horn is the mixer. The Gunn is on the very right of the plate, left of it is the cylindrical attenuator with a micrometer screw for tuning. On the left of the mixer are the different components of the PLL. The first stage broad band amplifier is in the foreground on the right of the horn.

RF. The mixers always operate at the same frequency so can be fixed tuned. Horn and mixers together have a noise temperature of about 1500 K - 2600 K depending on the frequency.

Gunn oscillator: The commercial Gunn oscillator, also called local oscillator (LO), operates at exactly half the centre frequency of the water line, i.e. at 91.665 GHz. By choosing this frequency and using a subharmonic double sideband mixer, small shifts in the oscillator frequency should not change the output dramatically, see 5.5.2.

Phase lock loop (PLL): The basic task of the phase lock loop is to ensure that the LO frequency does not change, but it also keeps the conversion (i.e. mixing capability) of the mixer constant. 10 dB [1] of the LO signal is coupled out before the mixer. This signal is mixed with a harmonic of a stable fixed frequency reference oscillator. In our case the ninth harmonic of 10.2 GHz produced by a dielectric resonance oscillator (DRO) is mixed with the LO signal. The difference between the two oscillator frequencies should be 145 MHz ($9 \times 10.2 - 91.655$ GHz). This difference is amplified and then compared to a very stable 145 MHz crystal oscillator in the phase detector (PD). If there is a difference between the two frequencies the PD produces a voltage proportional to this difference, which is added to the voltage driving the Gunn. A change in supply voltage will make the Gunn oscillate at a slightly different frequency. In this way the LO frequency is constantly monitored and adjusted.

Broad band amplifier: The first amplifier after the mixer is a commercial, low noise, broad band amplifier. It is important that there is little noise added by this amplifier, especially because the mixer has a conversion loss of about 7 dB, so that the sky signal is reduced by a factor of 0.2. The amplifier enhances the sky signal as well as any noise. (One can think of the system temperature being increase by a factor of 1/0.2 due to the loss in the mixer.) Once the signal from the sky has been amplified, added noise will usually only constitute a small percentage of the signal and not be of great importance. Since the radiometer probes the water line at several frequencies in the line wing (but not at the line centre) the amplifier needs to cover a frequency range from around 1 to 8 GHz (broad band). We choose an amplifier of nearly constant gain

[1] A deci Bell (dB) is defined as a ratio of power P_1 and P_2 such that x dB is equivalent to $\frac{P_1}{P_2} = 10^{-\frac{x}{10}}$.

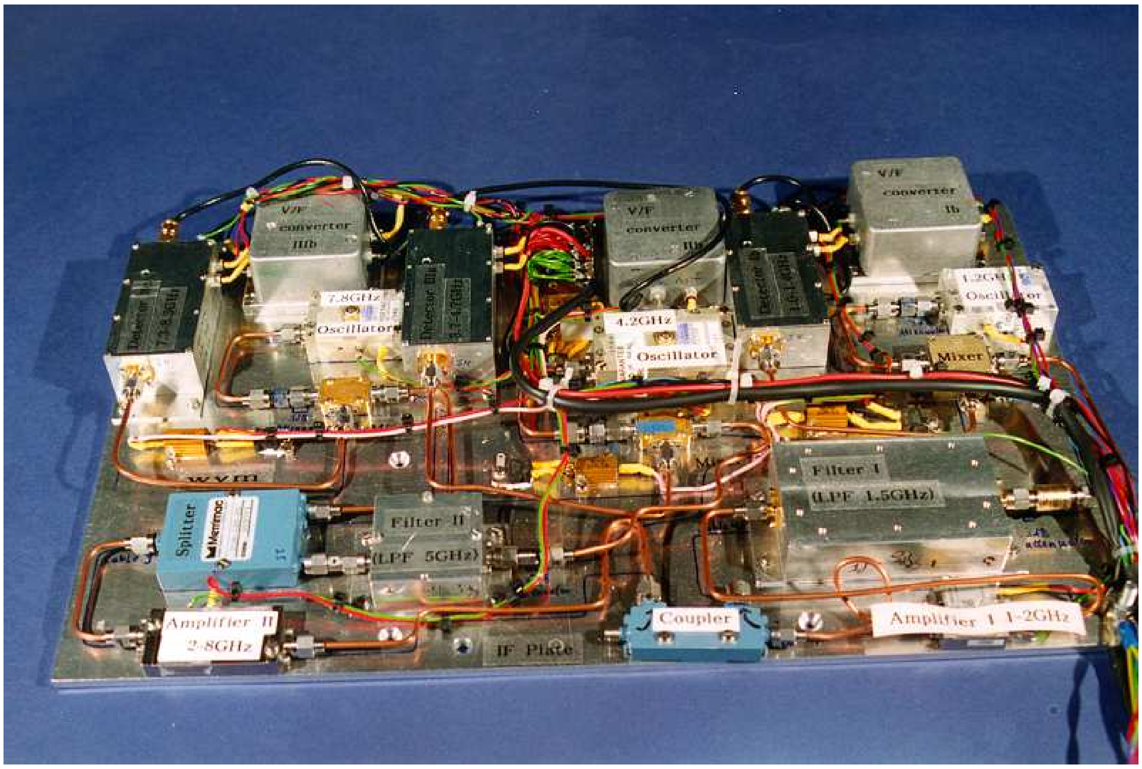


Figure 5.5. Picture of IF electronics. The signal from the RF plate enters the coupler in the middle of the foreground and is then split and down converted. The three crystal oscillators and mixers for the second down conversion are in the middle of the plate. The three elongated, dark boxes in the background are the detectors for the three channels. On the right of each detector is a square box containing the voltage to frequency converters.

within each frequency band, because we do not want certain frequencies within a band to be preferentially amplified (also see §5.2.2).

5.2.2 IF plate

Coupler and Splitter: The coupler and power splitter are used to divide the sky signal into three parts, each covering the whole IF range. From each part a different frequency band is filtered out for further processing.

An alternative to just dividing the signal in power would be to use a triplexer, which selects the signal according to its frequency. The 1 - 8 GHz IF signal would, for example,

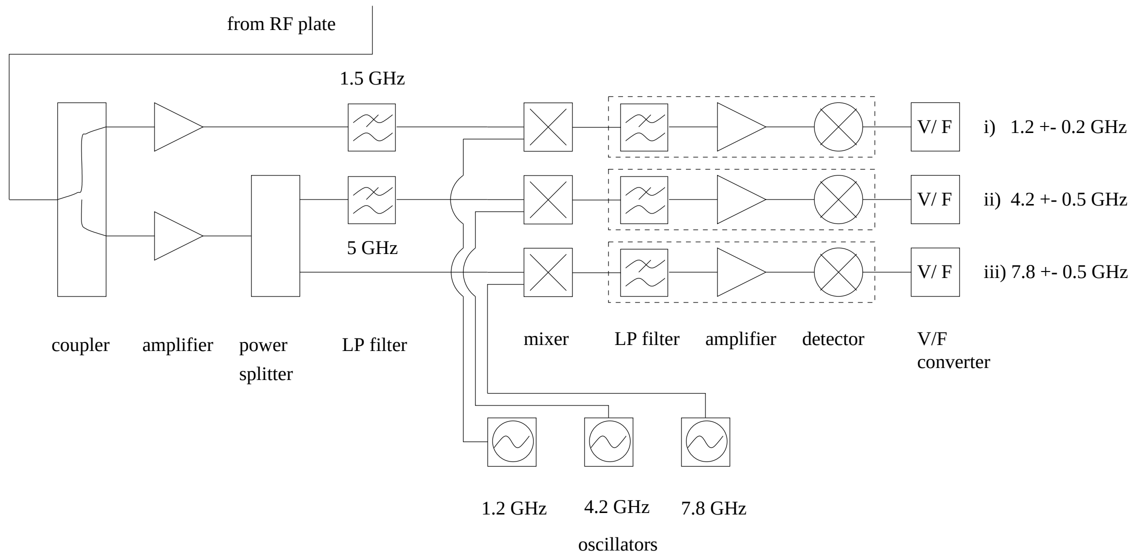


Figure 5.6. Intermediate frequency (IF) electronics.

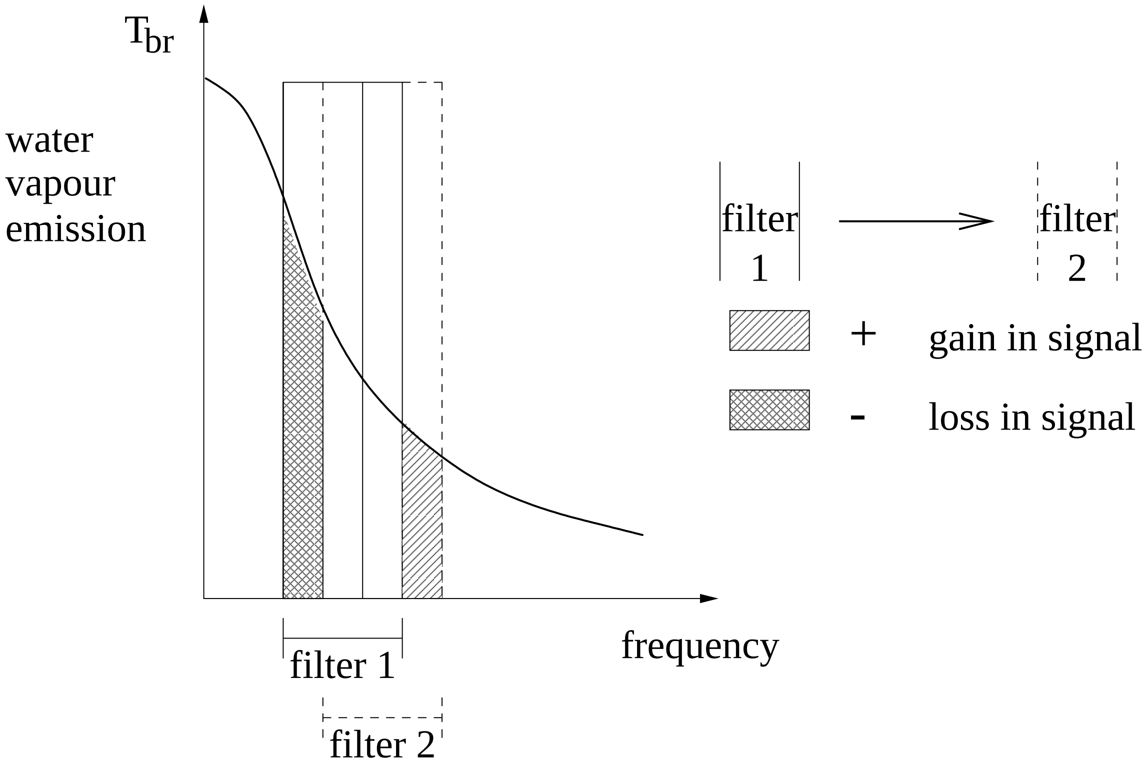


Figure 5.7. A small shift in the centre frequency of a filter results in a big change in the measured brightness temperature because the water line is very steep.

be split into a channel covering 1.0 - 1.4 GHz, another covering 3.7 - 4.7 GHz and a third of 7.3 - 8.3 GHz. By using a triplexer no signal is lost (as it is in the power splitter and filter set-up). However, suitable triplexers are very expensive. More importantly, in order to measure a optical path difference of $30\mu m$ the frequency bands of the triplexer need to be known to better than 5 MHz. At least the triplexers used in each of the two radiometer need to be the same within 5 MHz (see Fig. 5.7). None of the commercially available triplexers had this accuracy.

Second down conversion: Instead of passing the three IF channels through three different filters, which pass the desired frequency bands, the IF signal is down-converted a second time: The signal in the first channel is mixed with a coaxial resonator oscillator signal at 1.2 GHz. The difference frequency is then passed through a 200 MHz wide low pass filter. Since the mixer is double side band the first channel covers a frequency range of

1.0 to 1.4 GHz above and below the centre of the water line at 183.3 GHz. The second channel is mixed with a dielectric resonator oscillator at 4.2 GHz, the third with one at 7.8 GHz, and each is passed through a 500 MHz low pass filter.
The second down conversion has two big advantages:

1. The positions of the frequency bands are well defined, as it is easy to buy oscillators with an accuracy of $\Delta\nu/\nu \leq 10^{-5}$, i.e. the frequency is defined to better than 0.1 MHz.

2. No knowledge of the shape of the filter bandpass is required.

   To explain this I would like to consider first a system without a second down conversion, where the filter would sit on the wing of the steep water line (Fig. 5.7) and the intensity passed by the filter will depend on the shape of the bandpass: A filter with preferred transmission at low frequencies will pass most of the high intensity line emission but little of the low intensity emission, whereas a filter preferentially passing high frequencies will filter out some of the high intensity emission. As a result the latter filter will pass less intensity than the first. However, when observing a calibrator, which is usually a black body and effectively has constant intensity over the small bandwidth of the filter, the detected intensity will be independent of the shape of the bandpass, but will only depend on the integral of passband over frequency. Therefore the two filters discussed above would give the same output intensity when looking at a calibrator, but different intensities when looking at the line.

   However, if there is a second down conversion, the water line is folded at the oscillator frequency (Fig. 5.8). To first approximation the double sideband signal is flat, just as that of the calibrator, and the shape of the filter becomes irrelevant. Only the integrated bandpass is of interest, which can be easily determined using a black body as a calibrator. [2]

Filtering out harmonics of crystal oscillators: For the second down conversion a set of three oscillators is used. In order to avoid any harmonics (i.e. whole multiples of the oscillator frequency) appearing as spurious signals in the measurements, the frequency

[2]Using the same argument the gain of amplifiers after the second down conversion is of no importance, however the gain of amplifiers before the mixing needs be flat.

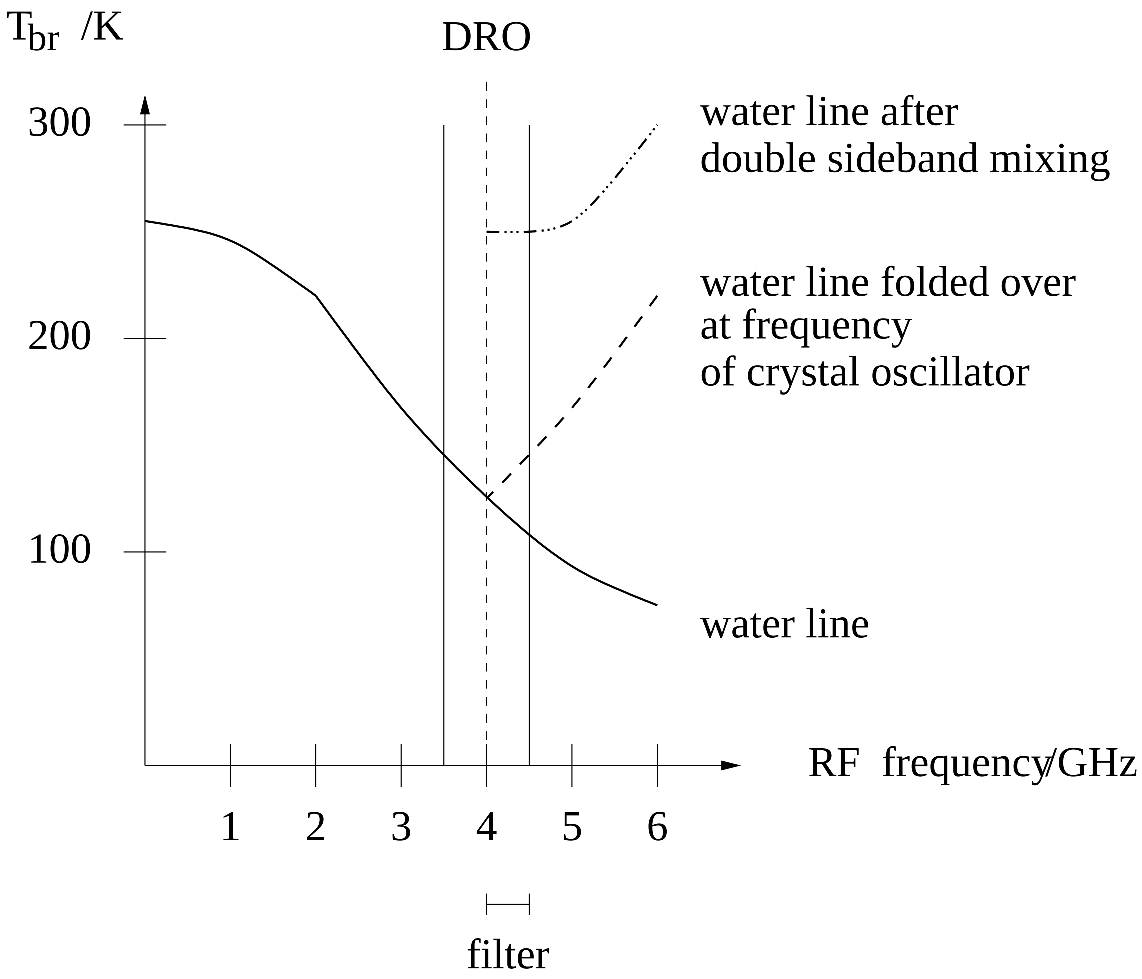


Figure 5.8. Double sideband mixing of the water line with a dielectric resonator oscillator (DRO) signal at 4 GHz. The water vapour line is folded at 4 GHz and the original line and the folded line are added. The sum of the signals will be flat, as long as the original line can be approximated by a tangent and the upper and lower sideband response are the same, i.e. the sum is not weighted.

bands are chosen so that any harmonics lie outside them. E.g. the sixth harmonic of 1.2 GHz falls outside the frequency range of the third channel of 7.3 to 8.3 GHz, and the seventh harmonic of 1.2 GHz as well as the second of 4.2 GHz are above 8.3 GHz. In addition a 1.5 GHz low pass filter was added between the second amplifier and the mixer to prevent any harmonics of the 1.2 GHz oscillator to travel backwards through the amplifiers and the coupler into the second or third channel. Similarly a 5 GHz low pass filter to stop harmonics of the 4.2 GHz oscillator was used (see Fig. 5.6).

Detector: The signal in each channel is measured by a Schottky diode detector (level detector). The voltage across the diode is determined by the input current, consisting of a constant bias current and the signal. This voltage is amplified and used to determine the intensity of the signal input. However, temperature changes in the diode can change the voltage across the diode with the current being constant. In order to eliminate temperature effects the voltage across a second diode, whose only input is the bias current, is subtracted from the voltage of the detector diode. Temperate changes will cause both diode outputs to change by the same amount. The voltage difference should therefore only depend on the signal input and the output signal is temperature independent.
It was observed that the detector output is not linearly proportional to the input brightness temperature. Instead the detected voltage V is proportional to the input brightness temperature T_{bri} to the power of β, $V(T_{bri}) = aT_{bri}^{\beta}$. β takes values between 0.8 and 1. However, since we are interested in brightness temperature changes in a very small regime $\leq 10\ K$ compared to the input brightness temperature T_0 of $\sim 2500K$ $V(T_{bri})$ can be approximated by its tangent at T_0. This approximation will lead to an error in the difference temperature of less than 1%.

Voltage to frequency converters and signal transmission: As it is difficult to measure the detector output voltage accurately to 10^{-4} with a voltmeter, the voltage is converted into a frequency (with a constant conversion factor of about 180 kHz/1 V). This frequency is sent to a computer, where it is counted and digitised. In order to avoid interference each frequency is transmitted differentially along a shielded pair of twisted wires: each wire is at about 2.5 V. The frequency is transmitted by a rise in voltage to 3.5 V on one wire and simultaneously decrease to 1.5 V on the other wire. Since the voltage of the two wires changes at the same time, but in opposite sense, the far field of the wire pair will be the same as that of a wire constantly at 2.5 V. A constant field does not emit radiation, hence the transmitted signal should not cause any interference.

Optoisolators: Each pair of twisted wires is connected to an optoisolator at the computer end, to decouple the water vapour monitor electrically from the computer. The current induced by the changing voltage lights a small light emitting diode (LED) at the frequency of the V/F converter output. A detector sees the light pulses and sends a signal to the counter board in the computer. The optoisolator faithfully reproduces the signal frequency but at a constant output amplitude optimised for the counter board. Therefore voltage spikes won't harm the computer. In addition the radiometer and the computer can be connected to different ground voltages.

5.3 Temperature regulation

Many of the electronic components are quite sensitive to temperature changes. If the temperature of our RF plate rises by 1 Kelvin, the voltage output of the detectors decreases by 0.5 %. For the IF plate a temperature rise of 1 K, reduces the output by 0.5 to 1.0 % depending on the frequency channel. In order to measure brightness temperature fluctuations of 0.5 K in a system temperature of about 2500 K, we need to measure fluctuations of the order of 0.05 % of the output signal. Besides frequently calibrating the radiometers (once a second), all the electronic components are temperature regulated. Temperature sensors and heaters are mounted on the aluminum plates and two fans circulate the air in the electronic box.

Heaters: Four heaters (transistors and resistors) are located on each of the aluminum plates, on which the electronic components are mounted. Sensors in the centre of the plates monitor the temperature and regulate the heater current. The RF plate is heated to $\sim 40^\circ$ C, the IF to $\sim 38^\circ$ C. In both cases this temperature lies slightly above the temperature to which the electronics warms up without heaters.

Fans: Each water vapour monitor box contains two fans, one for circulating the air inside the box, the other to extract hot air.
The first fan runs constantly ensuring that the air is well mixed. Regions of hotter air could otherwise form and would shift in an uncontrolled way when the telescope and

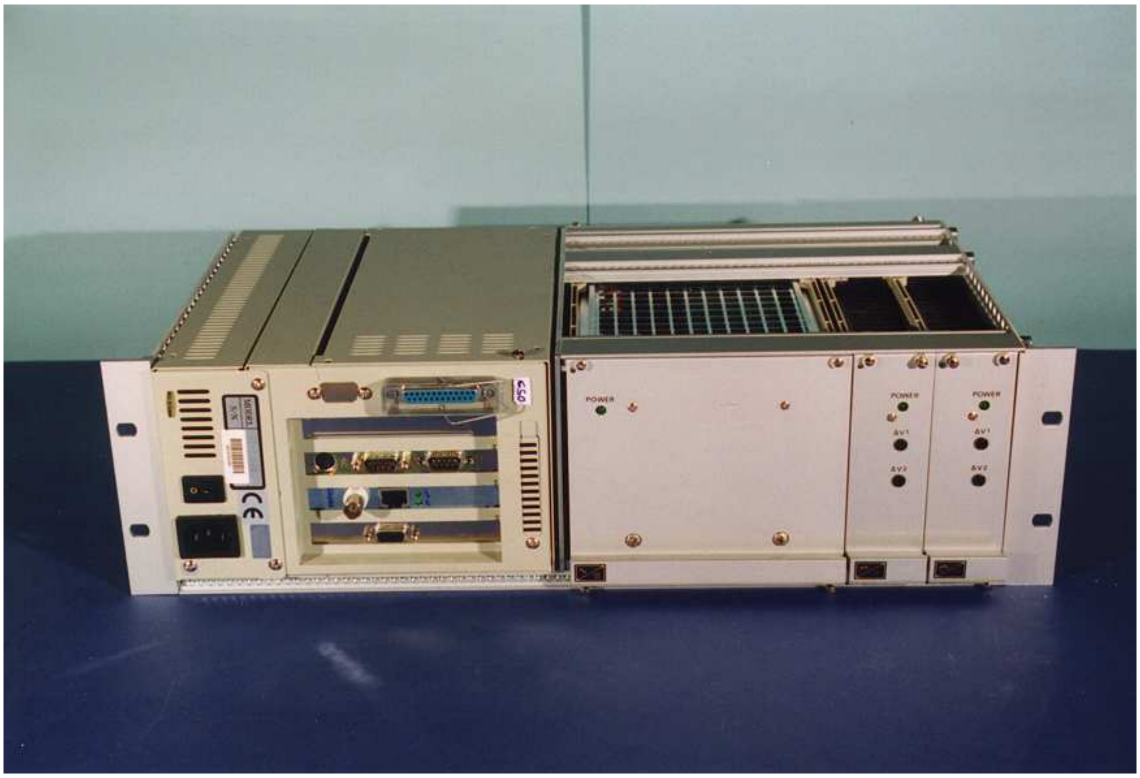


Figure 5.9. Industrial PC (left) and power unit (right) for WVMs in a two units high ($\sim$ 90 mm) 19 inch rack. The PC drives drives the flip mirror and digitises the data. The power unit supplies +15 V and -15V for the IF and RF electronic components and 26 V for the heaters.

therefore the WVM is tipped.

The extractor fan is coupled to the heating circuit of the RF plate since it is more likely to overheat than the IF plate. The voltage driving the fan is inversely proportional to that of the heaters. The extractor fan is turned off when the heaters are running at full power, but it slowly turns on as the plate gets close to its nominal temperature.

5.4 Computers

Each water vapour monitor has an industrial PC, which has three tasks (Fig. 5.10):

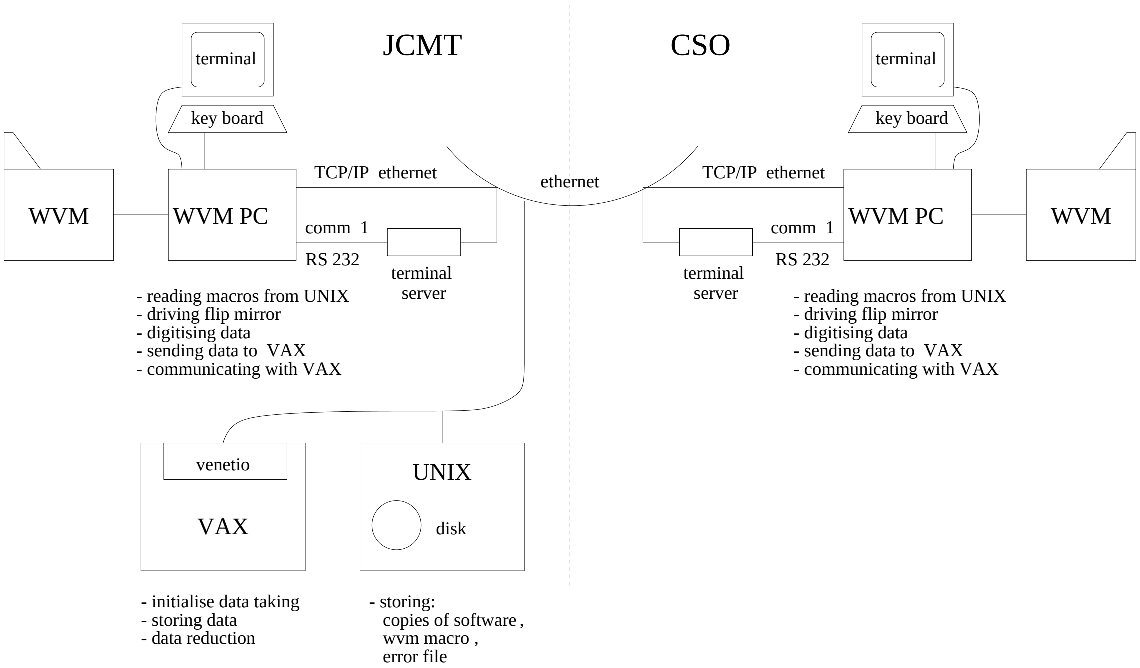


Figure 5.10. Schematic of the computers involved in running the WVMs. The PC drives the flip mirror and digitises the data. Since the memory and computing power (CPU) of the PC are fairly small the data is stored and reduced on the VAX at the JCMT. The VAX is also used to start and synchronise the data-taking of both WVMs. The macro which defines the calibration cycle is stored on the UNIX, where both PCs can access it.

1. The PC sends a logical signal to the motor control which rotates the flip mirror to different positions, so that the water vapour monitor can observe either the sky or one of the calibration loads (see §4.4).

2. Secondly, the PC digitises the signals from the water vapour monitors. The four V/F converters of the WVM send four frequencies representing the detector output of the three IF channels and a temperature measurement of one of the temperature sensors to a counter board in the PC. The frequencies are counted, digitised and stored. As the flip mirror moves, the detectors for the three IF channels will measure emission from the hot load, the warm load and twice from the sky. At the same time the fourth V/F will convert the temperature measurements of the sensor in the hot load, in the warm load, on the IF plate and on the RF plate. Therefore for each calibration cycle an array of 4×4 numbers is obtained.

3. Thirdly, the WVM PC needs to communicate with the VAX and the UNIX computer at the JCMT. The PC is connected to the Ethernet directly as well as via an RS232 link and a terminal server.

The macro which defines the calibration cycle is stored on the UNIX computer at the JCMT. The length of the calibration cycle can be changed easily on this computer. Both PCs are connected to the UNIX computer via the RS232 link, so that they will drive the mirror control with the same macro. In case the connection can't be established, each PC has a backup version of the macro stored in its own memory chips.

The VAX computer at the JCMT is connected to both WVMs via the Ethernet. Data-taking programs are run from the VAX. First a connection from the VAX to both PCs is established and they are synchronised. Thereafter every second the VAX reads the data array of 16 numbers from each PC. These 16 frequencies and the time are stored in a binary file, which is later used in data reduction (see chapter 6).

5.5 Performance of electronics

5.5.1 System temperature

The input power of the LO can be adjusted to give the best performance of the mixer. However, the optimum LO power is different for different IF frequencies. The LO power was adjusted to give a compromise between all channels. The current set-up has system temperatures between 2000 and 3000 K. As discussed above the thermal noise in a one second integration time from these system temperatures is just lower than the accuracy required.

5.5.2 Frequency stability of LO

The PLL keeps the Gunn at 91.665 GHz with an accuracy much better than 20 kHz. Fig. 5.11 is a sketch of the locked LO when monitored with a spectrum analyser. The width of the peak is determined by the resolution of the spectrum analyser. Since the peak did not move on the screen of the spectrum analyser the stability of the PLL must be much better than 20 kHz. Within the loop band width the phase noise comes from the 145 MHz crystal oscillator, outside it is determined by the Gunn. The "shoulders" indicate where the transition takes place.

Since the Gunn is locked with respect to the DRO and the crystal oscillator any shift in their frequency will also shift the LO frequency. Between 0 and 70 C the DRO has an accuracy of 10^{-5}, i.e. 102 kHz. This could result in a phase shift of 9×102 kHz in the LO frequency. The crystal oscillator also has an accuracy of $\sim 10^{-5}$, i.e. 1.45 kHz, resulting in a shift in LO frequency of only 1.45 kHz. The total accuracy in the LO frequency is therefore better than 1 MHz.

A change in LO frequency will also change the filter frequencies, but to first approximation the detected power will be independent of the LO frequency, see Fig. 5.12. Consider measuring the signal in a narrow frequency band located e.g. 1 GHz above and below twice the LO frequency. If the LO moves up in frequency, the signal in the lower frequency band will rise as it gets closer to the centre of the water line, but the signal in the upper band will decrease, see Fig. 5.12. If the water line can be approximated by a straight line over the width of the filter, the measured signal, which is the sum of these two signals, will stay constant.

For the correct line shape a LO frequency shift of 100 MHz will change the measured signal of a channel 0.8 to 1.8 GHz away from the line centre by 3.3×10^{-5}, i.e. by 0.007 K in

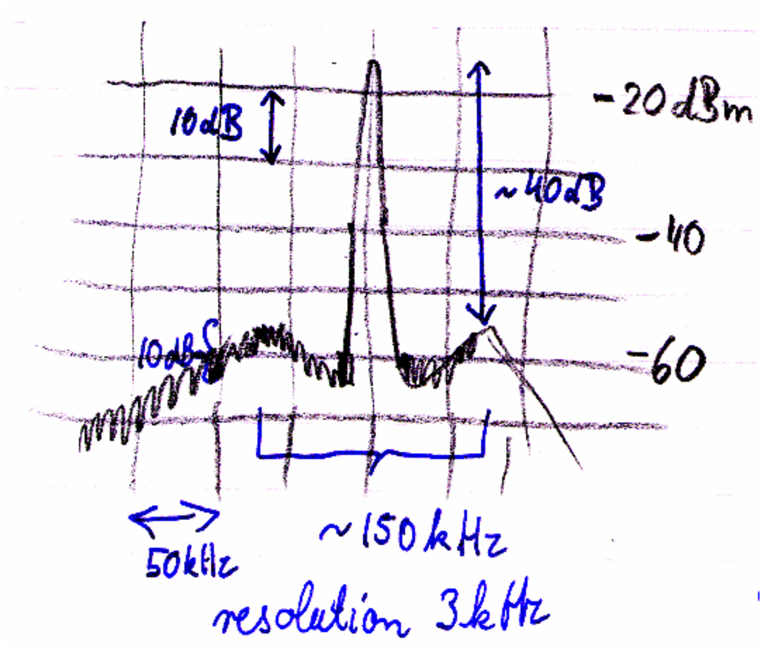


Figure 5.11. Sketch of locked LO taken from the notebook. The strength of the oscillator is plotted against frequency. Ideally, the LO should only oscillate at one frequency and the plot should show a delta function. The sketch shows a strong peak of the LO; Its width of 20 kHz is due to the limited resolution of the spectrum analyser. There is also some power at other frequencies. The loop bandwidth is $\sim$ 75 kHz, half the width of the "shoulders". Inside the loop bandwidth the 145 MHz oscillator, to which the LO is locked, determines signal stability and the phase noise. Outside the the loop bandwidth the phase noise is due to the LO.

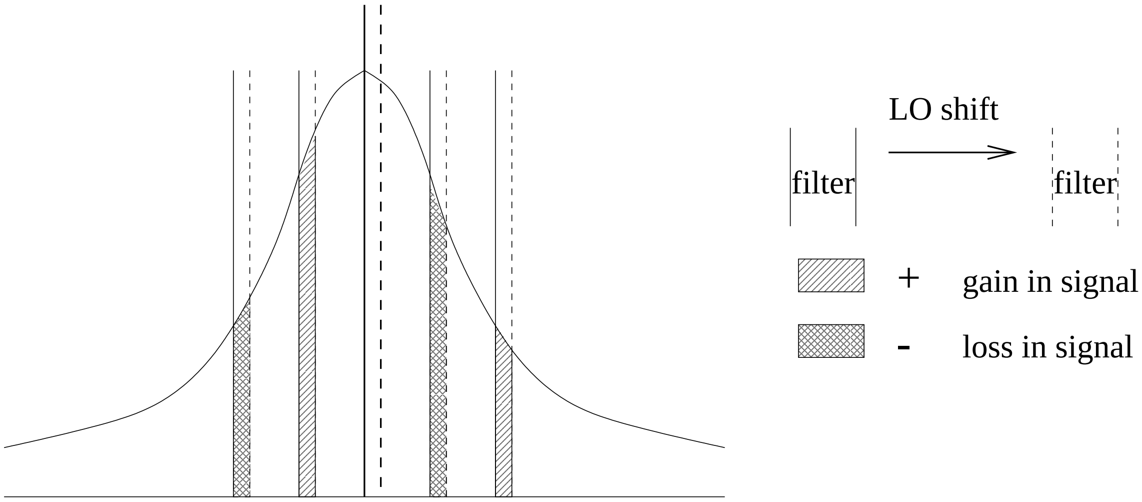


Figure 5.12. Effect of shift in LO frequency on detected power of water vapour line. If the LO frequency shifts to a higher frequency both filters shift up as well. There will be more power in the left filter, but less in the right filter. To first approximation the sum of the detected power in both filters will be constant.

$\sim$200 K. This change is very small and not significant.

In summary, the expected shift of LO frequency of 1 MHz will not change the detected power significantly. The error is so small that it can be neglected.

5.5.3 Frequency response

Ideally all amplifiers used should have a flat gain curve, i. e. their amplification should be independent of input frequency. Similarly the mixers should have a constant conversion factor and the detector should be equally sensitive to signals at different frequencies. To measure the combined response of IF electronics of the WVM a signal of constant power but at varying frequency was used as an input into the IF circuit (Fig. 5.6). In Fig. 5.13 detected power is plotted versus frequency for channel II. The response is not perfectly flat. No power is detected at 4.2 GHz the frequency of the crystal oscillator because after the second down conversion 4.2 GHz becomes $\sim$ 0 GHz which cannot be detected, because there is a capacitor at the detector input. According to the discussion about down conversion in §5.2.2 the passband of filters and the gain of amplifiers applied after the second down conversion is not

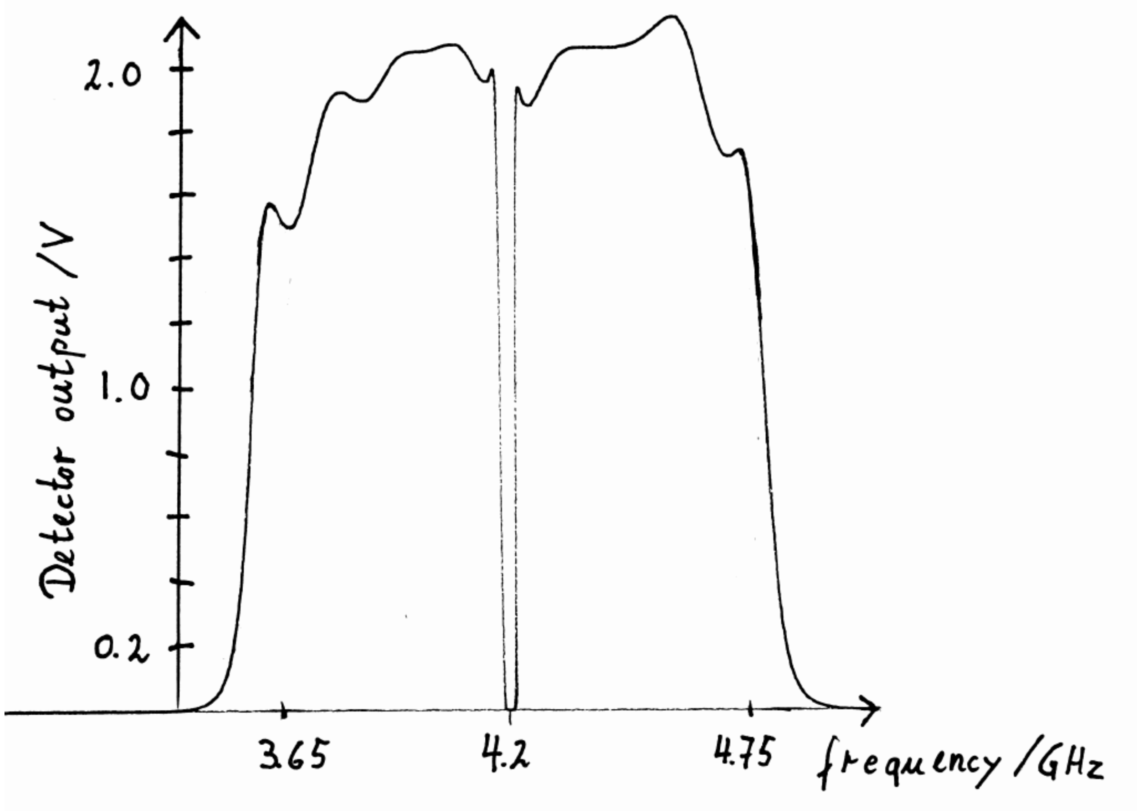


Figure 5.13. Plot of detected power versus input frequency. A signal of constant power but with a frequency varying between 3.4 and 4.9 GHz was used as an input to the IF channel (Fig. 5.6). The signal was amplified, mixed with the 4.2 GHz crystal oscillator signal, further amplified and detected. The plot gives an indication of the gain of the amplifiers,passband of the filter and response of the detector.

of importance. Any amplification and filtering applied after the second down conversion will be symmetric about 4.2 GHz in Fig. 5.13. Therefore only any asymmetry will contribute to an error.

To estimate an upper limit for the error in deduced brightness temperature we assumed that the response between 3.5 and 4.2 GHz is 10% lower than the response between 4.2 and 4.9 GHz, this results in an error of 0.5 K. The difference in response is expected to be nearly constant and vary by less than 10% or 0.05 K over one hour. When measuring the brightness temperature fluctuations we take differences therefore only the 0.05 K error will contribute. This error is much smaller than the accuracy of 0.5 K we are aiming for.

5.5.4 Gain fluctuations

Given an input signal at constant power the output voltage will fluctuate slightly due to changes in the overall amplification and attenuation of the input signal. The gain of the amplifiers is especially sensitive to temperature changes of the amplifier itself. Therefore all electronic components are mounted on temperature regulated plates. The gain fluctuations at the JCMT WVM are very small with a root mean square variation of 0.1 K comparable to thermal noise in a 3 second integration, see Fig. 5.14. The CSO WVM shows bigger gain variations of about 0.4 K rms over about a 6 minutes period. This is probably due to the CSO WVM being mounted outside and therefore being much more exposed to the weather. Every second the water vapour monitors are calibrated against loads at constant temperature, so that the amplification and therefore the gain variation of the system is known and can be taken into account during data analysis.

5.5.5 Temperature stability of plates

The temperature stability of the plates, on which all the electronics are mounted, is very high with an rms fluctuation of 4 mK. This ensures a constant temperature of the electronics and therefore reduces the gain variations.

5.5.6 Spikes in data

Most of the spikes have been eliminated by only using the good counters in the counter board and by sending a very clean signal from the voltage to frequency converters to the counter

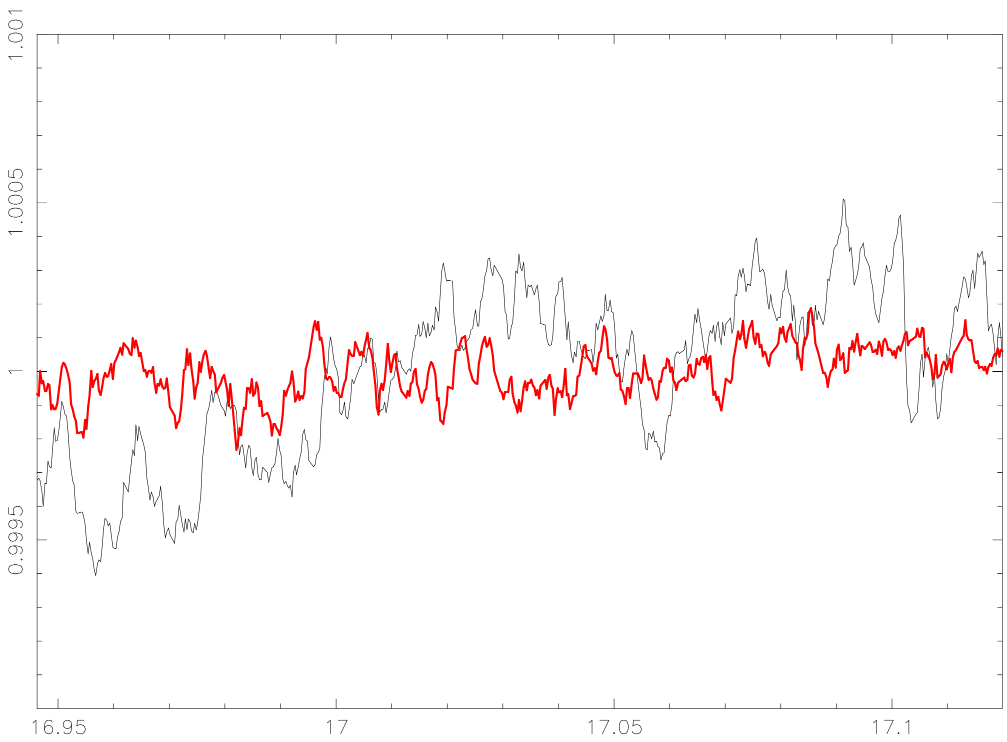


Figure 5.14. Gain variations. The thick (red) line shows the gain variations of the JCMT WVM against time in hours and the thin line the gain variations of the CSO WVM. The data were hanning smoothed over 10 s. Full scale of the y-axis corresponds to about 5 K.

board in the computer. Unfortunately, however, there are still some spikes in the data.

There seem to be at least two different types of spikes: The CSO data shows spikes in channel II and III, but only when integrating on the sky and not during observations of the loads. The spikes always occur at the same time in both channels, but at irregular intervals of 40 to 100 sec. The JCMT WVM has much fewer spikes, which occur in all channels. (See also §6.3.2 under spikes.)

The origins of these very sudden changes in detected power are unknown. Possible explanations are:

- Crosstalk of frequency outputs:
  All the frequency outputs of each WVM are sent to the data taking PC along one cable. Interference between the different frequency signals in this cable could cause spikes.

- Interference by astronomical receivers:
  Since the spikes in the CSO WVM only occur in the sky channel this might indicate that these spikes are caused by a signal from the telescope such as an interfering signal from the astronomical receivers.

- PC:
  Since the PC has very limited computing power, it might miscount the frequencies, when it is interrupted, for example, by the VAX sending commands.

It is desirable to eliminate the origins of the spikes as far as possible. Large spikes, however, won't degenerate the quality of the data very much as they are easily removed during the data reduction.

5.5.7 Summary

The water vapour monitors could still be improved slightly by reducing the gain variations of the CSO WVM and by eliminating the origin of the spikes. Overall, however, the electronics of the water vapour monitor performs very well and the radiometers fulfill the requirements for phase correction.

Chapter 6

Performance of water vapour monitors, data analysis and phase correction

This chapter describes the performance of the water vapour monitors (WVM) and the data taken with them. The section §6.1 gives a brief account of the observing trips on which we have taken the radiometers. It summarises the aims for each trip, the performance of the monitors during each trip, the problems encountered and the subsequent improvements made. Section §6.2 outlines a typical data analysis of the water vapour data used for phase correction. The next section §6.3 looks at the uncertainties in the data and estimates the errors of the corrected phase. In section §6.4 the most recent data set taken is presented and analysed. The phase predicted by the water vapour monitors is compared to the interferometer phase and the residual phase is computed. The last section §6.5 briefly summarises the performance of the WVMs and suggests some future experiments with the 183 GHz radiometers.

6.1 Description of observing runs

Work related to this project was carried out during four observing runs at the CSO-JCMT interferometer, in January 1996, November 1996, May 1997 and November 1997. Table 6.1 summarises the four observing runs with respect to the water vapour monitors.

6.1.1 January 1996

Goals

- Firstly, the best position for installing the water vapour monitors on the telescopes had to be found, exact measurements had to be taken and our plans had to be discussed with the maintenance crews of the telescopes.
- We had just started building the water vapour monitors. We wanted to take a simplified, preliminary version of only one radiometer to Hawaii to test its performance at the actual site. In particular, we wanted to measure system temperatures, determine the temperature dependence of the system and measure the system's stability over long periods of time.
- Thirdly, we wanted to get measurements of the water vapour and compare them to the CSO tau meter, which measures the atmospheric opacity at 225 GHz.

Preliminary water vapour monitor

At this stage the water vapour monitor consisted of an aluminium board with a temporary horn, the mixer and the IF circuit. This water vapour monitor had only 2 channels located 1.2 GHz (channel I) and 4.2 GHz away from the line centre (channel II). There was neither a phase lock loop, nor a temperature control, nor calibration loads. The electronics plate was installed in the receiver cabin of the JCMT above the tertiary mirror. An ellipsoidal mirror with a focal length of 77 mm directed the beam from the sky into the horn.

Tests performed and results

- The best positions to install the final water vapour monitors were found and the maintenance crew agreed to provide brackets to hold the instrument and mirrors. They also arranged connections to the computer network.

		Jan 96	Nov 96	May 97	Nov 97
		IF board	2 WVMs	2 WVMs	2 WVM
Latest improvements				added fans, temp. reg.	new motor contr., insulation
Weather		good	fair	bad	bad
Preparation:					
Tune attenuator		√	√	√	√
Aligning mirrors		√	√	√	√
Tune flip mir. contr.		—	√	√	√
T_{sys}/K JCMT	1	2200	2300	1800	2200
	2	2300	2300	2200	2800
	3	—	2200	2200	2800
T_{sys}/K CSO	1	—	2200	2900	2200
	2	—	2800	3800	2800
	3	—	2400	4500	2800
Temperature	1	-50mV/K			
dependence	2	-60mV/K			
Stability	1	0.5%/4h			tipping WVM:
	2	0.3%/4h			all channels:
					1% gain change
Sky dip	1	0.9-1mm	consistent with		
	2	1-1.2mm	τ-meter to		
τ-meter		0.7-1.2mm	better than 30%		
Coupling JCMT					95%
CSO					75%
Data obtained		only tests	good corr. with 3C273	—	good corr. with MWC349
Problems		—	spikes, overheating JCMT WVM	V/F break, flip mir. pos. bad weather	some spikes, gain varied, bad weather

Table 6.1. Summary of observing runs with WVMs.

- The system temperature of the water vapour monitor was 2200 K in channel I and 2300 K in channel II, compatible with our aim of a system temperature below 2500 K.

- Gain fluctuations of the water vapour monitor lay below 0.5 % over a 4 hour period. Most of these fluctuations were due to temperature changes in the electronics. The IF plate has a temperature dependence of -50mV/K for channel I and -60mV/K for channel II. Since the detector has an output of 8 V in channel I and 7 V in channel II, when looking at an ambient load, the temperature dependence is about 0.6 %/K. However, in order to detect fluctuations of 0.5 K with a system temperature of 2500 K we need an accuracy of 0.02%. Therefore, temperature regulation of the electronics, as well as calibration loads for monitoring the gain, were added to the subsequent versions of the water vapour monitors.

- A sky dip with the water vapour monitors suggested that there was between 0.9 and 1.2 mm of precipitable water vapour. This agreed well with the CSO tau meter, which measured 0.7 mm to 1.2 mm over the same time period. (The sky dip was taken during day time when the amount of water vapour fluctuated a lot.)

In summary the first test at the interferometer showed that the water vapour monitor is capable of detecting the water vapour as accurately as the CSO tau meter, but that its accuracy and stability needed to be improved in order to use it for phase corrections of the interferometer.

6.1.2 November 1996

Goals

- The performance of the first version of the two water vapour monitors was to be tested.

- It was to be investigated what improvements could be made.

- The response of the two instruments was to be compared, when they are observing the same atmosphere.

- We wanted to take the first measurements with the water vapour monitors and the interferometer in parallel to look for correlations between them.

Water vapour monitors

Both water vapour monitors had been built for the interferometry run in November 1996. Each had the full IF system with 3 channels, the temperature regulation for the electronics and and the phase lock loop as described in chapter 5, as well as the complete mirror system (see chapter 4).

Tests and results

- The system temperatures of both monitors lay between 2000 and 2500 K, a sky dip predicted amounts of precipitable water vapour comparable to the CSO tau meter.

- During one of the tests a layer of moist atmosphere passed over the monitors and was clearly visible in both monitors at the same time, as one would expect.

- In spite of the bad weather during the observing run we obtained some data, which proved that the radiometers work very well: Figure 6.1 shows that the *observed* phase of the bright quasar 3C273 agrees well with the *predicted* phase of the WVMs. When the quasar observations are corrected by the WVMs the root mean square (rms) of the phase is reduced from 55° to only 23°. (See Section §6.2 for a description of the data analysis.)

There were, however, still a few problems with the water vapour monitors:

- The JCMT monitor overheated in the thin atmosphere on Mauna Kea, because it is installed in the relatively warm receiver cabin.

- In addition the data exhibited many spikes. By spikes we mean single data points whose value differ greatly from neighbouring data points. This difference is so large that it cannot represent thermal noise or a change in brightness, but must be an artifact of the data taking process.

In summary the tests in November 96 at the interferometer proved that the water vapour monitors give a good estimate on the phase fluctuations due to water vapour and can considerably improve the quality of the astronomical data.

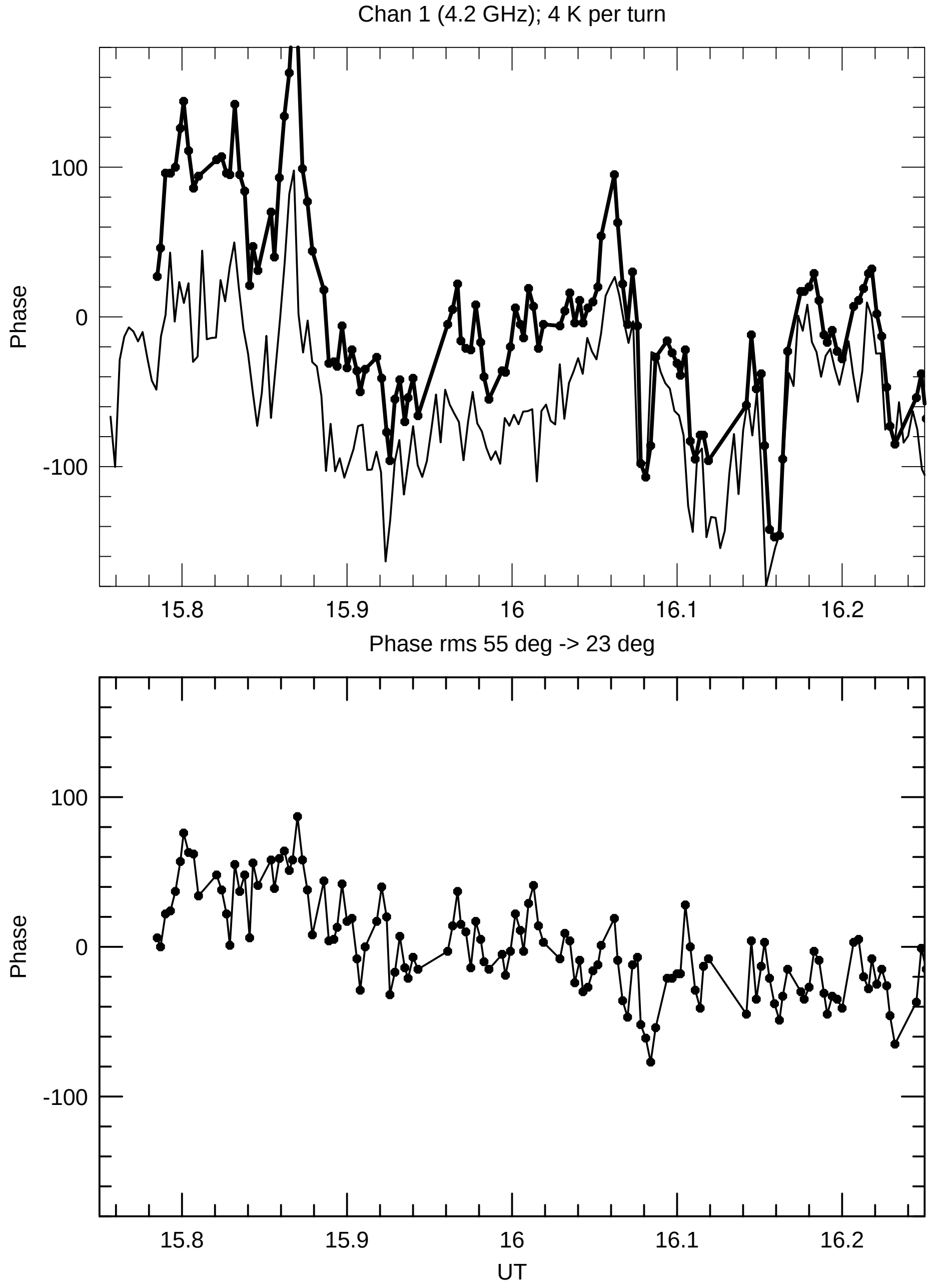


Figure 6.1. In the upper plot the difference in the brightness temperature observed in channel II (thick line) is plotted in comparison to the interferometer phase (thin line). The temperature difference was divided by the conversion factor of 4 K/360° to match the interferometer phase. The lower plot shows the radiometer phase minus the interferometer phase. After phase correction the rms phase is reduced from 55° to 23°.

6.1.3 May 1997

Goals

- The improved water vapour monitors were to be tested.
- The number of spikes observed in the previous run were to be reduced.
- We wanted to take more data with the water vapour monitors and the interferometer running in parallel and confirm the good correlation between them and correct the interferometer phase.
- The data reduction software for the monitors had to be written and installed.

Tests and results

- Two fans had been added to each water vapour monitor, which successfully prevented the electronics from overheating.
- Similar to the previous run the sky dips with the two monitors running in parallel confirmed the measurement of the CSO tau meters.
- Measurements at different elevations taken with the WVM mounted on the CSO suggested that between 15% and 26% of the beam power was lost and did not couple to the sky. In fact there is a small absorber in the middle of the CSO secondary mirror. From the size of the absorber a theoretical signal loss of 22% was predicted, which confirmed the measurements. (§6.2.2 explains how the coupling of the beam is calculated.)
- The software running the industrial PCs was improved, which reduced the number of spikes in the data considerably. However, there were still some spikes left.
- Some software was written to graphically display the measurements of the water vapour monitors, to calculate system temperatures and gains and to average data.

There were, however, two major problems with the water vapour monitors:

- **Voltage to frequency converters**
  Three voltage to frequency converters broke. The causes were trivial, e.g. a dry solder

joint and subsequent burning of a transistor. As we had two spare units we exchanged them for the broken voltage to frequency converters. One of the converters was repaired. Although these tasks are relatively simple, it proved infeasible to repair the CSO WVM without taking it off the telescope as it is difficult to work while climbing in the telescope structure. Taking the WVM off the telescope, however, took a considerable amount of time, mainly because the mirrors needed readjusting after remounting the WVM.

- **Motor control of flip mirror**
  The second big problem was that the CSO system temperature appeared to rise to very high values. Initial tests confirmed the expected system temperatures between 1700 K and 2500 K. However, once installed at the telescope the indicated system temperature rose up to 4500 K at the CSO water vapour monitor. The reason for this was that the flip mirror did not move to its nominal position. Instead of seeing the hot load at 100°C the horn saw part of the hot load and part of the box at ambient temperature, therefore detecting something at a temperature much lower than the expected 100°C. We discovered that the mirror position was a function of the temperature of the rotary solenoid, which sets the mirror to the different positions. Therefore, the solenoid was surrounded by thermal insulation and the mirror positions where readjusted after the system and the atmosphere had reached stable temperatures. However, the position of the mirror at the CSO still seemed to shift slightly.

In the rather small data set brought back to Cambridge I could find no correlation between the water vapour monitor phase and the interferometer phase. This might have had such a simple reason as, for example, that the clock of the computer which takes the SBI data might not have been synchronous with the clock of the VAX computer, which stores the WVM data.

In summary the improvements on the heating system worked well and reduction software was further developed, but, disappointingly, we did not obtain any more WVM data that showed good correlation with the interferometry phase. The main problem was the motor control of the flip mirror, which needed to be improved.

6.1.4 November 1997

Goals

- The motor control, which sets the flip mirror to the different positions, had to be improved.
- We wanted to confirm that the local oscillator of the WVMs does not interfere with the astronomical receivers.
- The stability of the radiometers when the telescope is changing elevation was to be tested.
- We planned to retake data with the water vapour monitors and the interferometer running in parallel and try phase correction.

Tests and results

- An integrator circuit was added to the motor control in order to measure the distance moved from one motor position to the next. The circuit ensures that the motor always moves by the same distance. The motor control now works well.
- The WVM worked well with system temperatures between 2200 K and 3000 K. The measured noise was comparable to the thermal noise and sky dips predicted similar amounts of water vapour as the CSO τ-meter.
- There were still large spikes in the CSO data of channel III. It turned out that one of the counters or the readout of this counter produced the spikes, as they appeared even when the PC clock was fed into the counter. Therefore channel III was connected to a different counter and the large spikes disappeared. There were, however, still some small spikes in some of the data.
- Tuning the astronomical receivers to a harmonic of the 183 GHz line showed no signal and therefore proved that the local oscillators of the WVMs are not leaking into the astronomical receivers, i.e. that the astronomical receivers do not pick up spurious signals from the WVMs.

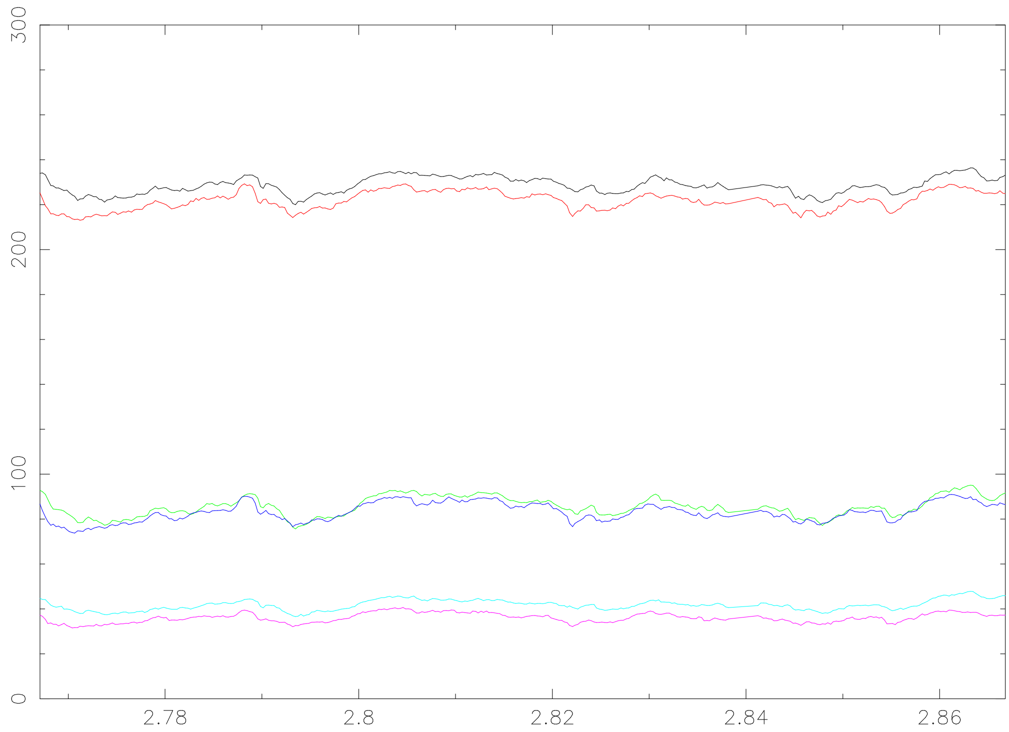


Figure 6.2. JCMT and CSO WVM observing the same atmosphere. From top to bottom the curves represent: JCMT channel I, CSO channel I, JCMT channel II, CSO channel II, JCMT channel III, CSO channel III.

- The WVM were tipped to different angles. The temperature sensor readout of the calibration loads varied by less than 0.025% suggesting that the physical temperature of the loads changes by less than 0.025 K, which is negligible. The gain of the system, however, changed by up to 1%, which stresses the necessity of frequent gain calibrations.

- Some data was taken with the WVMs running next to each other (Fig. 6.2). The graphs show similar noise. Most of the features are seen by both WVMs simultaneously. However, the monitors' signals do not agree perfectly.

  The brightness temperature measured in the JCMT and the CSO WVM is slightly offset. This might be due to errors in the determination of the load temperature and the coupling. However, as long as the difference is constant it will translate into a constant phase offset which does not degrade the phase correction.

Changes in the measured signal are very similar in the CSO and the JCMT WVM. The small differences probably arise because the mirrors which couple the WVMs to the sky had not been carefully adjusted. It is therefore possible that the monitors observe the sky in slightly different directions. The difference in detected signal can be explained, if the beams only overlap to 80%.

- Three data sets were taken to compare the phase predicted by the WVMs to the phase measured by the interferometer:

  1. On 1st Nov 3C273 was observed. The tau meter at the CSO indicated 0.6 mm precipitable water vapour (pwv), which converts to 0.9 mm along the line of sight to the quasar. The uncorrected phase had 27° rms over 36 minutes once the linear drift, which is probably due to instrumental effects in the interferometer, had been subtracted. Some of the features in the water vapour data agree well with the SBI phase, but phase correction does not improve the data considerably. One of the problems was that this data was taken on the first observing night before the CSO pick up mirror was aligned. Therefore the coupling of the WVM to the sky is not known and it might well be that part of the WVM beam misses the secondary mirror.
  2. A second data set was taken on the bright hydrogen recombination maser MWC349 under very bad weather conditions with 6.7 mm pwv along the line of sight to the maser. The atmosphere was very unstable. Even in the shortest possible integration time of the interferometer, which is limited by the computer speed to 4 sec, the phase of the maser was not measurable.
  3. The third data set was also taken on MWC349, with about 4 mm pwv along the line of sight. The correlation between the interferometer phase and the WVM phase was quite good and phase correction improved the data substantially. This last data set taken will be discussed in more detail in §6.4.

The last interferometer run gave some very good results, which proved that phase correction can considerably improve the phase stability. The water vapour monitors are now working well and should not need much more attention. With a consistent set of software it should be possible to employ the radiometers routinely for interferometric phase correction. The monitors are also able to determine accurately the sky opacity for single dish observations.

6.2 Typical data analysis

As a first step in the data analysis the temperature of the calibration loads (§6.2.1) and the coupling efficiency to the sky (§6.2.2) have to be determined. Once these factors are known the WVM data is smoothed and converted into sky temperatures (§6.2.3). The conversion factor is determined in order to translate the difference in water vapour emission into a phase difference (§6.2.4). This phase difference can then be subtracted from the interferometer phase (§6.2.5).

So far both the WVM data as well as the interferometer data are recorded so that the correction can be applied afterwards during the data reduction. Once the methods are better established real time correction will be possible.

6.2.1 Temperature of calibrating loads

The temperature control on the calibration loads was roughly adjusted to give physical temperatures of 35° C and 100° C. We are not interested in the physical load temperature but their brightness temperature in relation to the sky brightness. Therefore, it would be ideal if we could calibrate the loads against a source, which fills the water vapour beam outside the telescope and whose brightness temperature is well known. Unfortunately, this is not possible, because the only calibrators big enough are the sun and the moon, but their brightness is not well known at 183 GHz since the atmospheric water vapour absorbs a significant amount of their radiation. Instead we calibrate the loads with a piece of absorber at ambient temperature and a piece dipped in liquid nitrogen, which are held in front of the WVM. The ambient temperature can be measured directly with a thermocouple device (electric thermometer) and the load dipped in liquid nitrogen is assumed to be at 80 K. (The boiling point of liquid nitrogen is 77 K, about 75 K at the lower pressure on Mauna Kea, but since the absorber is not a perfect black body and scatters some emission, the brightness temperature of the absorber will be about 5 K higher than its physical temperature.) The load temperature as seen by the detector can then be determined by

$$\begin{aligned} T_{load} = \quad & (\nu_{load} - \tfrac{1}{2}(\nu_{amb} + \nu_{N_2})) \quad \times \frac{T_{amb} - T_{N_2}}{(\nu_{amb} - \nu_{N_2})} \\ & +(\tfrac{1}{2}(T_{amb} + T_{N_2}) \end{aligned} \tag{6.1}$$

where	T_{load}	is the brightness temperature of the calibration load,
	ν_{load}	the frequency counted in the computer,
		when the WVM is looking at the load,
	ν_{amb}	frequency when looking at load at ambient temperature,
	ν_{N_2}	frequency when looking at load dipped in liquid nitrogen,
	T_{amb}	brightness temperature of ambient load,
	T_{N_2}	brightness temperature of load dipped in liquid nitrogen (80 K).

We can easily determine the load temperature to an accuracy of 0.1% of the total detected temperature of around 2500 K (system temperature plus brightness temperature), i. e. to an accuracy of a few K. During one measurement for example, the temperatures determined using channel II suggested 377.5 K and 307 K for the hot and ambient load respectively, using channel III we obtained 380 K and 306 K.

According to a temperature sensor, which measures the physical temperature inside the load, the temperature of the loads are very stable with an rms off less than 3 mK. Therefore, it is sufficient to calibrate the loads occasionally.

6.2.2 Coupling to the sky

Ideally the water vapour monitors would only detect emission from the sky, when the flip mirror is in sky position. However, some emission and scattered light from the telescope will reach the detectors as well. Imperfect mirrors will introduce some scattered light, whereas, for example, the absorber in the centre of the secondary mirror at the CSO will itself emit at ambient temperature. Therefore, the sky emission will only contribute a certain percentage of the total detected signal, this percentage is called the coupling. Clearly we have to know this coupling efficiency in order to deduce the correct sky temperature from the detector readout.

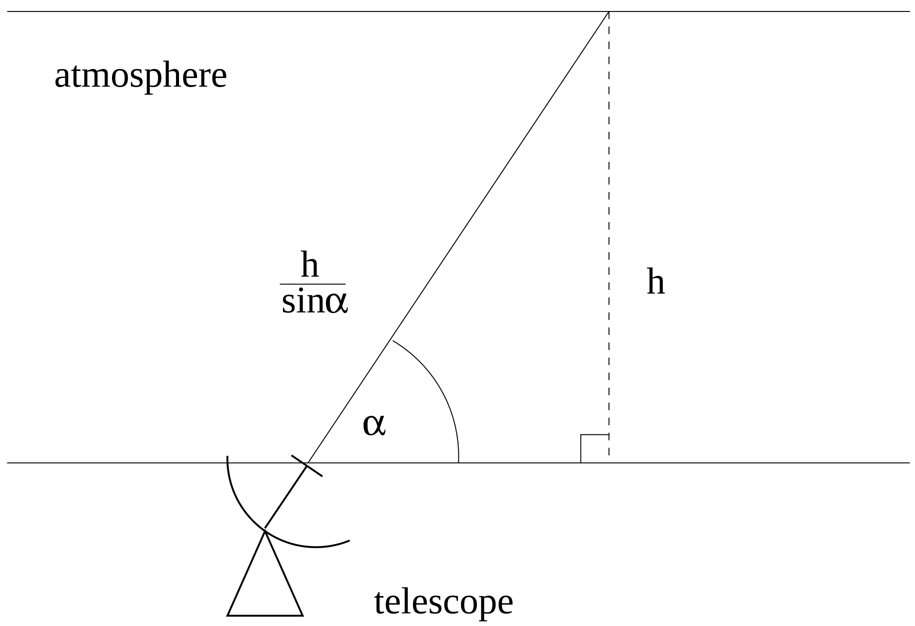


Figure 6.3. Slab model of the atmosphere. The atmosphere has thickness h. The telescope looks at a source of elevation α, so that the optical path length through the atmosphere is $\mathrm{h}/\sin\alpha$.

In order to determine the coupling, the atmospheric emission at different elevations is measured. This is called a sky dip. To first approximation the atmosphere can be imagined as a flat slab (Fig. 6.3). The optical path length in the atmosphere will depend on the elevations α as $1/\sin\alpha$. In the optically thin case the sky emission will also change as $1/\sin\alpha$. Therefore the WVMs will detect:

$$T_{measured} = c\frac{T_{sky}}{\sin\alpha} + (1-c)T_{amb} \tag{6.2}$$

where	$T_{measured}$	the temperature measured by the WVM,
	c	the coupling factor,
	T_{sky}	brightness temperature of the sky,
	α	elevation angle (90° means looking straight up),
	T_{amb}	brightness temperature of telescope parts which reach the horn of the WVM.

All contributions from the telescope (scattering of mirrors, absorber in secondary) are assumed to be at same temperature. For the CSO T_{amb} was assumed to be the air temperature as measured in the telescope dome and for the JCMT the temperature in the receiver cabin.

The sky temperature, T_{sky}, is not known, but by tipping the telescope to at least two different elevations α_1 and α_2 the sky temperature can be eliminated in the above equation. This uses the assumption that the sky emission stays constant over the two measurements. The sky dip measurement should therefore only be conducted under realtively stable weather conditions, where the amount of water vapour fluctuates as little as possible.

Rearranging the above equation we get

$$c = \frac{T_{m1} - \frac{\sin\alpha_2}{\sin\alpha_1} T_{m2}}{(\frac{\sin\alpha_2}{\sin\alpha_1} - 1) T_{amb}} + 1 \tag{6.3}$$

where	T_{m1}	is the temperature measured by the WVM at elevation $\alpha 1$,
where	T_{m2}	the temperature measured by the WVM at elevation $\alpha 2$.

In the case where the emission is not optically thin the equation above gives a lower limit for the couping factor.

Figure 6.4 shows such a sky dip measurement performed with the JCMT, the three lines correspond to the three channels of the WVM.

Channel I is completely saturated at all elevations and detects a constant temperature of 275 K.

The emission at 187.5 GHz, channel II, is also in the optically thick regime because the temperature step sizes are smaller than in channel III. In the optically thin case the step sizes in channel II would be expected to be bigger than in channel III as the step size is $T_{sky} \times c \times (1/\sin\alpha_1 - 1/\sin\alpha_2)$ and T_{sky} is higher in channel II. The amount of pwv confirm that the emission measured by channel II is not optically thin.

The noise on channel III represents the stability of the sky brightness. The constant slope however is due to the telescope tracking some rising protostars. Therefore the elevation of the telescope increases slowly and the detected sky brightness decreases.

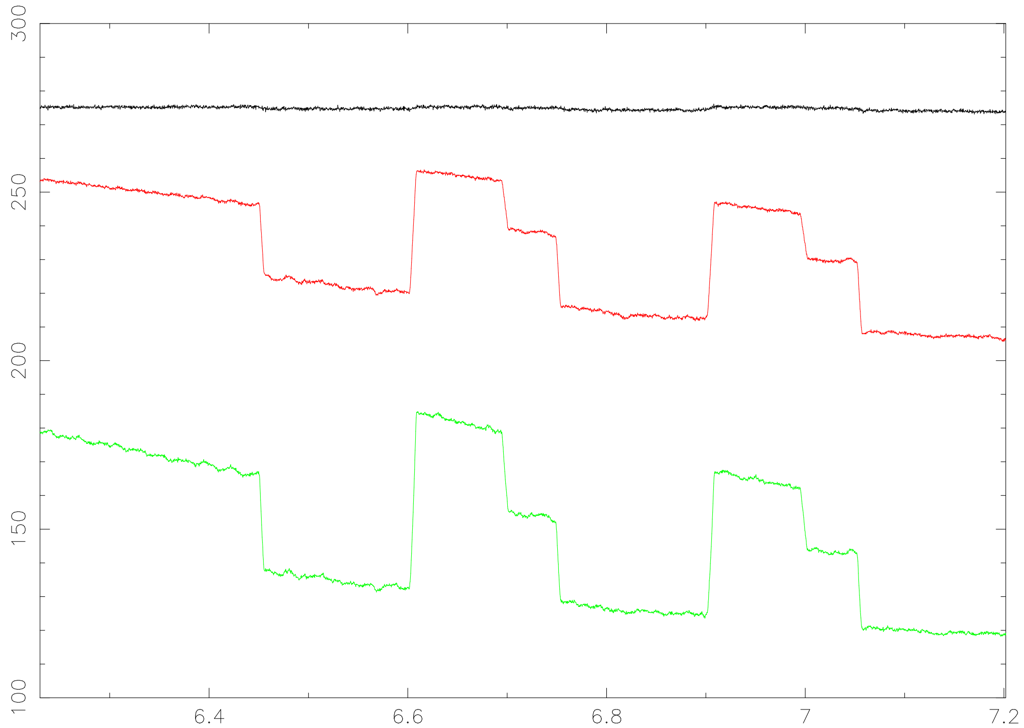


Figure 6.4. Sky dip performed with JCMT. The three lines represent data from the three IF channels of the WVM. Channel I at 183.3 ± 1.2 GHz measures the highest brightness temperature. The line below represents the data of channel II at 183.3 ± 3.7 GHz and the lowest line shows the measurements of channel III 183.3 ± 7.8 GHz. In this case the telescope changed elevation as it was observing three rising sources: first source 2, then source 1,then sources 3, 2, 1, and again 3, 2, 1, where source 1 was at the highest elevation, source 3 at the lowest.

The third channel was used to determine the coupling. Since even the third channel is not completely optically thin, the formula only gave a lower limit for the coupling. The correct value was found, using the fact that the amount of pwv depends on elevation as $1/\sin\alpha$ independent of optical depth. Using a conversion table which takes the effects of optical depth into account (see chapter 3) the sky temperature can be converted into precipitable water vapour (pwv). The coupling factor was then increased until the ratio of pwv equalled $\sin\alpha_1/\sin\alpha_2$.

The determined coupling was 72% for the CSO and 94% for the JCMT with an error of about 2%. The coupling efficiency at the CSO is much lower because of the big piece of absorber in the middle of the secondary mirror. (The absorber has a diameter of 72 mm and accounts itself for a loss in coupling of 22%, assuming the WVM beam is centred on the secondary.) The data from the second channel are consistent with the above results.

The coupling should stay constant and does not need to be remeasured, unless the alignment of the mirrors, or any other relevant parameters, have been changed.

6.2.3 Averaging WVM data

Once the temperatures of the calibration loads and the coupling are determined the WVM data can be reduced routinely and the interferometer phase corrected. First, data points which vary by more then 1% from the previous data point are removed. A 1% change at a system temperature of 2500 K would represent a 25 K temperature change. Such big changes are unlikely to occur in the atmosphere and are therefore most probably artifacts of the instrument. Next the system temperature is calculated and exponentially smoothed. Since the system temperature is expected to be very stable over time a large smoothing factor of exp(t/500sec) was chosen. The frequency output is averaged over half the integration time of the interferometer (usually 4 or 5 sec). Using the smoothed system temperature the averaged frequency is converted into a sky temperature.

6.2.4 Converting temperature to phase

Next the difference in sky temperature between the CSO and JCMT is taken and has to be converted into a phase difference.

This conversion factor can be determined in two ways:

1. From the sky temperature the amount of pwv and the conversion factor which is the inverse of the sensitivity, can be looked up in tables or graphs, which are the output of the atmospheric model (see Fig. 3.11).

2. Alternatively, a bright astronomical source, whose phase is easily detected by the interferometer can be observed. By scaling the water vapour data to fit the phase data best the conversion factor can be found.

Until we have a better feeling for the accuracy of the theoretical predictions of the conversion factor, it is best to measure it with the help of a calibrator located at an elevation similar to that of the observed source.

6.2.5 Phase correction

We chose the two WVM averages immediately before and after the time of the interferometer measurement and average them. (The time in the middle of each integration is used to label the interferometer measurements.) Finally the averaged predicted phase is subtracted from the observed phase.

6.3 Error estimates of corrected phase

In an ideal case the phase predicted by the water vapour monitors should agree perfectly with the phase measured by the interferometer. This is unfortunately not the case, as there are a number of inaccuracies and uncertainties in the measurements and computations, which I will divide into three main classes:

1. Phase shifts introduced by the interferometer itself due to instrumental noise or inaccurate tracking of the astronomical source (§6.3.1).

2. Noise introduced by the water vapour monitors, in particular thermal noise, gain variations, spikes in data and uncertainty in coupling (§6.3.2).

3. Uncertainties in determining the conversion factor from sky brightness to phase (§6.3.3).

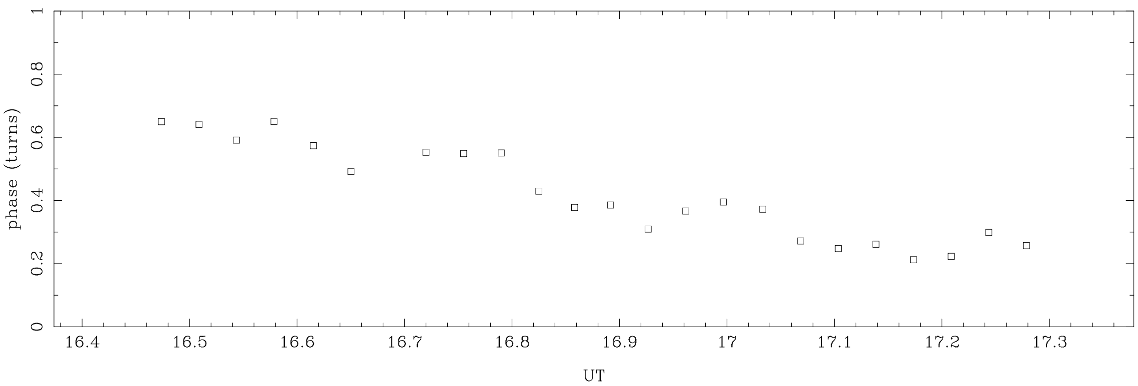


Figure 6.5. Phase measurements of the quasar 3C273 over 1 hour. The linear drift is due to the interferometer, the scatter in the phase is either due to instrumental noise or the atmosphere (1 mm pwv, seeing $0.''6$).

6.3.1 Phase fluctuations caused by the interferometer

With a two element interferometer there is no experiment which could distinguish between phase fluctuations caused by the atmosphere and those introduced by the interferometer itself. However, from past observing runs with the CSO-JCMT interferometer we are quite confident that the interferometer does not introduce substantial phase fluctuations, because the phase is very stable under good weather conditions. On time scales under 10 minutes the rms phase fluctuations due to the azimuth wrap, the azimuth track and the central bearing at the JCMT are expected to lie below 12° (Lay, 1998). Due to instrumental drifts slow linear phase drifts of up to about 1 turn in two hours are observed (Lay, 1998). An example of these linear drifts can be seen in Fig. 6.5. Fortunately, these slow, linear phase drifts are easy to distinguish from phase changes caused by the atmosphere, most of which have a period smaller than 2 minutes. Therefore, a constant slope was removed from the residual phase before calculating the rms of the data.

6.3.2 Instrumental noise introduced by the water vapour monitors

The water vapour monitors introduce most of the uncertainties and noise to the residual phase. These error sources can be divided into 4 classes:

1. thermal noise of WVM,

2. gain variation,
3. spikes in data,
4. coupling of WVM to sky.

Thermal noise

The water vapour monitors are uncooled and their system temperatures are about 2500 K. Let's assume the data is averaged over 8 sec, i.e. the integration time on the sky is $\Delta t = 3.2$ s (8×400 msec, the rest of the time is spent on the calibration loads and on moving the mirror). The thermal noise can be determined using

$$\Delta T = \frac{T_{sys}}{\sqrt{\Delta\nu * \Delta t}}, \tag{6.4}$$

where	T_{sys}	the system temperature,
	$\Delta\nu$	the bandwidth of the WVM channel,
	Δt	the integration time.

Channel I has a bandwidth of $\Delta\nu = 200\ MHz$, giving an rms of 0.1 K, channel II and III each have a bandwidth of 500 MHz giving an rms of 0.06 K. (A lack of coupling will increase the thermal noise. We will discuss its contribution later.) In order to predict the phase the difference of the sky temperature of the two water vapour monitors had to be taken, i.e. the noise is increased by $\sqrt{2}$. The rms in sky temperature is related to the noise in the predicted phase by the conversion factor from Kelvin to phase. As the sensitivity, and therefore the conversion factor, depends strongly on the observing frequency and the amount of underlying water vapour, so does the noise in the residual phase. To illustrate this Table 6.2 gives two examples for observations at 350 GHz and an integration time of 8 sec.

The sensitivities are calculated from the standard atmospheric model and the line profile as described by Waters (1976), see §3.3.4. An exponential distribution of pwv, T_{ground} of 275 K and a pressure of 600 mbar were used.

The above examples show that by using an appropriate channel we can have low errors due to thermal noise even in relatively high water vapour. Also note that observations with an rms noise of $10°$, for example, the noise power is only $\sin 10° = 9\%$.

pwv	channel	sensitivity	$noise = \Delta T/sen$
1 mm	channel I	12 $K/turn$	4°
	channel II	7 $K/turn$	4°
	channel III	$3.5K/turn$	9°
4 mm	channel I	0.5 K/turn	100°
	channel II	3 $K/turn$	11°
	channel III	$2.5K/turn$	13°

Table 6.2. Sensitivity and thermal noise for the three WVM channels for different amounts of precipitable water vapour.

Gain variations

In addition to the thermal noise the water vapour monitors show gain variations, see §5.5.4. In order to determine the gain the calibration loads are observed every second. To reduce the thermal noise several measurements are averaged. The longer the averaging time the lower is the thermal noise, but the worse is the time resolution of the gain variation. Averaging over 8 sec will give an rms uncertainty of approximately 0.2 K for the CSO and 0.05 K for the JCMT. To calculate the phase noise from the thermal noise $\Delta T_{thermal}$ plus the gain fluctuations ΔT_{gv} the uncertainties have to be added quadratically, i.e. $\Delta T = \sqrt{\Delta T_{thermal}^2 + \Delta T_{gv}^2}$ and then the temperature has to be converted into phase, as described above.

Spikes

The are some spikes in the data, see also §5.5.6. At the CSO they have an amplitude of about 2.5 K in channel II and 4 K in channel III, see Fig. 6.6 and Fig. 6.7. The JCMT WVM has much fewer spikes.

Unless the source of the spikes is eliminated or the spikes are removed they will contribute to the error of the predicted phase. The bigger spikes are removed by the data reduction program (see §6.2.3), but spikes of about the same amplitude as the gain variations are not detected by the computer. As the spikes are of different amplitude and frequency one would need to characterise all of them to theoretically calculate their contribution to the phase error. Instead I have chosen a small set of the MWC349 data observed on 8 Nov 1997 at

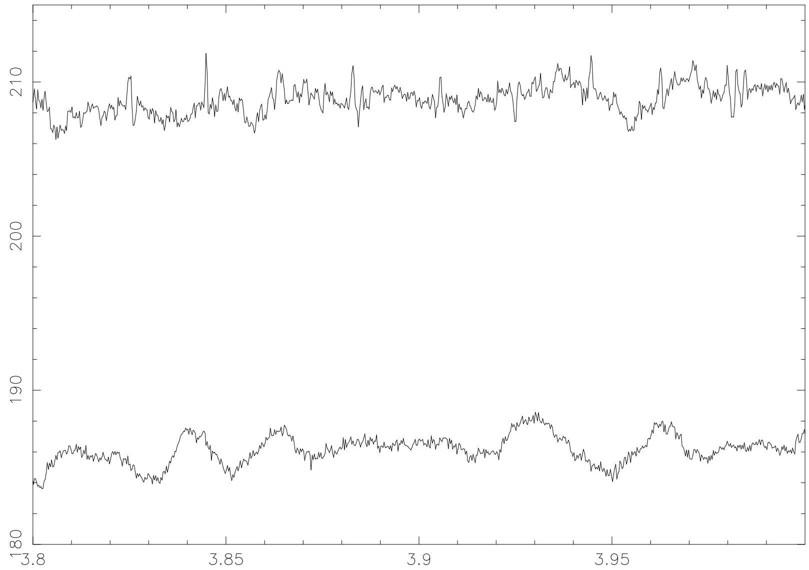


Figure 6.6. Temperatures measured by channel 2 of the CSO WVM (top) and the JCMT WVM (bottom). The CSO data has spikes of ∼2.5 K amplitude and seems noisier (fast variations of ∼ 2 sec). The temperature changes are not as pronounced because the coupling is worse. The scale on the x-axis is in hours and on the y-axis in Kelvin. The temperatures measured by the JCMT and CSO WVMs are different because the coupling has not been corrected for.

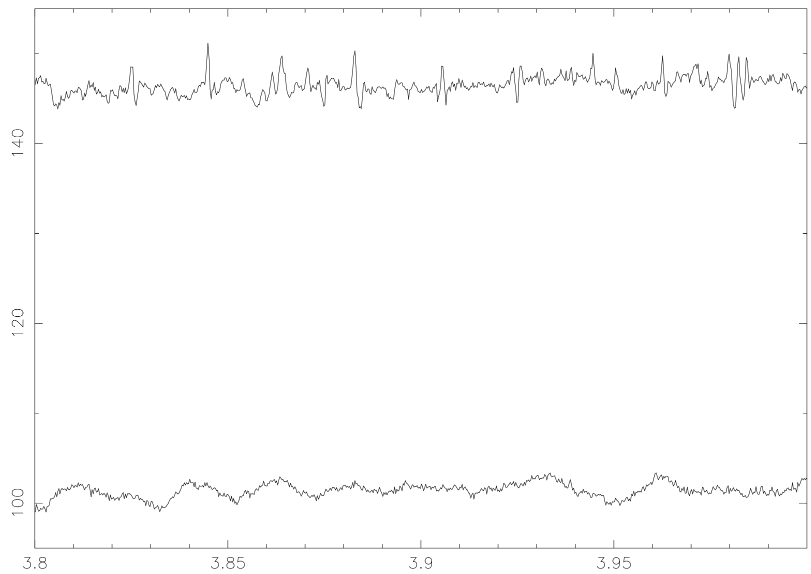


Figure 6.7. Plot similar to Fig. 6.6 but for channel 3 of the CSO WVM (top) and the JCMT WVM (botton).

UT = 3.8 to 4.0 (see Fig. 6.11) and calculated the rms with the spikes in the data and after removing some spikes from the WVM data. All of these spikes are from the CSO wvm. For phase corrections with channel II the rms of the residual phase is 48°, after removing the most obvious 10 spikes in this 12 minute sample the rms goes down to 46°. Removing 5 further small spikes, which were basically picked due to the residual phase deviating rather than the WVM data looking suspicious, lowered the rms to 38°. As the spikes are bigger in channel III, the rms phase is improved more by removing these spikes, i.e. from 84° to 80° after removing 10 and to 68° after removing 15 spikes. We can attribute an rms error to the spikes by $rms^2_{spikes} = rms^2_{total} - rms^2_{without\ spikes}$. The rms_{spikes} lies between 14° and 30° for channel II and between 25° and 49° for channel III. The size and frequency of the spikes, however, vary considerably between the data sets and it is therefore impossible to attribute one single rms value to all of them.

Coupling to sky

The water vapour monitors do not couple perfectly to the sky, i.e. part of the detected signal is not due to the sky, but is for example caused by the absorber in the secondary mirror of the CSO. The imperfect coupling raises the phase noise in two ways.

1. If the coupling, c, were precisely known the thermal noise and the noise due to gain variations rise by 1/c because the sky temperature is determined from the measured temperature by:

$$T_{sky} = \frac{1}{c}(T_{measured} - T_{absorber}) + T_{absorber}. \tag{6.5}$$

   One can also think of it as the system temperature being raised by 1/c.

2. The second effect the coupling has on the noise is that it will not be precisely known and therefore the uncertainty in the coupling will add to the phase error.

   For example, if the coupling is underestimated, firstly, the deduced sky temperature will also be underestimated (because $T_{measured} < T_{absorber}$ in Equation 6.5). Secondly, any temperature changes due to fluctuating water vapour will be enhanced.

   (a) The first effect of underestimating the sky temperature is to give a constant phase offset, but this will not degrade the performance of the phase correction; since it

	JCMT		CSO	
Error source		subtotals		subtotals
Thermal noise	0.06 K		0.06 K	
Gain variations	0.05 K		0.2 K	
Quadratic sum		0.078 K		0.21 K
Coupling	95%		75%	
Error/coupling		0.082 K		0.28 K
Inaccuracy in coup.	0.01 K		0.02 K	
Quadratic sum		0.083 K		0.28 K
Conversion to phase		7.9°		26.5°
Spikes	0°		30°	
Total error		8°		40°

Table 6.3. Theoretical errors for the JCMT and the CSO WVM data.

does not matter whether the residual phase lies at 0° or any other constant value. (If we determine the conversion factor theoretically from the the sky temperature this error will, however, reflect in the uncertainty of the conversion factor, see below.)

(b) An enhancement of the phase fluctuations however has a direct effect on the phase error. The coupling factor is known to within 2%. Typical sky brightness fluctuations have an rms of ∼0.5 K. In the worst case of the CSO coupling the error introduced is $2\% * 0.5\ K/c^2 = 0.02\ K$. This error should be (quadratically) added to the other errors and then converted into a phase error as demonstrated at the beginning of this section.

Comparison of theoretical noise with measured noise

In order to verify to what extent these theoretical error estimates correspond to the real residual phase error, let us compare it to the maser data from 8 Nov 1997 UT = 3.8 to 4.0 corrected by channel II of the WVM data. The measured rms was 48°, the conversion factor was $(3.8\ \mathrm{K/turn})^{-1}$ and the averaging time was 8 seconds (integration on the sky was 3.2 seconds). The contributions to the theoretical error are listed in Table 6.3. The total rms phase error is the quadratic sum of the JCMT phase error and the CSO phase error giving 40.5°. It is dominated by the gain variations and the spikes in the CSO WVM.

The calculated rms lies below the measured rms of 48°, i.e. a phase error of 26° is not accounted for. The WVM data and the interferometer data were not always taken at exactly the same time, though their 8 second integrations overlapped by at least 5 seconds. This might be part of the reason for the higher observed rms. Spikes in the JCMT data could have increased the noise. Other contributions to the rms phase could come from the interferometer itself (§6.3.1) or an uncertainty in the conversion factor (§6.3.3).

6.3.3 Uncertainty in conversion factor

For the maser data presented earlier a conversion factor from Kelvin to phase was chosen to give the best phase correction. In this case an uncertainty in the conversion factor should not significantly increase the phase error. Clearly, in the case of observing a faint astronomical object the interferometer phase will be dominated by noise and we will have to determine the conversion factor by either observing a calibrator or ideally by predicting it theoretically. For faint sources the uncertainty in the conversion factor will contribute to the phase error.

Determining conversion factor with calibrator

When determining the conversion factor with a bright calibrator and applying the same factor to the faint source, there are three sources of error:

1. There will be a small error in determining the correct conversion factor for the calibrator. By using a computer program that determines the conversion factor by minimising the rms phase one should be able to keep this error below 5%.

2. Faint source and calibrator are observed at different times, over which the conversion factor will change, due to changing amounts of water vapour. For an estimated change in pwv of 10% over one hour during stable observing conditions the error in the conversion factor in channel II would lie between 5% and 10%, depending on the amount of pwv.

3. Calibrator and astronomical source will be at different positions, which will cause an error in the conversion factor. As long as both are at roughly the same elevation, so that there is roughly the same amount of water vapour along the line of sight, the error should not exceed the 10% estimate quoted above.

In addition the source will change elevation with time, so that the amount of water vapour along the line of sight and therefore the conversion factor will change. When the source is at low elevation and the airmass changes rapidly, more frequent calibration will be needed.

Theoretical prediction of the conversion factor

There are two error sources in theoretically predicting the conversion factor:

1. Firstly, the models used to predict the conversion factor (see chapter 3) could be inaccurate. The little good data obtained so far do not allow us to verify the models.

2. Assuming the models are correct, the error in the conversion factor depends on the uncertainties of the many variables the conversion factor is calculated from. In section §3.3.4 we estimated the percentage error in the sensitivity [1] for an unknown height of the pwv between 200 m and 1000 m above the telescope and an uncertainty of 1% in the extrapolation from the load temperature to the brightness temperature. Depending on the channel and the amount of water vapour the errors lie between 2% and 33%.

Since the error in the conversion factor depends on how it was determined as well as on the weather conditions it is difficult to attribute a single phase error to it. But in order to illustrate its effect, let us look at a typical example, where the error of a conversion factor of $(3\ \mathrm{K}/360^\circ)^{-1}$ is 10%. This will give rise to a phase error of 6.5° assuming a rms temperature fluctuation of 0.5 K.

Summary of error estimates

The errors mentioned above all contribute to the rms in the residual phase. The most dominant errors are caused by spikes and the gain variations in the CSO WVM. Since the spikes seem to be an artifact of either the data taking program or a problem with the signal transmission, it should be easy to remove them, which would considerably reduce the rms phase. But even with all these errors we corrected the interferometer phase to an rms of 23° at 350 GHz in November 1996, which is already a very good result.

[1] The conversion factor is the inverse of the sensitivity. Their percentage errors are therefore identical.

6.4 WVM data: correlation functions and phase correction

To judge the performance of the WVMs this section will analyse the last data we took in great detail. First the auto- and cross-correlation of the water vapour data is taken, which gives a good indication of the quality of the data. Secondly, the phase predicted by the water vapour monitors is compared to the interferometer phase and the results of phase correction are discussed.

The last WVM data was taken in the evening of the 8th November 97, when we observed the bright maser MWC349 at 356 GHz with the interferometer and simultaneously monitored the atmospheric water vapour. The weather was relatively poor with 3.3 mm of pwv and an unstable atmosphere.

6.4.1 Auto- and cross-correlation of WVM data

The auto- and cross-correlations of the water vapour monitor data were taken according to the following formula:

$$J_iC_j(\tau) = \frac{\int_0^T (J_i(t) * C_j(t+\tau)dt)}{\left(\int_0^T J_i(t)^2 dt \int_0^T C_i(t+\tau)^2 dt\right)^{1/2}} \tag{6.6}$$

where $J_iC_j(\tau)$ the cross-correlation between JCMT channel i and CSO in channel j with time delay τ

$J_i(t)$ sky temperature of JCMT WVM channel i at time t,

$C_j(t+\tau)$ sky temperature of CSO WVM channel j at time $t+\tau$.

The cross-correlations of all three frequency channels of the WVM are plotted against the time delay τ in Fig. 6.8. The sky temperatures were exponentially smoothed with a time constant of 0.005/s and then subtracted from the original data. This process removes any large scale changes in the amount of water vapour either due to weather changes or due to changes in airmass (i. e. changes in elevation of the source), so that only the short time variations (5 - 100 s) are left. In this normalization a correlation of 1 indicates perfect correlation, whereas a value of 0 represents random data.

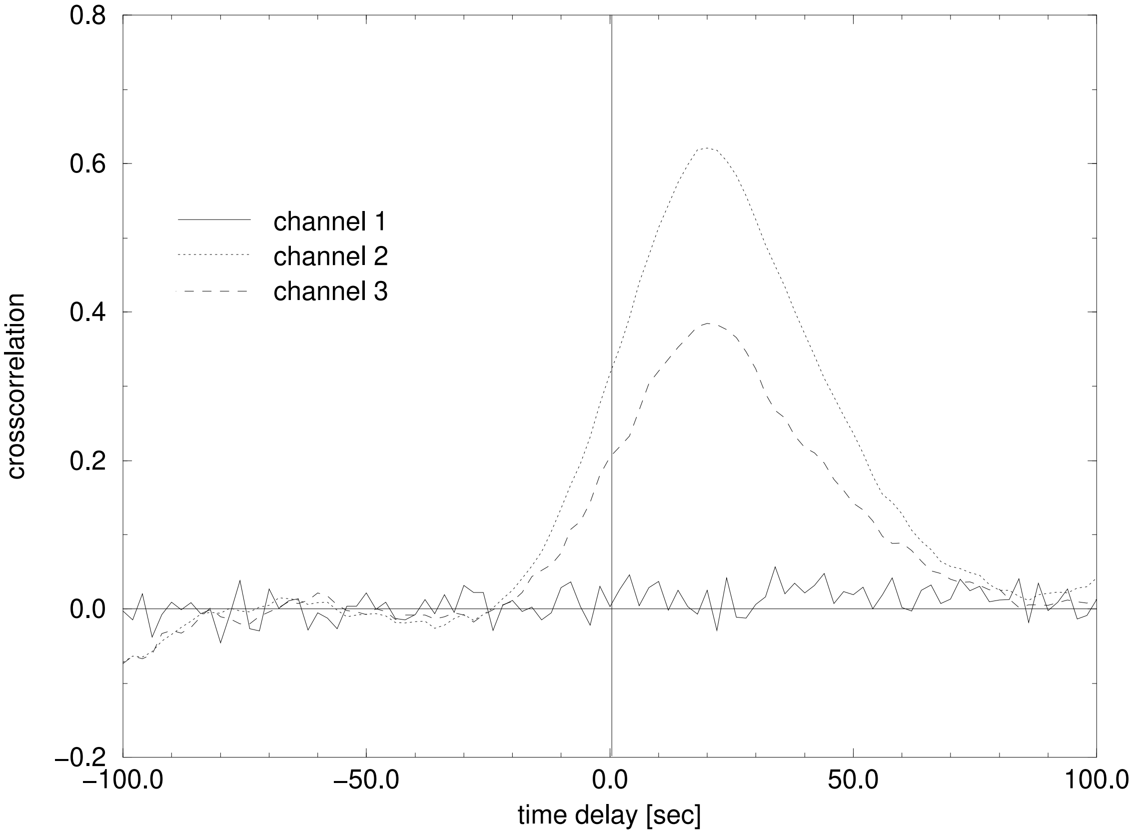


Figure 6.8. Cross-correlation between the JCMT and the CSO WVMs.

The cross-correlation J_1C_1 is zero because the first channel is saturated. Any fluctuations measured in channel I are due to thermal noise, hence these fluctuations are not correlated between the radiometers.

The cross-correlation J_2C_2 and J_3C_3 peaks at $\tau = 22$ sec, suggesting that the JCMT data is similar to the CSO data taken 22 sec later. The frozen screen model can explain this easily: According to this model the water vapour distribution is not broken up, instead wind spacially shifts the undisturbed screen of vapour. Since the baseline of the interferometer is 160 m a windspeed of 7.3 m/s parallel to the baseline is needed to explain the maximum in the cross-correlation at 22 sec. In fact, a west-north-westerly wind (300°), i. e. parallel to the interferometer baseline, of 13 mph (5.8 m/s) was measured on the weather station next to the JCMT. The higher windspeed seen in the interferometer data is presumably due to the water vapour being at higher layers in the atmosphere, where the windspeed is expected to be slightly higher than at ground level.

The auto-correlations is shown in Fig. 6.9. As expected the auto-correlation (e.g. J_2J_2) is symmetric around $\tau = 0$ and has a value of 1 at $\tau = 0$. It drops slowly with $|\tau|$ increasing, because data sets taken shortly after each other will still be correlated to some extent, as the amount of water vapour won't change abruptly. Over time scales greater than $\sim$50 s the fluctuations are not correlated any more. The CSO auto-correlation drops much more quickly due to the spikes in the CSO data. A data point with a spike in it is not well correlated to the data point just taken before or after it. The auto-correlation of channel I drops immediately to zero because channel I is saturated and measures (uncorrelated) thermal noise.

The auto- and cross-correlation function give a good indication of the quality of the water vapour data. Channel I is saturated and there is the problem of spikes in the CSO data. But the maximum in the cross-correlation function which predicts the wind speed correctly indicates that the data quality in channel 2 and 3 is good.

6.4.2 Phase correction

The best test for the water vapour monitors is to compare the phase predicted by the radiometers to the phase measured by the inteferometer.

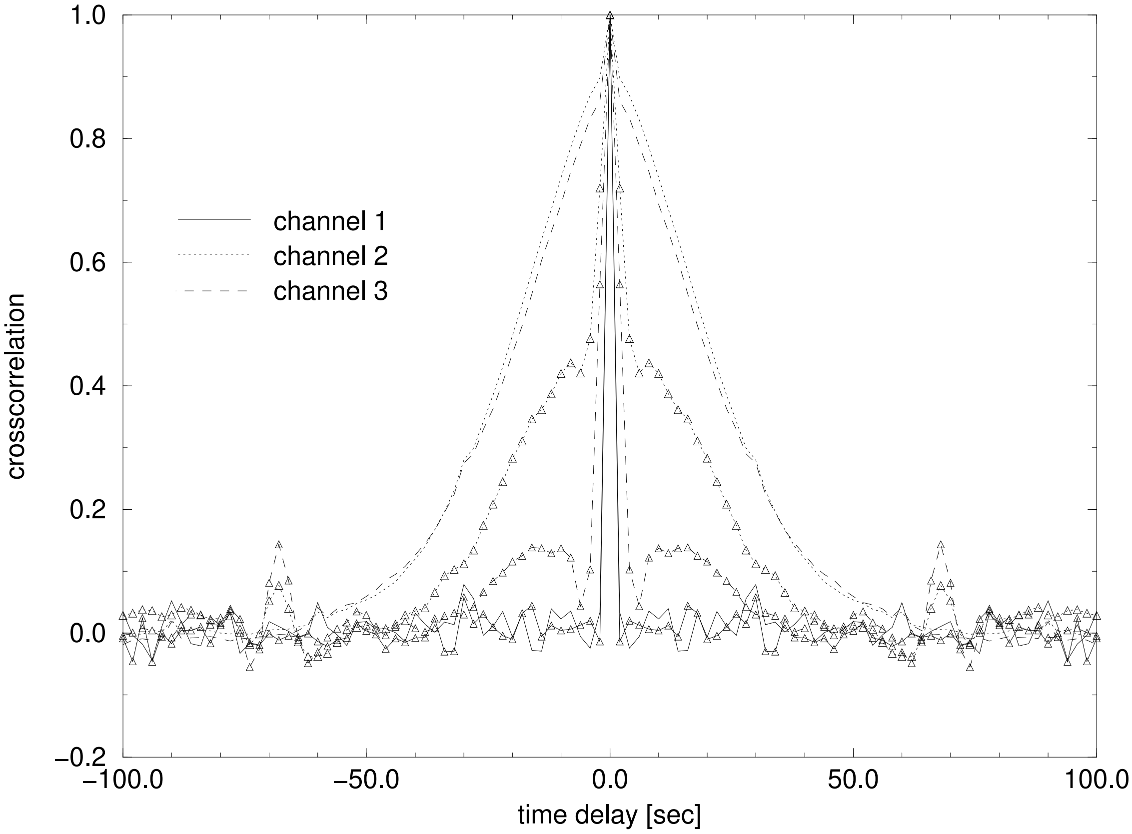


Figure 6.9. Auto-correlation of JCMT WVM (no symbols) and CSO WVM (triangles).

For this comparison the WVM data were analysed as described in section §6.2 of this chapter averaging the data over 4 seconds.

The interferometer data were reduced with the standard software (Lay, 1994). The maser data consists of spectra of 500 MHz bandwidth, which clearly show highly peaked redshifted and blueshifted emission of the hydrogen recombination line $(27 \rightarrow 26)^2$, see Fig. 6.10. Scatter of the phase, symbolized with dots in the figure, is due to noise from atmospheric emission. Clearly in the emission peaks the noise is very small, so that the phase measures the optical path difference between the JCMT and the CSO. The phase averaged over channels 600 and 860 for every 8 second integration should therefore correlate with the difference in sky brightness measured by the two WVMs.

Figure 6.11 shows the interferometer phase in comparison to the phase deduced from the WVM data.

1. The upper most graph shows the phase predicted from the measurements of channel I (solid line), the channel closest to the line centre. This channel is saturated and the fluctuations on the plot represent thermal noise rather than fluctuations in the amount of water vapour.

2. The data from channel II fits the interferometer phase best. A conversion factor of $(3.8\ \mathrm{K}/360°)^{-1}$ was used. (This value was chosen because it fits the interferometer phase best. The theoretically predicted sensitivity lies between 3.1 and 3.7 K/360°.). Correcting the interferometer phase with this data reduces the 127° rms of the interferometer phase to 48° in the 12 minute period. The original interferometer phase and the residual phase after subtracting the predictions from channel 2 are plotted in the fourth graph. Most of the spikes in the residual phase, e. g. at 3.98 hours UT, come from spikes in the CSO data. These spikes are also clearly visible in the WVM data in channel 2 and are even more dominant in channel 3. By manually removing the 15 most prominent spikes from the data the rms of the residual phase drops to 38°.

[2] The hydrogen recombination lines are due to electrons recombining with a hydrogen atom and then cascading down through different levels to the ground state. Here we observe electrons falling from level 27 to 26.

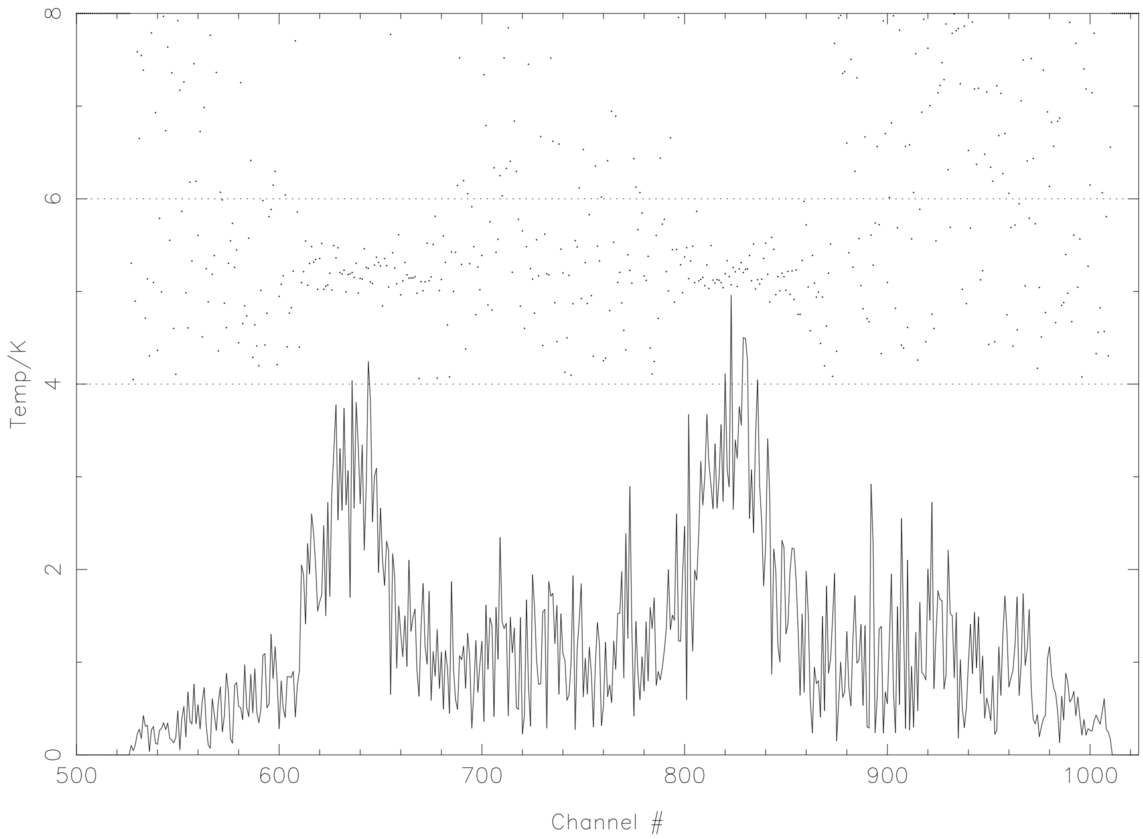


Figure 6.10. Spectrum of MWC 349 integrated over 160 sec. The dots represent the maser phase between −180 (lower dotted line) and 180° (top of graph). The solid line represents the amplitude of the maser, the scale is marked on the y-axis. (An antenna temperature of 1 K is equivalent to 65 Jy.) At the location of the peaks in the amplitude spectrum the phase is clustered. Outside the peaks the phase is random, as it is dominated by noise from atmospheric emission.

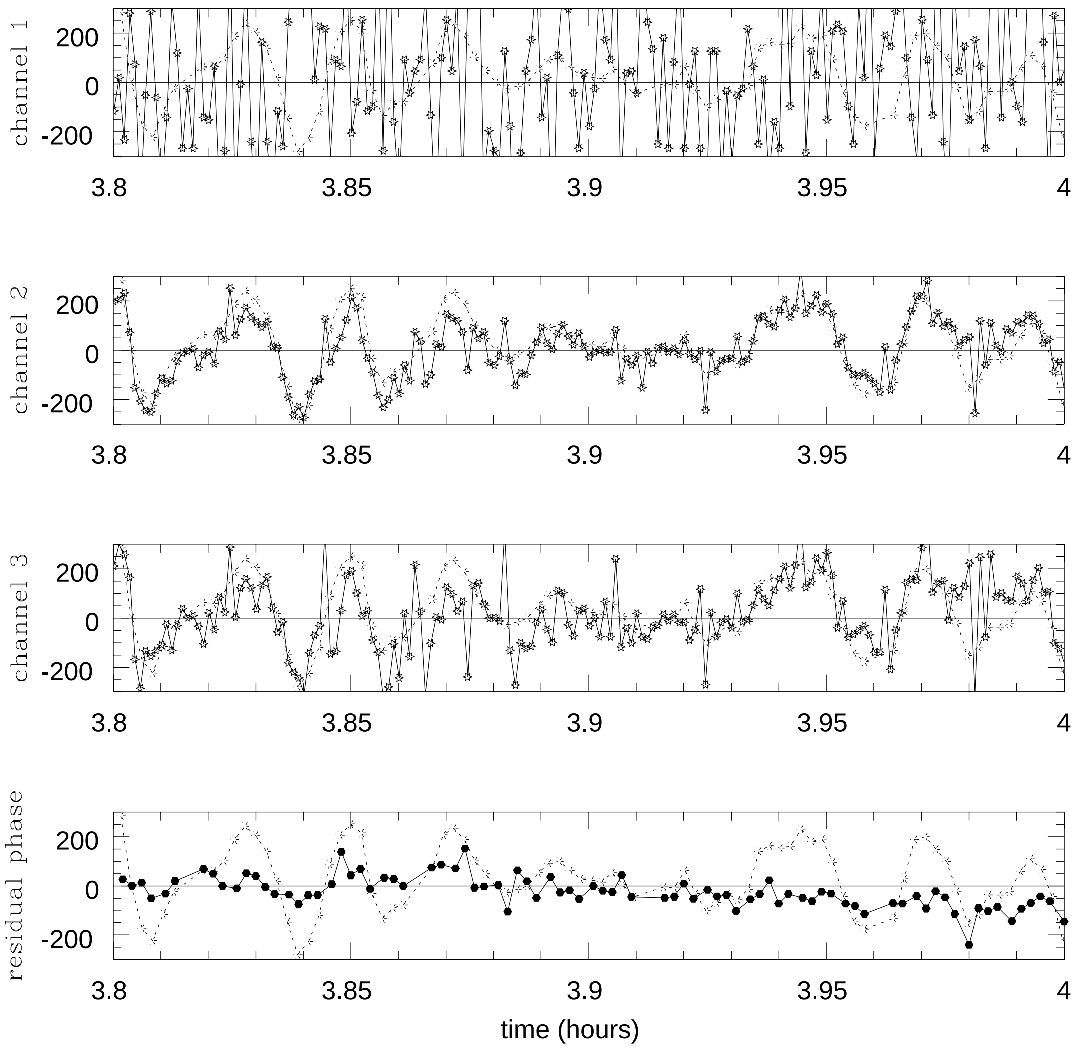


Figure 6.11. Phase of MWC349 measured by the interferometer (dashed line) overlaid on phase predicted by channel 1 to 3 of the water vapour monitors (solid line in first, second, third plot, respectively). The rms of 127° of the original interferometer phase (dashed line) is reduced to 48° after correction with channel 2 (solid line).

3. The third plot shows the phase predictions from channel III of the water vapour monitor overlaid over the interferometer phase. With a conversion factor of $(3.4\ \mathrm{K}/360°)^{-1}$ the rms of the corrected interferometer phase was reduced to 84°. (Theoretically predicted sensitivities between 2.4 and 2.6 K/360° gave worse results. It might be that the higher sensitivity of 3.4/360° gives lower rms because devision by a bigger sensitivity will make the spikes less dominant and therefore reduce the rms.) Correction with the third channel is not as good as that with the second channel partly due to the lower sensitivity and therefore greater contribution of thermal noise, but mainly due to the much more prominent spikes.

To attach some meaning to the rms value one can imagine that in an average visibility vector detected by the interferometer deviates from 0° (or any other constant value) by e.g. 48°. When integrating the signal the interferometer basically takes the vector average of theses voltage vectors. The resultant vector will be shorter by cos 48°. (Fig. 7.9 illustrates such a vector average. The more the individual vectors deviate from the mean phase the shorter the resultant vector gets.) The power detected in the integrated signal is $\cos 48° = 0.67$ of a signal with no phase fluctuations (rms=0°). The phase noise is defined as sin 23°. Reducing the rms from 55° to 23° reduces the phase noise by 52%. (This analysis is not valid for large rms values ($\geq 60°$) because the angle of the phase vector can only take values between $+180°$ and $-180°$ and a different analysis is needed. This has not been further investigated because we are most interested in small rms values.)

The rest of the data from the same evening look very similar and give similar rms values. Some more examples are shown in the appendix.

6.5 Future work and summary

6.5.1 Future work

The data analysis above showed that the performance of the water vapour monitors could be even better by reducing the gain variations in the CSO WVM and by eliminating the spikes in the data.

So far the conversion factor has been chosen so that the predicted water vapour phase fits the interferometer phase best. When observing a faint source the conversion factor can be determined every few hours with the help of a bright calibrator. However, there should be enough information in the radiometer data to theoretically predict it. Current research (Prado-Carrión, 1996) is taking place to confirm the validity of the atmospheric model and to find an inversion algorithm to predict the amount of water vapour and its physical conditions to higher accuracy. It would be also very interesting to investigate the distribution of water vapour with height and its physical conditions to determine the range of typical conditions. It should then be possible to theoretically predict the conversion factor to higher precision.

To prove that the water vapour is the cause for most of the phase fluctuation, the phase predicted from two further 183 GHz radiometers, which have been modeled on the instruments described here, will be compared to the phase measured by a phase monitor.

Besides investigating the models and the distribution of the atmospheric water vapour it would be very interesting to compare the 183 GHz radiometers with other phase correction methods. There has been a suggestion to mount two 183 GHz radiometers on the IRAM Plateau de Bure interferometer. Since the IRAM interferometer uses total power observations at 230 GHz for its phase correction and since additionally some 22 GHz radiometers are being installed, a direct comparison would be possible. This comparison would demonstrate the advantages and disadvantages of the different phase correction methods.

6.5.2 Summary

Though we only have limited data so far, data analysis above already demonstrates that the water vapour monitors work well. Under good weather conditions the rms phase was reduced from 55° to 23°, reducing the phase noise by 52%. Even under bad weather conditions the rms of the interferometric phase could be reduced from 127° to 48°. This clearly demonstrates that the WVMs can considerably reduce the phase noise due to the atmospheric phase fluctuatuions. Moreover the results prove that the water line at 183 GHz is very sensitive to changes in the amount of atmospheric water vapour and is well suited for phase corrections. Radiometers at 183 GHz are a viable system for future interferometers especially where the optical path length needs to be known to high accuracy. Only sensitive monitoring of the water vapour content and corrections of the phase fluctuations makes high frequency interferometry (in particular for long integration times) possible, which will open up a whole new spectrum of astronomic research which will lead to many exciting discoveries.

Chapter 7

Interferometric observations of Arp 220

This chapter describes one of the first extragalactic observations performed with the JCMT-CSO interferometer. Arp 220 was considered the ideal candidate for interferometric submillimetre observations of a galaxy because it is the closest ultraluminous infrared galaxy and can therefore be studied at high physical resolution. Moreover observations have high signal to noise because of the high fluxes. Arp 220 is also easy to model since it has been extensively studied at other wavelengths.

A single baseline interferometer observes only along a single (u,v) track and hence full two dimensional images cannot be produced. However, fluxes and spectra can be measured and existing models and predictions can then be verified or rejected. This chapter therefore starts with a detailed account of previous measurements and suggested models of Arp 220 (§7.1). The next section (§7.2) introduces the observations performed. §refsec: basic reduction explains the basic data reduction techniques. There then follows a discussion of the continuum (§7.4) and the line data (§7.5) in more detail - explaining how the information was derived from the data, as well as comparing the results with previous observations and models. §7.6 summarises the results.

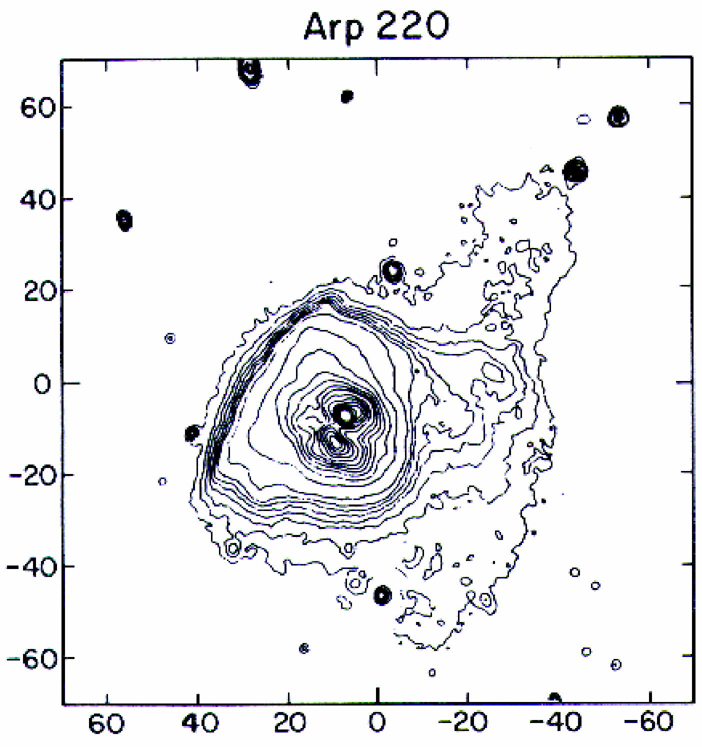


Figure 7.1. Contour map from CCD image at 6500 Å. Contour intervals are a factor of two apart. The axes are labeled in arcseconds. Figure taken from Sanders et al. (1988).

7.1 Background on Arp 220

At a distance of 73 Mpc ($H_0 = 75$ km s^{-1} Mpc^{-1}) Arp 220 is the nearest ultraluminous infrared galaxy (ULIRG). ULIRGs are defined to have far-infrared luminosities ($8 - 1000\mu$m) exceeding $10^{12} L_\odot$ (Sanders et al., 1988), Arp 220 has a flux of $1.5 \times 10^{12} L_\odot$ (Sanders et al., 1988). Most ULIRGs are recent mergers.

In the optical Arp 220 exhibits faint tidal tails (Joseph and Wright, 1985) and two bright knots on either side of a dust lane, which suggest a recent collision of two galaxies (Sanders et al., 1988), Fig. 7.1. Two nuclei separated by $0\farcs95$ in the east-west direction (position angle (P.A.) 93° east of north) are seen in the near infrared at 2.2μm (Scoville et al., 1998) (Fig. 7.2). In the radio at 1.3 cm, 2 cm and 6 cm emission peaks at a similar location to the infrared are observed (Becklin and Wynn-Williams, 1987).

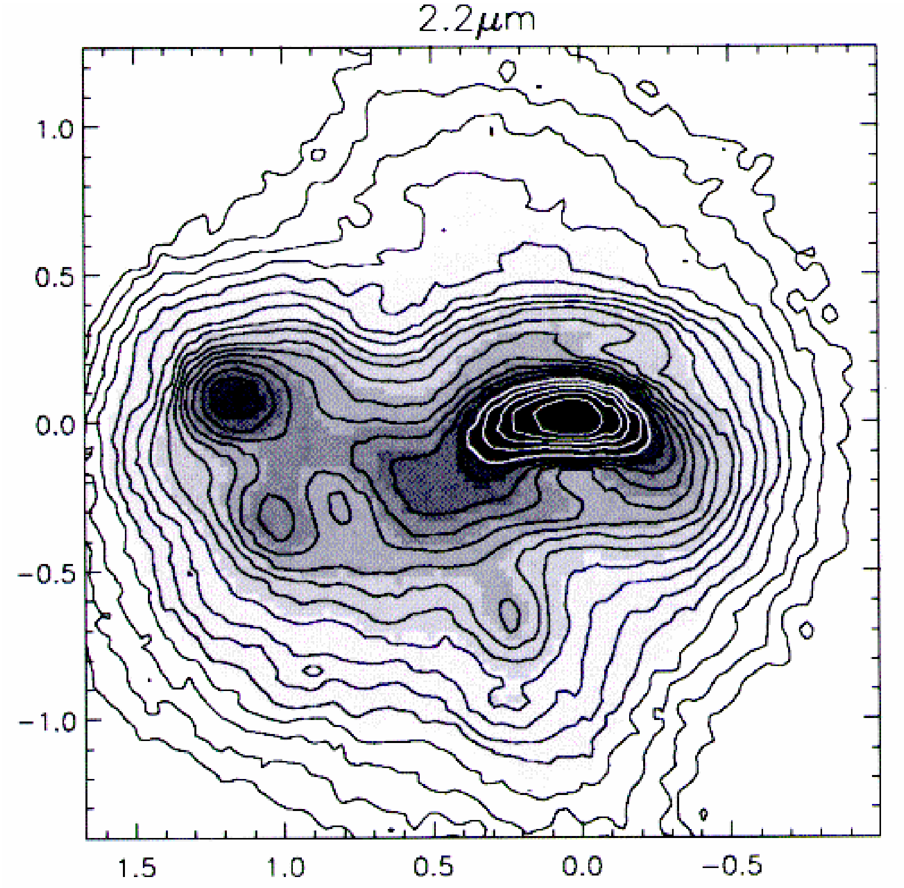


Figure 7.2. Contour and grey scale images for the the central 2.″5 of Arp 220. The contours are spaced logarithmically by a factor of 1.16 down from the peak on the western source at 20 mJy arcsec^{-2}. The coordinates are offset in α and δ from the peak ($\alpha_{1950} = 15^h32^m46.^s90, \delta_{1950} = +23°40'07.''94$). Figure taken from Scoville et al. (1998).

Arp 220 contains high molecular gas masses of $\sim 9 \times 10^9 M_\odot$ (Scoville et al., 1997) consistent with the $4 - 40 \times 10^9 M_\odot$ (Sanders et al., 1991) measured in other ULIRGs (c.f. our Galaxy $2 \times 10^9 M_\odot$). Furthermore this gas is highly concentrated, as 2/3 of the total CO $1 \rightarrow 0$ emission arises from the central 600 pc (Scoville et al., 1991). The high concentration of gas and dust results in an extinction of $A_V = 50 - 1000$ mag, dependent on whether a screen or a mixed model is assumed for the dust distribution (Sturm et al., 1996, Downes and Solomon, 1998) .

This high abundance of interstellar matter plays a critical role in the energy output of ULIRGs and of Arp 220 in particular as it could either fuel an active galactic nucleus (AGN), as advocated by Scoville et al. (1997), or it could provide the material for a starburst (Lutz et al., 1996).

A starburst hypothesis for Arp 220 is supported by 18 cm VLBI observations, the flux of which is probably emitted by luminous radio supernovae (Smith et al., 1998) rather than by a single point source. The observed X-rays are believed to be created in superwinds driven out from the nucleus by a starburst (Heckman et al., 1996). The high intensity of low excitation lines in the near infrared further support the starburst scenario (Sturm et al., 1996). Finally 35 μm and 53 μm absorption measurements in OH-megamasers suggest that they are radiatively pumped in a warm, extended region characteristic of a starburst (Skinner et al., 1997).

Our observations will be compared to the results of the following four papers:

1. Baan and Haschick (1995) have observed Arp220 with the VLA in continuum at 4.83 GHz and in a transition line of formaldehyde (H_2CO) . They propose that the continuum, produced by synchrotron emission, traces the true nuclei, which are separated by $1\farcs06$, P.A. $= 98°$. The nuclei orbit each other at about 200 km s^{-1} in a plane inclined at 45°. The OH maser emission comes from a region slightly ahead of the continuum emission, while the formaldehyde emission is trailing slightly behind.

2. Near-infrared observations with the NICMOS camera on HST at a resolution of $0\farcs11$ – $0\farcs22$ show two peaks of emission at 1.1 and 1.6μm, but three at 2.2μm (Scoville et al., 1998). However, even the 2.2μ m observations are strongly affected by dust obscuration: the eastern nucleus is bifurcated by a dust lane, so that the infrared emission peaks lie south and north of the position of the true nucleus; the western nucleus has a crescent

shape, consistent with a circumnuclear ring or a spherical cluster obscured by a dust disk (Fig. 7.2).

3. CO and millimetre continuum observations trace the molecular gas and dust. Arp 220 has been mapped by OVRO (Owen's Valley Radio Observatory) in CO $1 \rightarrow 0$ and $2 \rightarrow 1$ as well as in the adjacent continuum at a resolution of $1'' - 2''$ (Scoville et al., 1997, hereafter SYB). The deconvolved 229.4 GHz continuum emission comes from an elliptical region with minor and major axes of 200 pc ($0\overset{''}{.}63$) and 370 pc ($1\overset{''}{.}07$) respectively at P.A. of 91°. Approximately two thirds of the CO $1 \rightarrow 0$ emission is contained in a nuclear source of radius ~300 pc. The central CO $2 \rightarrow 1$ emission has minor and major axes of 450 pc ($1\overset{''}{.}28$) and 680 pc ($1\overset{''}{.}94$) elongated along P.A. 53°, parallel to the dust lane observed at optical wavelengths. The "cleaned" CO $2 \rightarrow 1$ maps show two nuclei separated by $0\overset{''}{.}95$ at P.A. 101° and with a velocity difference of 250 – 300 km s^{-1} (Fig. 7.3). SYB suggest that $5.4 \times 10^9 M_\odot$ of molecular gas is concentrated in a thin (16 pc), rotating disk inclined at $40 - 50°$ to the line of sight (see Fig. 7.4). The high brightness temperatures imply an area filling factor of ~ 0.25. SYB infer a mean gas density in the disk of 2000 cm^{-3}. The two nuclei rotate with this disk at 310 km s^{-1} at a radius of 235 pc, and also have smaller accretion disks of their own, as the CO spectra of the nuclei are double peaked. Assuming that the accretion disk of the nuclei has a radius of $0\overset{''}{.}25$ (50 pc) the dynamical mass inside this radius is $4 \times 10^8 M_\odot$.

4. Downes and Solomon (1998, hereafter D&S) have independently mapped Arp 220 in CO $2 \rightarrow 1$, CO $1 \rightarrow 0$ and in the continuum with the Plateau du Bure interferometer (Fig. 7.5). Though most of their observations and the main conclusions are similar to those of SYB there are a few differences. Firstly D&S 1.3 mm continuum observations show two emission maxima separated by $0\overset{''}{.}8$ at P.A. 101°. Secondly, their CO observations trace two different features approximately at the position of the eastern nucleus: red-shifted gas at 5580 – 5740 km s^{-1}, $1\overset{''}{.}3$ away from the western nucleus at P.A. 85°, in the same orientation as the K-band NE source (Scoville et al., 1998); as well as blue emission between 5220 and 5420 km s^{-1}, $0\overset{''}{.}85$ away from the western nucleus at P.A. 110°. Thirdly, D&S extended SYB's model by adding an outer disk of $8''$ diameter and an eastern streamer extending $7''$ perpendicular to the disk (Fig. 7.6). Fourthly,

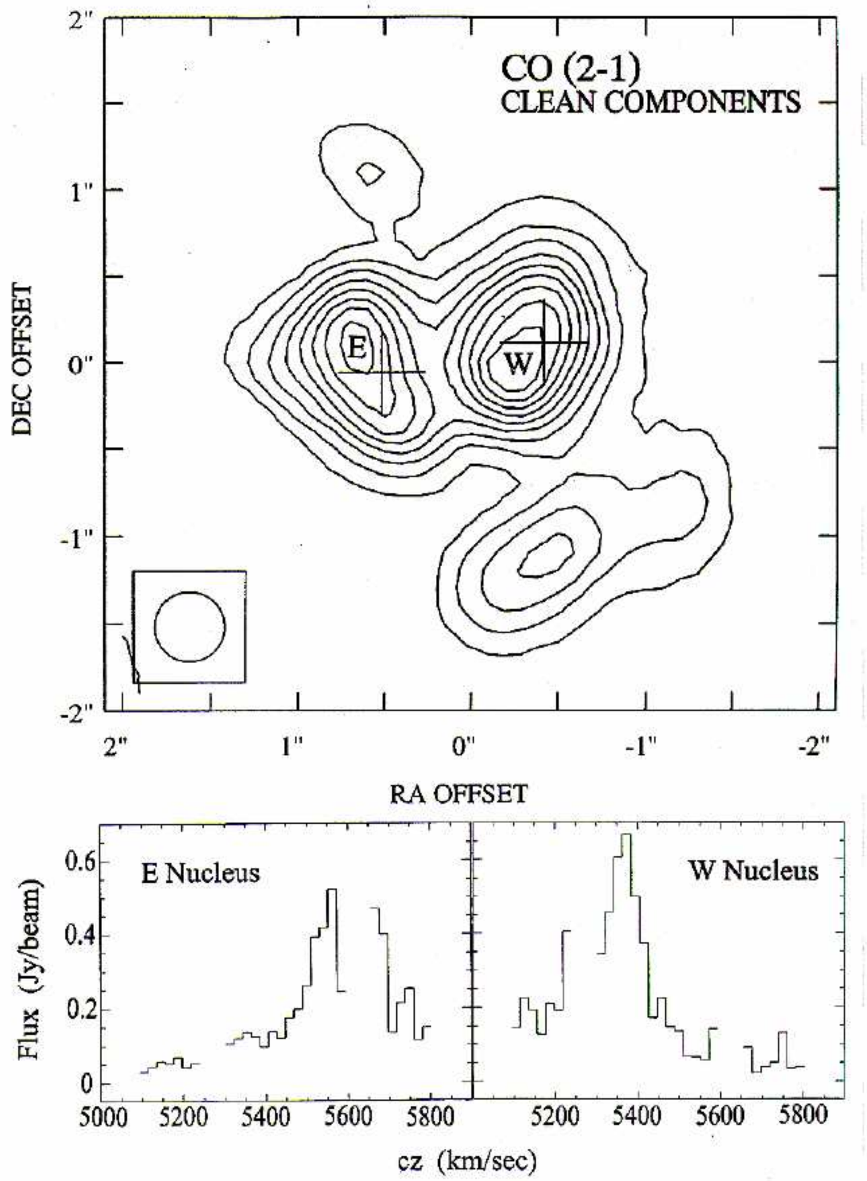


Figure 7.3. Integrated intensity map (top) for the brightest CO $2 \rightarrow 1$ "clean" components peak at the positions of the near-infrared nuclei, which are marked with the "+" symbols. In the bottom panels the CO $2 \rightarrow 1$ spectra at each of the peaks are shown. The gaps in the spectra correspond to the missing frequencies not covered by the spectrometers. The mean velocities derived from the "clean" components at the western and the eastern peaks are ~ 5300 and ~ 5600 km s^{-1}, respectively. Figure taken from Scoville, Yun & Bryant (1997).

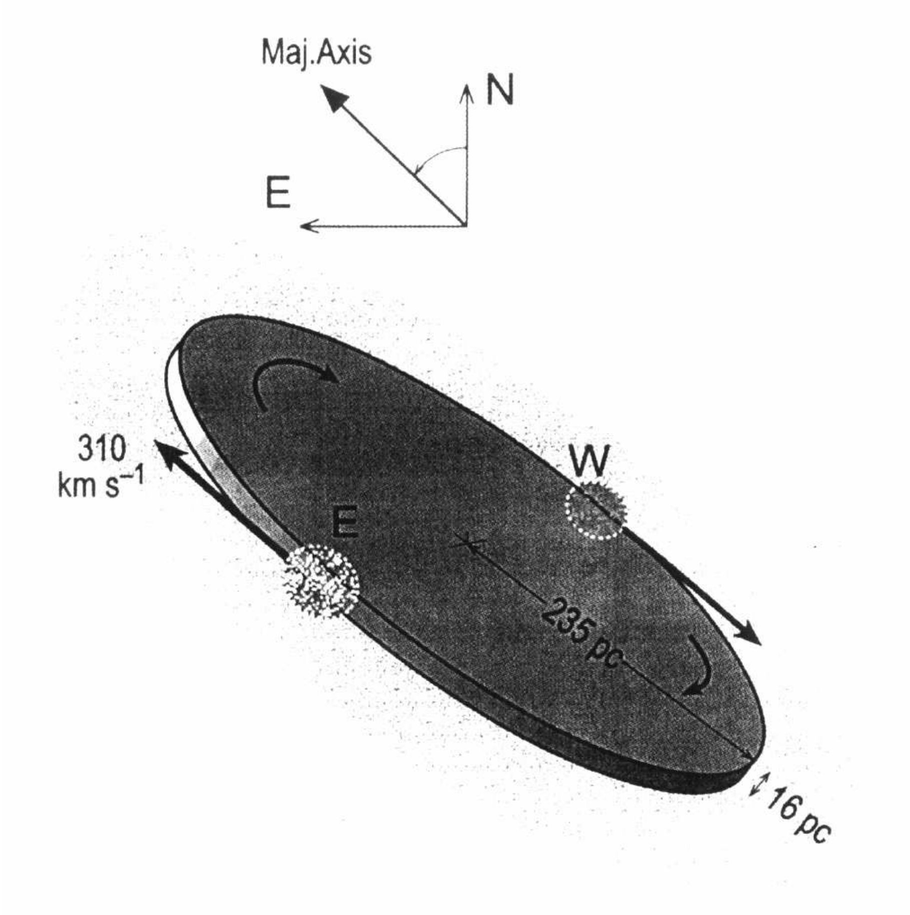


Figure 7.4. Schematic of the nucleus of Arp 220 showing the central molecular gas disk with the double infrared/radio nuclei orbiting at the outer edge of the nuclear gas disk. Based on the elongation of the molecular disk, the adopted major axis is at P.A.= 45°, and the disk is inclined at approximately 40° to the line of sight. The infrared nuclei then lie along a line between the major and minor axes of the gas disk (69° from the major axis on the disk plane). Approximately two thirds of the total molecular emission arises from this disk, the remaining from a more extended disk with approximately 10 times lower surface density. Figure taken from Scoville et al. (1997).

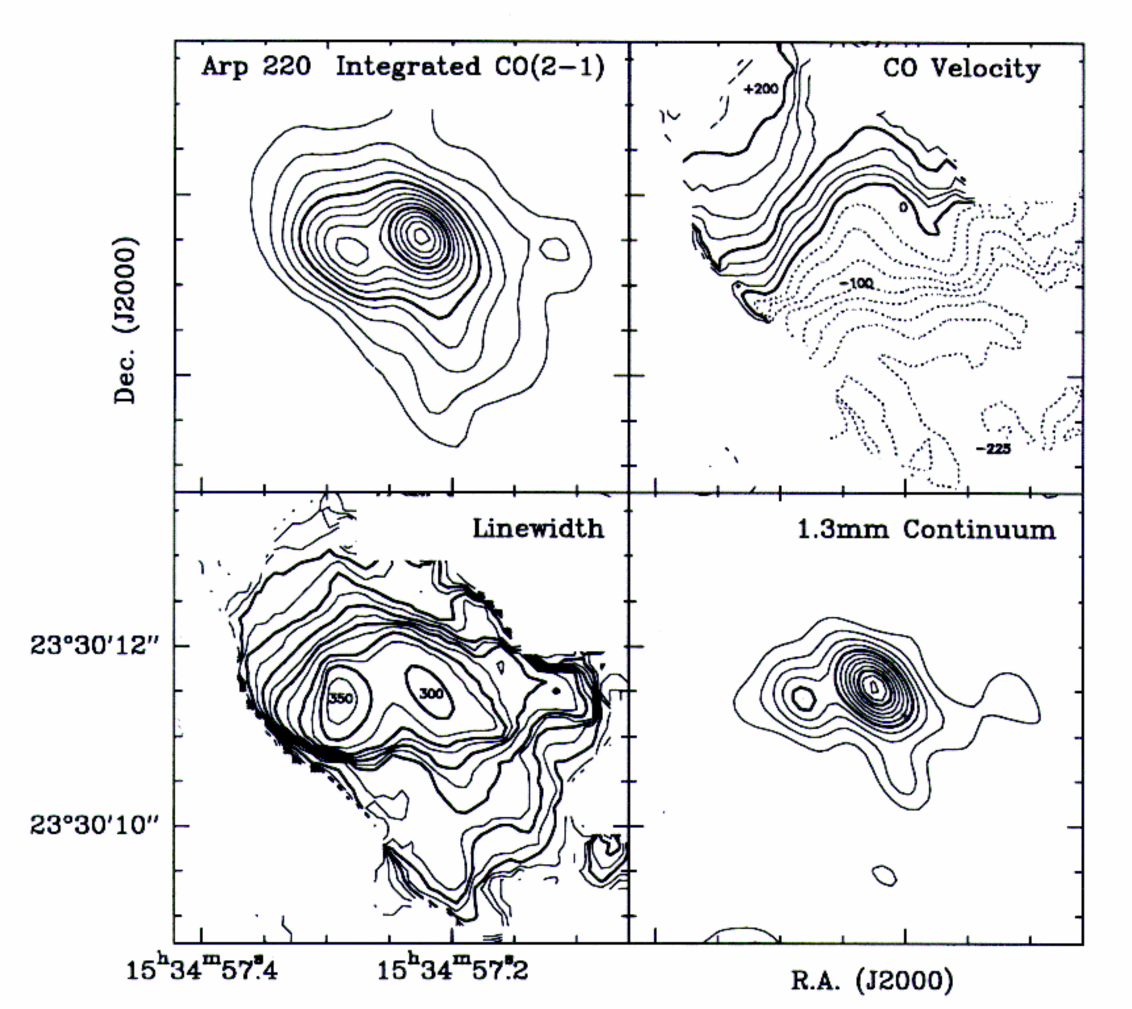


Figure 7.5. i) CO 2 → 1 intensity integrated from −320 to +300 km s^{-1} (relative to cz_{lsr} = 5450 km s^{-1}), linear contours in units of 11.4 Jy beam^{-1} km s^{-1}. ii) CO velocity from −225 to +225 km s^{-1}, in steps of 25 km s^{-1}. iii) CO line width from 25 to 350 km s^{-1}, in steps of 25 km s^{-1}. iv) 1.3 mm continuum: 1 to 12 by 1, in units of 5.7 Jy beam^{-1}. Figure taken from Downes and Solomon (1998).

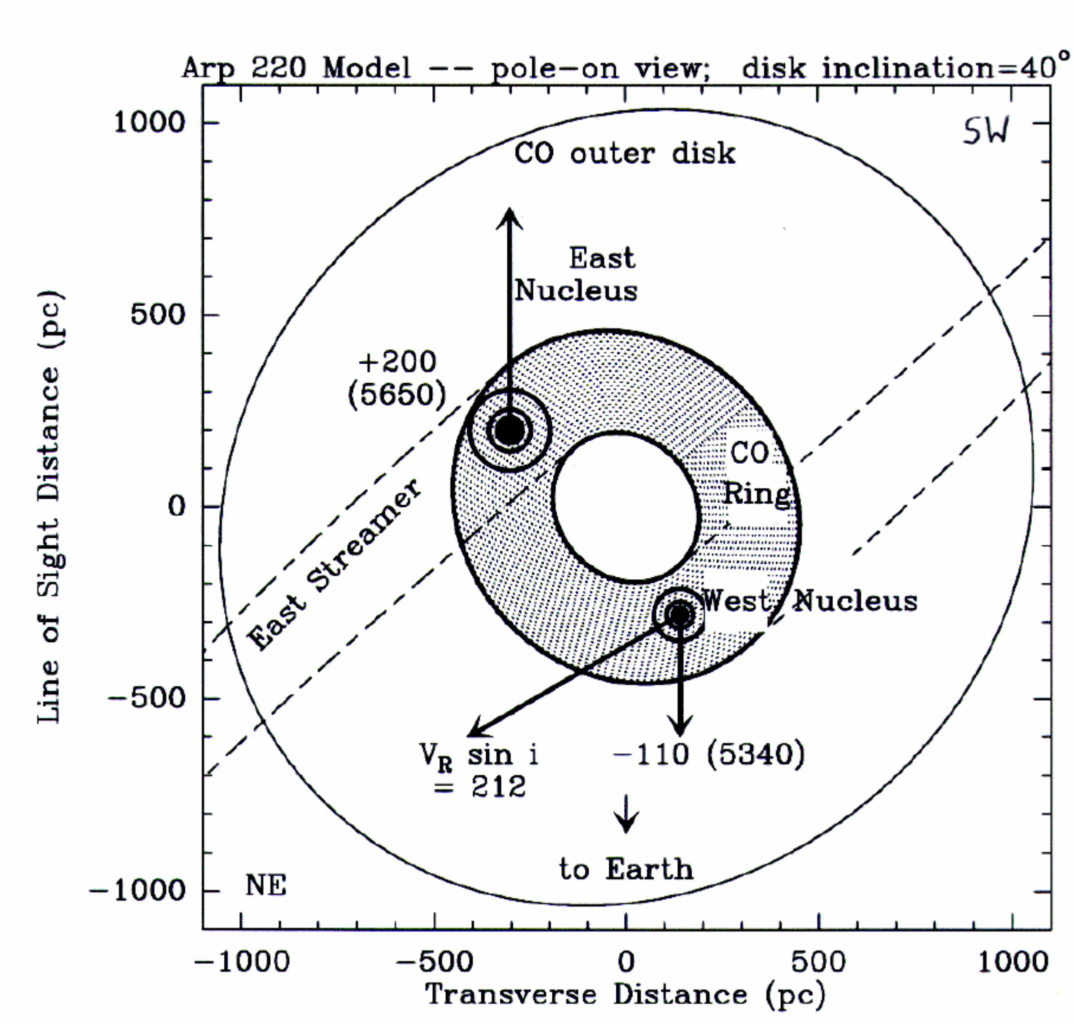


Figure 7.6. Model of the eastern and western nuclei in the molecular disk. The rotation velocity is 330 km s^{-1} and the view is pole-on. The arrows indicate the observed radial components along the line of sight, at a disk inclination of 40° from face-on. Velocities are relative to a systemic velocity of $cz_{lsr} = 5450$ km s^{-1}. Labels are km s^{-1}, values in parentheses are cz_{lsr}. Figure taken from Downes and Solomon (1998).

in contrast to SYB they do not assume that the molecular gas contains basically all of the dynamic mass, but allow stars to contribute. Their prediction of the disk thickness is 80 pc, significantly more than the 16 pc predicted by SYB. Therefore the mean density in the disk drops to 900 cm^{-3}. D&S predict gas masses of $0.6 \times 10^9 M_\odot$ for the western nucleus, $1.1 \times 10^9 M_\odot$ for the eastern nucleus and $2.0 \times 10^9 M_\odot$ for the disk. The bulge mass within a radius of 480 pc is $8 \times 10^9 M_\odot$ showing that stars contribute significantly (50%) to the dynamic mass. Finally, in contrast to SYB, D&S conclude that the luminosity of Arp 220 is produced by a starburst because firstly there is no obvious AGN, secondly given the high extinction enough ionising photons are observed at near-infrared wavelengths to justify the starburst model and thirdly there is enough dense molecular gas to fuel a starburst.

7.2 Observations

On 9th May 1997 we observed the CO $3 \rightarrow 2$ transition and continuum emission at 342.5 GHz of Arp 220 with the JCMT-CSO interferometer. In order to monitor the gain of the system these observations were interleaved with those of the quasar 3C345. We observed the radio galaxy 3C273 to give an absolute calibration. The weather was fair with 1.4 mm of precipitable water vapour. Fig. 7.7 shows the uv-track of the CO $3 \rightarrow 2$ data, giving a fringe spacing between $1\farcs1$ and $5\farcs9$. Unfortunately, we did not obtain any good data from the water vapour monitors during these observations (§6.1.3).

7.3 Basic data reduction

The Arp 220 and 3C345 data were reduced in the same way. First the data were divided by a bandpass created from observations of 3C273. Then ten 10-second integrations in a data file were vector averaged to a 100-second integration. The relative gain of the interferometer was calibrated out by applying the gain curve of the calibrator 3C345 to the Arp 220 data. The rest of this section explains these reduction processes in detail.

The signal measured by an interferometer can be thought of as the vector sum of a noise vector and a vector from the astronomical source (see Fig. 2.2). The angle of the signal vector rotates as the source moves through the fringes. Compensating for the path difference (the

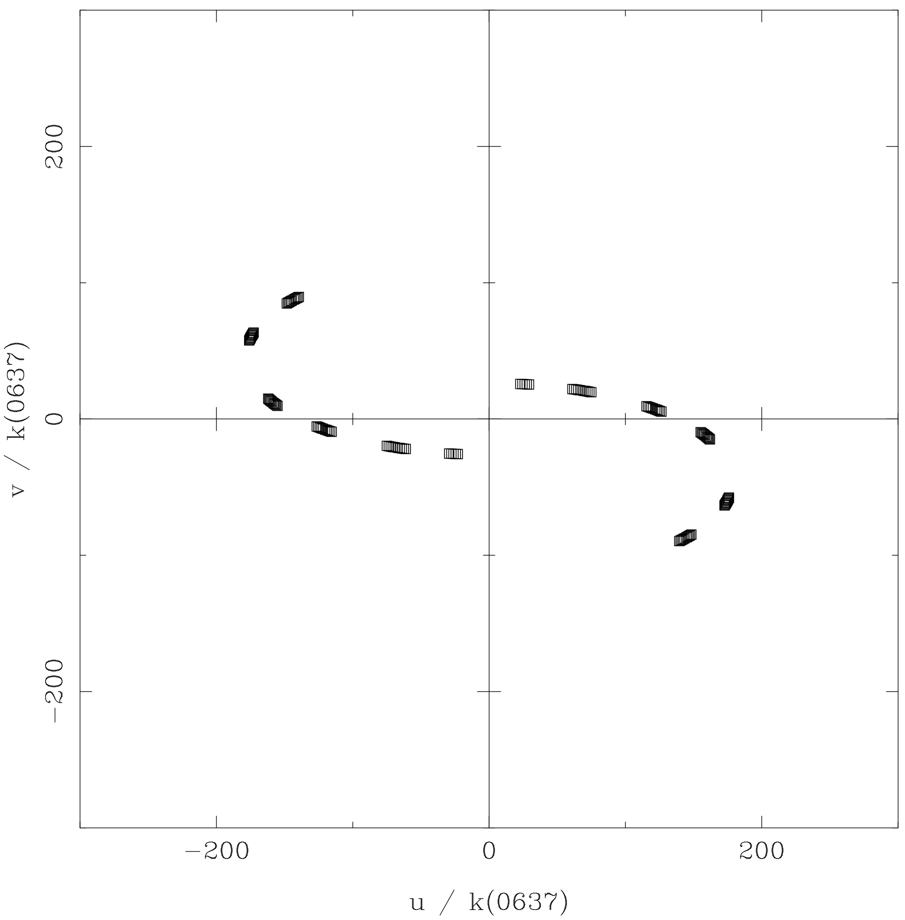


Figure 7.7. (u,v) track of CO $3 \rightarrow 2$ observation of Arp 220. (The units on the u and v axes are in kilo wavelengths.)

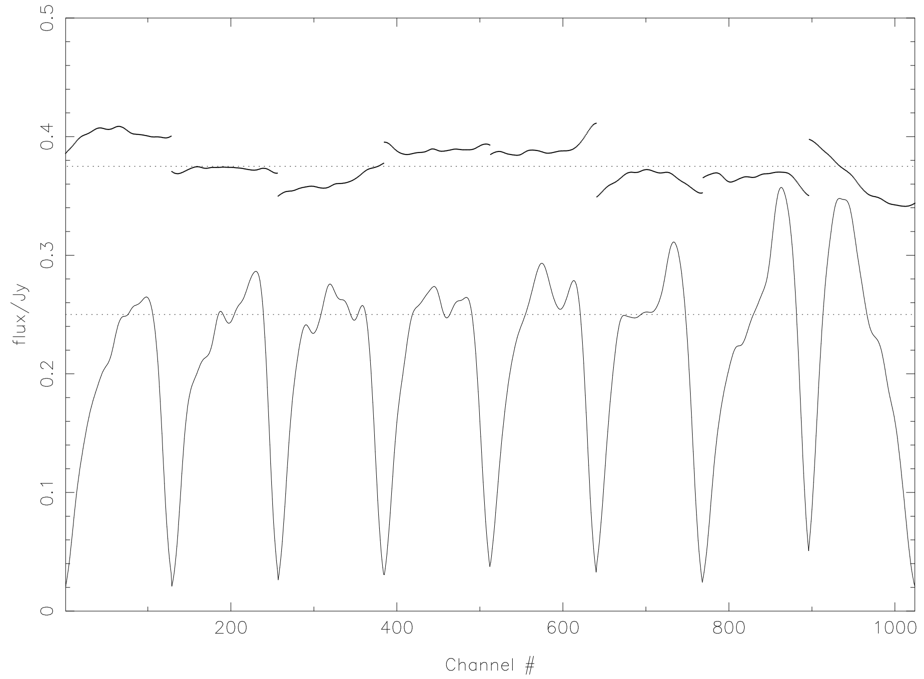


Figure 7.8. The smoothed spectrum of the quasar 3C273 used for bandpass calibration. The solid lines in the top of this spectrum represents the phase. The top of the box marks 180° phase, the first dotted line 0°, the lower dotted line −180°. The solid line in the lower part of the diagram represents the amplitude. The units for the amplitude are marked on the y-axis.

distance between the source and JCMT minus the distance between source and CSO) can be pictured as counteracting this rotation. The noise component will have random phase and mostly cancel out over the integration. The signal is then split into 1024 frequency channels and its amplitude and phase are measured in the DAS (Dutch Autocorrelation Spectrometer) and integrated for 10 seconds. Each 10-second data file therefore contains 1024 phases and amplitudes, as well as information such as observing frequency and time. Lay (1994) gives a detailed account of the JCMT-CSO interferometer, data taking and data reduction.

7.3.1 Passband calibration

The data of a point source with a flat frequency spectrum should have constant phase and amplitude over the 1024 channels, i.e. the observed spectrum should also be flat. However,

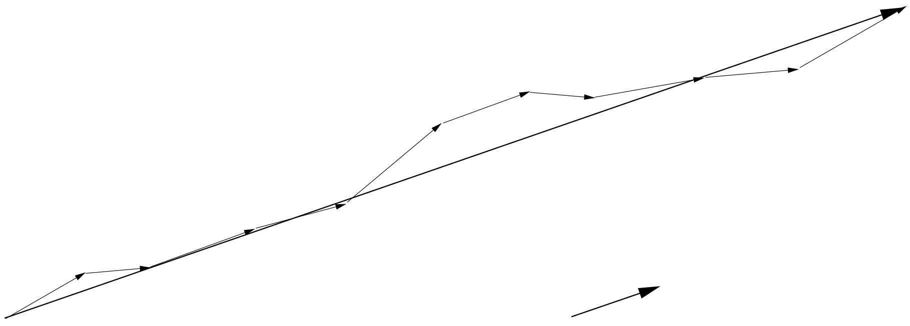


Figure 7.9. In the data reduction ten 10-second integrations were vector averaged to 100-second integrations.

since the gain in the DAS is a function of frequency this is not the case (Fig.7.8). From the difference between the expected flat spectrum and the actually measured spectrum one can easily deduce the gain of the DAS as a function of frequency. Every night a bright quasar is observed, whose spectrum measures the response of the DAS. Fig. 7.8 shows such a spectrum or passband obtained by observing the quasar 3C273. The spectrum has been smoothed by a Gaussian of 21 channels full width half maximum in order to improve the signal to noise. The gaps in the spectrum are an artifact of the DAS. All observations taken during that night are divided by the passband. (Division of complex numbers means, of course, that the observed amplitude is divided by the passband's amplitude, but the phase of the passband is subtracted from the measured phase.) This division is called passband calibration as any instrumental difference between the channels is removed.

7.3.2 Vector averaging

During the observations of Arp 220 the interferometer integrated over 10 seconds on source and wrote the result to a data file. A data file contains ten cycles of these 10-second integrations. In order to increase the signal to noise ratio the vector average of the ten 10-second integrations was taken, i.e. the sine and cosine components of the voltage vectors were added and divided by 10 (Fig. 7.9).

In May 1997 we took ten such 100-second data files observing Arp 220, and then took two

100-second data files of 3C345. In Fig. 7.7 the uv-plot of the CO $3 \rightarrow 2$ data shows 6 clusters of such 100-second data-files. During the big gaps between the data files the calibrator as well as the continuum emission of Arp 220 were observed.

7.3.3 Gain calibration

The gain of the interferometer changes slightly with time. However, the amplitude and phase of the calibrator are a measure of such changes. Therefore the vector averaged amplitudes of 3C345 were plotted versus time and a gain curve was drawn connecting the data points. We calculated the average amplitude and normalised the gain curve by dividing it by the average value. The Arp 220 data were then divided by this normalised gain curve, so that changes in the response of the interferometer were calibrated out. [1]

The amplitudes are measured in antenna temperature. To convert the antenna temperatures into Jansky the flux of 3C273 was measured with the interferometer as well as with the JCMT alone. By comparing single dish measurements of Mars, which has a well known flux, with those of 3C273 a flux of 12 ± 0.5 Jy was derived for 3C273. 3C273 has an antenna temperature of 0.183 K (0.194 K at 342.5 GHz) when observed with the interferometer, therefore the conversion factor is 65.6 ± 3.1 Jy/K for the line data and 61.9 ± 2.7 Jy/K for the continuum data. The error quoted is the statistical error from the measurements. However, there will also be an additional systematic error due to gain variations, uncertainties in the planet flux, etc.. The total error in the calibration is estimated to lie around $\pm 20\%$ (Lay, 1994).

7.4 Continuum data

7.4.1 Amplitude as a function of hour angle

The continuum flux of Arp 220 was measured at 342.5 GHz. After reducing the continuum data of Arp 220 as described above, the amplitude was plotted as a function of hour angle (squares in Fig. 7.10). The change in amplitude contains information about the source extent and whether the emission originates from a single source or multiple components. An interferometer detects a constant amplitude when observing a point source. For a single extended

[1] If the 3C345 data were divided by the gain curve, the resulting amplitudes would all take the average value and the amplitude of the calibrator would be constant as expected for a perfect interferometer.

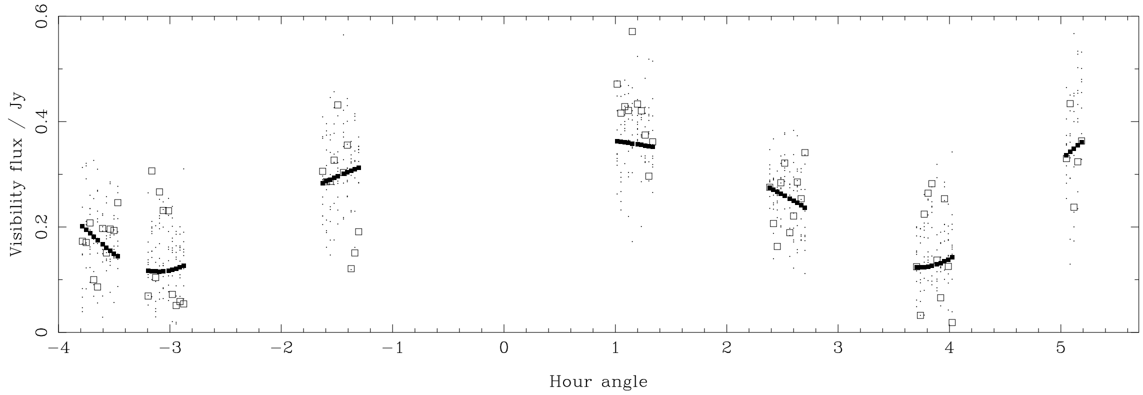


Figure 7.10. Continuum data of Arp 220. The data are represented by squares, the model is the thick line and the dots are simulated data given the model.

source the amplitude decreases as the projected baseline increases and some of the flux is resolved out. If there are two equal point sources the detected amplitude can even drop to zero, when the source separation is a multiple of half the width of the fringe spacing.

7.4.2 Maximum likelihood method of finding the best model

It is possible to calculate the expected amplitude for different models of the emitting source and find the model that is most likely to represent the data (maximum likelihood method). For Arp 220 we calculated the amplitude an interferometer would detect when observing an uneven binary, where each source is assumed elliptical, with a Gaussian brightness distribution along each axis (Lay et al., 1995). Our uneven binary is characterised by 10 parameters: the separation of the two sources and the position angle (as measured east of north), as well as the flux, minor axis, major axis and the position angle of the major axis of each source. The expected amplitude as a function of hour angle was calculated for different models, i.e. different parameter sets. Since the noise of the data is known the probability of each data point given the model can be calculated. In the case of high signal to noise the probability is described by a Gaussian distribution. However, in our case the noise is comparable to the signal and therefore the Rice distribution describes our data better. The product of the probability of each of the data points given the model gives the likelihood of the model. The model with the greatest product of probabilities is the most likely.

In Fig. 7.10 the squares represent the data, the thick line the calculated amplitude and the

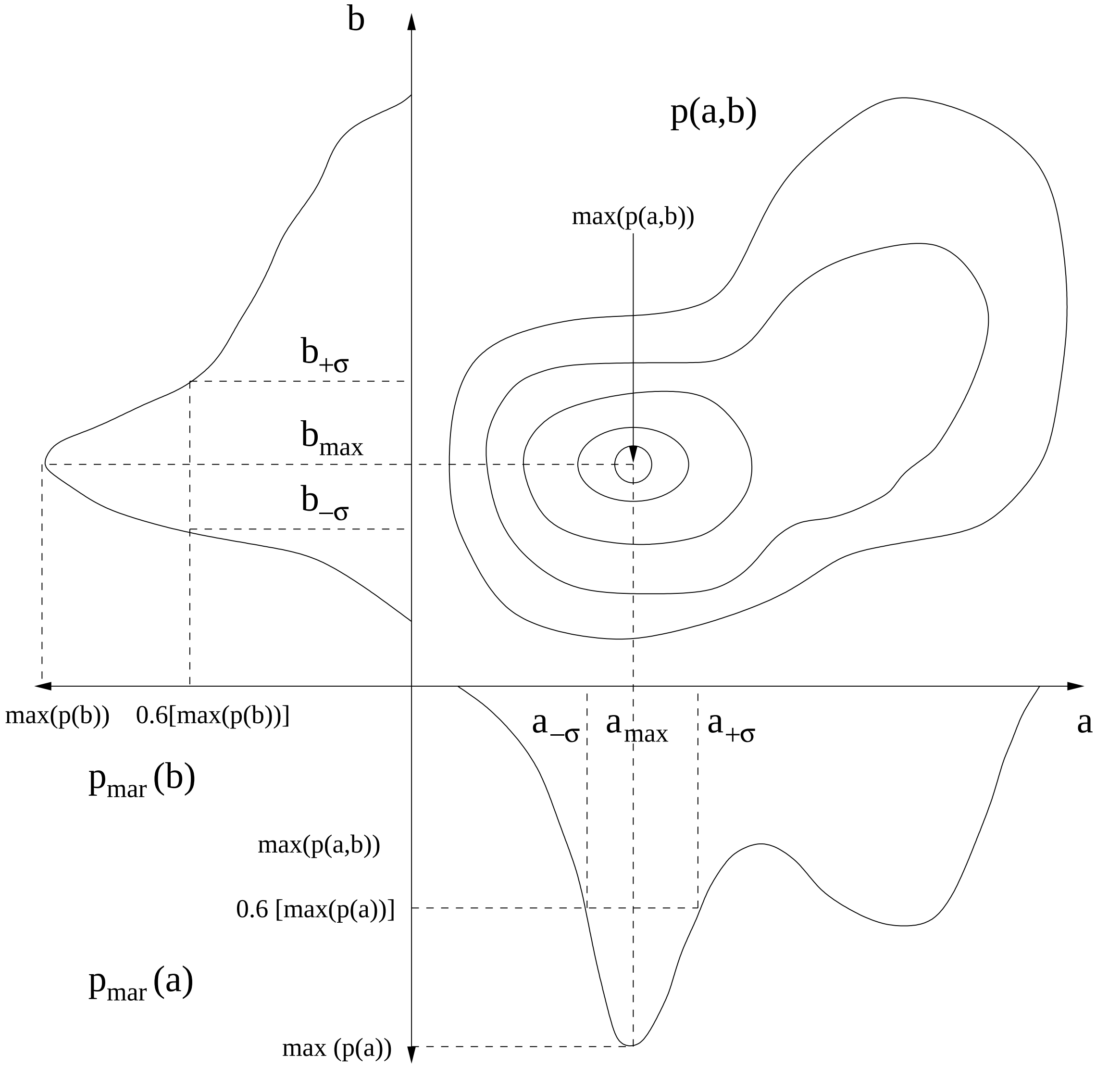


Figure 7.11. The contour plot represents the probability $p(a, b)$ of two parameters a and b. The marginal probabilities $p_{mar}(a)$ and $p_{mar}(b)$ are obtained by integrating $p(a, b)$ over b and a, respectively. We defined the Bayesian 1σ error of a_{max} as the difference between a_{max} and a_σ, where the marginal probability $p_{mar}(a_{max})$ has dropped to $exp(-1/2)$ of its maximum value, $p_{mar}(a_\sigma) = exp(-1/2)p_{mar}(a_{max})$. This error is independent of b.

dots the simulated data given the model and the noise. Note that in the case of low signal to noise the best model does not go through the middle of the data points, but lies lower. This is due to the Rice distribution being asymmetric around the amplitude of the signal. The cause of this asymmetry becomes clear in the extreme case of pure noise without any signal. The interferometer will always measure a positive amplitude. Therefore the distribution of measured amplitudes cannot be symmetric around zero, which is the amplitude of the signal in this case. The measured amplitudes should, however, fall in the middle of the simulated data.

The Bayesian error of each of the ten parameters for the most likely model was calculated in the following way: First the probabilities were normalised by dividing them by the maximum probability. Then the probabilities of all models with the parameter a having a certain value a_1 were summed. This sum is the marginal probability of a_1, which is independent of the other nine parameters. This process was repeated for different values a_2, a_3 etc. of a. [2] Fig. 7.11 is an example for the two dimensional case. A contour map of the probability $p(a,b)$ of a and b is plotted in the upper right corner. In the upper left graph the marginal probability $p_{mar}(b)$ of b integrated over all a is plotted. The lower right graph shows the marginal probability $p_{mar}(a)$. Usually, but not necessarily, the value a_{max} with the highest marginal probability is also the value found for the most likely model. In a small region around a_{max} the probability distribution can be approximated by a Gaussian. We defined the one sigma Bayesian error of a_{max} as the distance between a_{max} and a_σ, where the marginal probability drops to $exp(-1/2)$ of its maximum value. [3] The error on each parameter gives an indication on how well the measured data determine each parameter in a given model, but the error contains no information about the validity of the model.

From the maximum likelihood method we can obviously only determine which of the models *tried* is the most likely. It is possible that an entirely different model we have not considered fits the data much better. For this reason it is very helpful to have maps of the source at other wavelengths (e.g. around 230 GHz), which show the morphology of the source

[2] The probability $p(a,b,c,..,j)$ of the ten dimensional parameter space was integrated over all parameters but a to give the one dimensional marginal probability of a: $p_{mar}(a) = \int \cdot \cdot \int\int p(a,b,c,\ldots)\, db\, dc \ldots dj$.

[3] More precisely a_σ should define the boundaries in which 68% of the marginal probability lie. In the case where the marginal probability can be approximated by a Gaussian, a_σ will have the value $exp(-1/2)p_{mar}(a_{max})$. Since in our case the marginal distribution appears Gaussian we used the latter method to determine the error.

and make the choice of models more obvious.

7.4.3 Models for continuum emission from Arp220

Since the continuum map at 229.4 GHz as observed by SYB shows an elliptical source, models with a single source of a diameter between $1''$ and $2''$ and a position angle of around $90°$ were tried. However none of these models gave a good fit to the data. (Such models cannot reproduce the minimum at HA $= -3$ and $+4$ hours.)

Therefore an uneven binary models was fitted to the data. The most likely model is shown in Fig. 7.10, its parameters are listed in Tab. 7.1. The errors are one sigma Bayesian errors of the marginal probability distribution as discussed above. For comparison the parameters observed (not fitted) at other wavelengths are also listed in Tab. 7.1. Note that without using the phase information we can only determine the separation and the position angle of the two nuclei, but we cannot tell which of the two nuclei lies in the west. Since the western nucleus has twice the flux of the eastern nucleus at 229 GHz, it was assumed that in the 342 GHz continuum the brighter nucleus also lies in the west.

7.4.4 Discussion of the model which fits the continuum observation best

Since the model fits the data well and since it at least partly agrees with observations at other wavelengths we assume that it gives a good representation of the continuum emission of Arp 220. Therefore we will discuss the results in more detail.

1. Total flux

   According to our model the total continuum flux of Arp 220 is 450 mJy $\pm$ 21 mJy. Most of the continuum flux is expected to come from dust emission, whose spectrum increases with frequency to the power of $2+\beta$. β lies between 0 and 2 depending on the size of the dust grains and their radiation properties (Hildebrand, 1983). From SYB's flux measurement at 110.2 and 229.4 GHz we find $\beta = 0.3$. From this we predict a flux of 480 mJy at 342.5 GHz, which agrees well with our data. Assuming that the synchrotron radiation does not contribute significantly to the flux at 229 GHz, the constant β indicates that it does not contribute significantly to the flux at 110.2 GHz either. Instead the continuum observed is grey body emission from dust and all the

	These data 342.5 GHz	SYB 229.4 GHz	D&S 229 GHz	Scoville 2.2μm	Baan 4.83 GHz
Morphology	binary	ellipse	binary	three sources	binary
Total flux	450±21 mJy	192±20 mJy	175 mJy	—	154 Jy
Separation	$1\farcs0 \pm 0\farcs04$	—	$0\farcs8$	$1\farcs13$ (NE-W)	$1\farcs06$
				$1\farcs05$ (SE-W)	
Position	$80^\circ \pm 3^\circ$	—	100°	87° (NE-W)	98°
angle				108° (SE-W)	
Eastern n.:					
flux	160±15 mJy	—	34.2 mJy	—	62.8 Jy
size	$< 0\farcs1$	—	$< 0\farcs1$	$0\farcs34 \times 0\farcs26$ (NE)	$0\farcs29 \times 0\farcs20$
				$0\farcs26 \times 0\farcs23$ (SE)	
Western n.:					
flux	290±15 mJy	—	68.4 mJy	—	91.0 Jy
size	$(0\farcs2 \pm 0\farcs2)^2$	—	$0\farcs3 \times 0\farcs3$	$0\farcs49 \times 0\farcs22$	$0\farcs40 \times 0\farcs30$

Table 7.1. Parameters of the binary model that fits the 342.5 GHz continuum data best. For comparison measurements at millimetre wavelengths (Scoville et al., 1997 (SYB) and Downes and Solomon, 1998 (D&S)), in the infrared (Scoville et al., 1998) and at radio wavelengths (Baan and Haschick, 1995) are also listed. (The radio fluxes quoted are from the naturally weighted 4.83 GHz continuum maps.)

measurements lie on the raising slope of the grey body emission[4] before the spectrum turns over. (For a dust temperature of 80 K as expected for Arp 220 (D&S) the *black* body turnover lies at $\sim$ 5000 GHz (Rohlfs, 1986), well above our highest frequency measurement.)

2. Binary separation and position angle
   The model predicts the binary to be separated by $1''.0 \pm 0''.04$ approximately east-west at position angle $80 \pm 3°$. In D&S's continuum map the emission peaks are slightly closer together and at a larger position angle. Our data do however agree reasonably well with the position of the formaldehyde emission, whose peaks are separated by $1''.1$ at 79°. Formaldehyde is associated with star bursting gas (Baan and Haschick, 1995), which will also contain hot dust and therefore emit continuum radiation at submillimetre wavelengths. It is therefore not surprising that the submillimetre flux peaks at the same location as the formaldehyde, though it is not clear why the continuum at 230 GHz originates from a different region.

3. Source fluxes
   The fluxes of the sources are 290 ± 15 mJy and 160 ± 15 mJy, giving a ratio of 1.8. The peak fluxes in D&S's continuum map are 68.4 and 34.2 mJy/beam giving a similar ratio. The constant ratio over wavelength suggests that the continuum flux is created by similar physical processes in both nuclei. Since the continuum flux is proportional to the dust mass multiplied by the temperature and the temperature should not vary by more than a factor of 3 between the two nuclei, the masses of the two nuclei must be of the same order of magnitude.

4. Source sizes
   According to our model the continuum emission of the western source originates in a region of $0''.2 \pm 0''.2$ FWHM, the eastern source is practically a point source. D&S's continuum maps at 230 GHz suggest similar source sizes. The 2.2μm emission, which

[4]The low value of $\beta = 0.3$ suggests that the dust grains are much bigger than the wavelength observed at and that the continuum emission nearly has the black body spectrum, for which β would be zero. This is a bit surprising as even dense T-Tauri star disks, which contain large dust grains, have $\beta \approx 1$ (Richer, 1998). Before drawing too many conclusions, however, it should be carefully checked that the continuum emission detected at different wavelengths with different interferometer beams is coming from the same region.

traces dust enshrouded young stars comes from slightly larger region (Scoville et al., 1998).

In summary, the continuum emission at 342.5 GHz does not come from a single extended component as observed by SYB at 230 GHz. Instead a binary model fits the data well. From the flux ratios we conclude that the emission is grey body emission from dust. The dust is probably associated with two starburst regions of $\sim$ 70 pc ($0\farcs2$) diameter at the position of the nuclei.

7.5 CO 3-2 line data

The CO $3 \rightarrow 2$ line data consist of amplitude and phase data as a function of frequency and hour angle, $A(\nu,t)$ and $\phi(\nu,t)$. In this section we will first look at spectra $A(\nu)$. Next we discuss the best model model fitted to the amplitude data $A(t)$, as done with the continuum data. A spatial distance between the red and blue component of the line can be determined by analysing the phase changes as a function of hour angle, $\phi(\nu,t)$. Finally the data are divided into a blue and a red component and the most likely model is found for each component, $A(\nu,t)$.

7.5.1 Spectra

Data reduction

The data were reduced according to the basic reduction process described in §7.3. In order to increase the signal to noise in the spectra, each set of ten 100-second averages, which were taken approximately at the same time, was averaged into a 1000-second file. (The sixth set only contains five 100-second samples.) Over time scales longer than 100 seconds the average phase might change due to instrumental phase shifts or the atmosphere. Therefore the average phase over the whole frequency range of each 100-second sample was determined. This average phase was then subtracted from the phases in each of the individual channels. The resultant data were then averaged as described above, i.e. the arithmetic means of the sine and cosine components were determined. The frequencies were converted into velocities

v_{cz} using

$$v_{cz} = \frac{\nu_{rest} - \nu_{obs}}{\nu_{obs}} c - v_{rad}, \tag{7.1}$$

where ν_{rest} is the frequency of the CO $3 \rightarrow 2$ transition for CO at rest with respect to the observer

ν_{obs} is the frequency observed

c is the speed of light

v_{rad} is the radial velocity of the interferometer with respect to the local standard of rest (due to the motion of the earth etc.).

Discussion of spectra

The six spectra are plotted in Fig. 7.12 in order of increasing hour angle. Fig. 7.7 shows the position of the spectra in the (u,v) plot. The first spectra was taken at an hour angle of -2.3 and corresponds to the data in the upper left side in the (u,v) plot at (-140kλ, 90kλ). Depending on the position in the (u,v) plot and therefore on the resolution, the spectra look different. All of them, however, have a line width of approximately 400 km s^{-1} as well as a double peaked structure suggesting that two components at different velocity contribute to each spectrum. The last spectrum was taken with the interferometer having the largest fringe spacing, 6$''$. It contains significantly more flux in both the blue and the red component than the spectra taken at shorter fringe spacings. This shows that some of the line emission comes from an extended source, which is resolved out at smaller fringe spacings. This emission could originate from the rotating molecular disk and the eastern streamer suggested in the models of SYB and D&S (Fig. 7.4 & 7.6).

Tab. 7.2 lists the characteristics of the CO $3 \rightarrow 2$ spectrum, where the fringe spacing was 6$''$, as well as those of SYB's CO $2 \rightarrow 1$ and CO $1 \rightarrow 0$ spectra, which include all data with a fringe spacing greater than 5$''$. Some parameters measured by D&S are also included in the table.

1. Line width
   The CO $3 \rightarrow 2$ spectrum has full width half maximum of about 420 km s^{-1} (after subtraction of the continuum). This compares well with the line width of the lower transitions of CO.

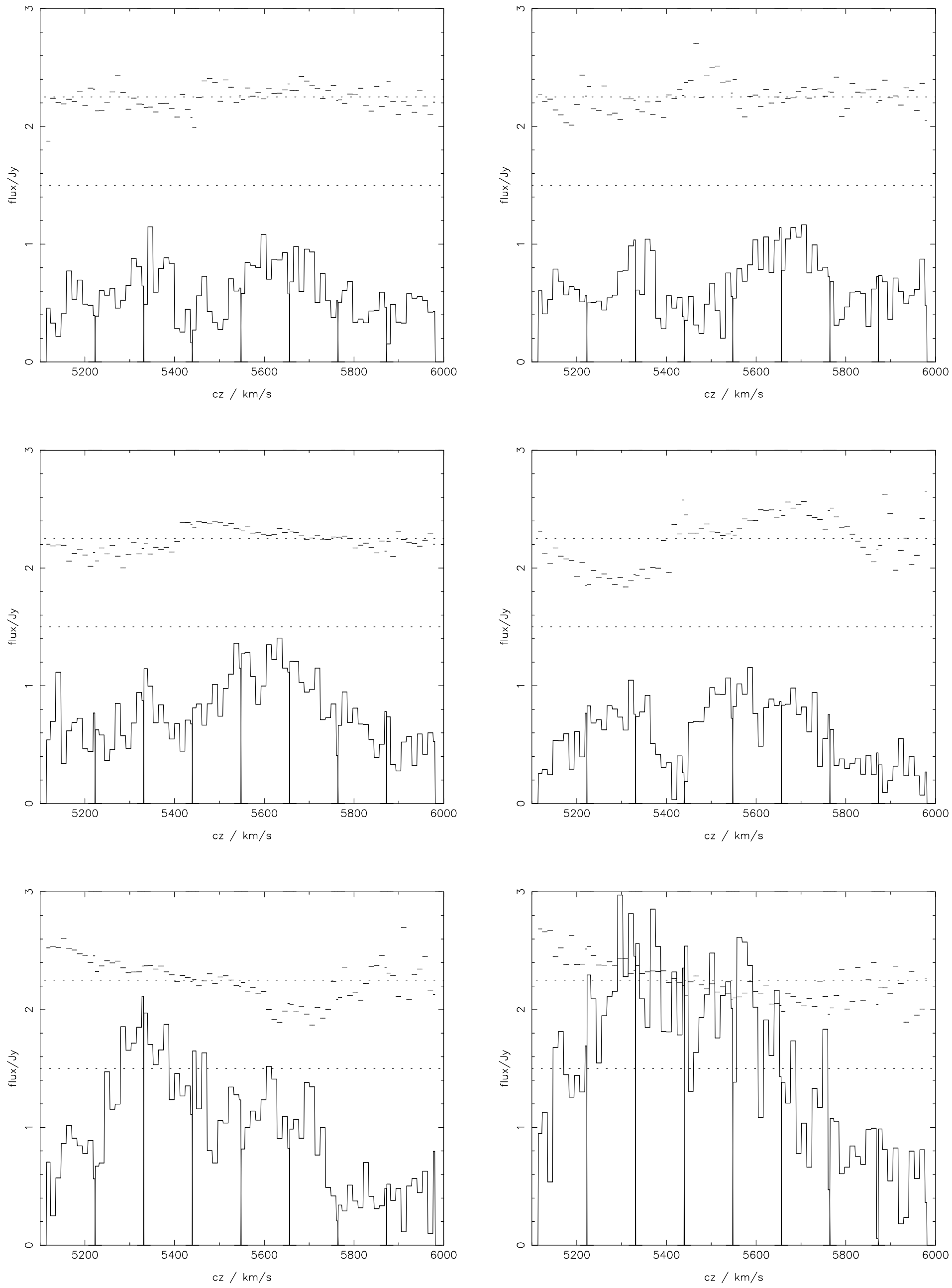


Figure 7.12. CO 3 → 2 spectra of Arp 220. The first spectrum on the top left was taken at −2.3 h (HA) at (u,v) coordinates (−140 kλ,90 kλ). Top right HA= −0.8 h (−170 kλ,60 kλ), middle left HA=2 h (−160 kλ,10 kλ), middle right HA=3.2 h (−125 kλ,−10 kλ), bottom left HA=4.5 h (−70 kλ,−20 kλ), bottom right HA=5.5 h (−25 kλ,−25 kλ) (as shown in uv-track in Fig. 7.7)

	These data CO $3 \rightarrow 2$	SYB CO $2 \rightarrow 1$	SYB CO $1 \rightarrow 0$	D&S CO $2 \rightarrow 1$	D&S CO $1 \rightarrow 0$
Shape	double peaked	double p.	double p.	—	—
v of peak [km]	5350, 5550	5300, 5550	5350, 5550	—	—
Peak flux [Jy]	2.65, 2.25	2.2 , 1.8	0.88, 0.77	—	—
Δv_{FWHM} [km s^{-1}]	420±30	447	504	—	—
Integrated flux [Jy km s^{-1}]	1200±100	1071	410	1100	410

Table 7.2. Characteristics of CO $3 \rightarrow 2$ spectrum of Arp 220, taken at $6''$ fringe spacing (bottom right spectrum in Fig. 7.12).

2. Double peak
   All spectra show two peaks at around 5350 km s^{-1} and 5550 km s^{-1}.

3. Peak fluxes
   Within the errors the ratios of the peak flux of the red and blue components are similar for all CO transitions. This suggests that the physical conditions influencing the CO intensities are similar for the red and the blue-shifted gas.

4. Integrated intensity
   The integrated flux of the CO $3 \rightarrow 2$ line, 1200 Jy km s^{-1}, is similar to that of the CO $2 \rightarrow 1$ line. This yields a brightness temperature ratio of 0.5, since $T_{br} = S_\nu/(2k\Omega\nu^2)$ and the beam sizes Ω are approximately the same for both transitions. Such a low brightness temperature ratio suggests that the line is subthermally excited. [5] This means that collisions are not frequent enough to fully populate the J = 3 level of CO. Subthermally excited emission implies H_2 densities below 1000 cm^{-3} (Harrison, 1998) to 10 000 cm^{-3} (Richer, 1991), depending on the model for the collisional excitation. [6] SYB's model of a thin disk predicts densities above 2×10^4 cm^{-3}. Since D&S assume

[5] Alternatively, gas well below the excitation temperature of the CO $3 \rightarrow 2$ line of 33 K would have a similar ratio. However, infrared data suggest dust temperatures between 75 and 80 K (D&S).

[6] Large velocity gradient (LVG) models for an expanding sphere at a single temperature predict a density of 500 cm^3 from the CO $2 \rightarrow 1$ to CO $1 \rightarrow 0$ line ratio (Goldsmith et al., 1983). The morphology of Arp 220 is obviously much more complex, but it still suggests that the densities in Arp 220 probably lie below 10^3 cm^{-3}.

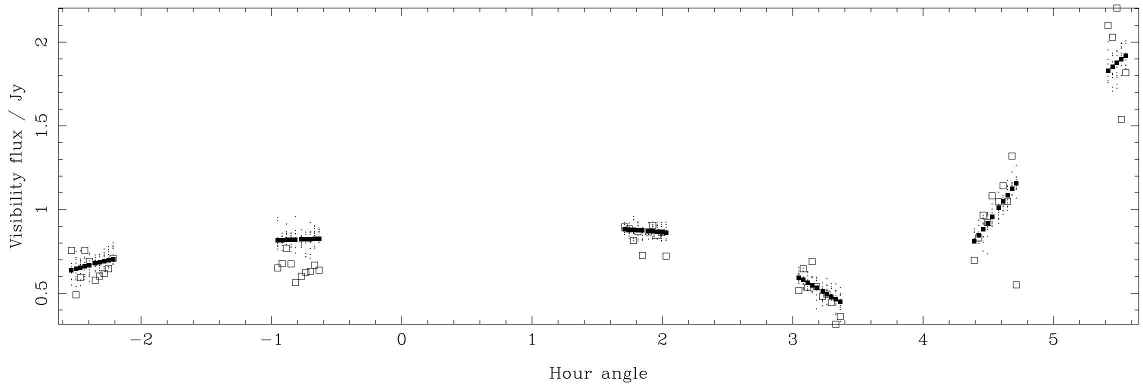


Figure 7.13. CO 3 → 2 line data of Arp 220. The data are represented by squares, the model is the thick line and the dots are simulated data given the model.

a thicker disk and that some of the dynamic mass is in stars they predict densities of 900 cm^{-3}. Our data favour their model.

7.5.2 Model for line emission deduced from amplitude data

Fig. 7.13 shows the amplitude averaged over frequency as a function of hour angle. As for the continuum data (see §7.4) we looked for the model that fits the data best. A single elliptical source did not agree well with the data, nor could two very extended sources represent the data. The best fit is obtained from an uneven binary model (thick line in fig 7.13), whose parameters are listed in table 7.3.

1. Total flux

   The total flux suggests that the CO 3 → 2 line is subthermally excited and that the molecular density therefore lies below 10^4 cm^{-3} (See §7.5.1 under integrated intensity).

2. Binary separation and position angle

   The model predicts that the two sources are separated by $1''\!.0$ in the east-west direction (P.A. = 90°). This agrees well with the separation and position angle of the emission peaks in SYB's cleaned maps. D&S quote a wider separation of $1''\!.3$. However, this value was obtained by attributing the flux measured to more than two different features. The separation of the emission peaks in D&S's integrated intensity map is only $0''\!.85$, which is more consistent with our results.

	These data	SYB	D&S
Morphology	binary	binary + disk	binary + disk
Total flux	1200±35 Jy km s^{-1}	1071 Jy km s^{-1}	1100 Jy km s^{-1}
Separation	$1''.0 \pm 0''.04$	$1''$	$1''.28$
Position angle	$90° \pm 4°$	86°	95°
Eastern nucleus:			
flux	900±50 mJy	450±330 Jy km s^{-1}	220 Jy km s^{-1} + disk
size	$(0''.7 \pm 0''.2)^2$	$0''.25$ (theory)	$0''.9 \times 0''.9$
Western nucleus:			
flux	1200±50 mJy	500±330 Jy km s^{-1}	130 Jy km s^{-1} + disk
size	$(0''.5 \pm 0''.2)^2$	$0''.25$ (theory)	$0''.31 \times 0''.28$

Table 7.3. Parameters for binary model fitting CO 3 $\rightarrow$ 2 data of Arp 220 best. For comparison CO 2 $\rightarrow$ 1 data taken by SYB (Scoville et al., 1997) and D&S (Downes and Solomon, 1998) are listed.

3. Source fluxes

   The CO 3 $\rightarrow$ 2 fluxes of the two binary components are similar. This agrees with an estimate of the CO 2 $\rightarrow$ 2 fluxes as measured by SYB. D&S predict a different ratio, but in their analysis the line flux is not attributed to the binary alone, instead ~50% of the flux comes from the disk.

4. Source sizes

   The CO 3 $\rightarrow$ 2 data suggest FWHM source sizes of $0''.5 \pm 0''.2$ and $0''.7 \pm 0''.2$. Downes measurements lie just inside the error bars, but SYB's theoretical prediction of the radii of accretion disks around the nuclei are inconsistent with our measurements. It is interesting to note that the line emission comes from a more extended region than the continuum emission. This probably means that the CO envelopes a much smaller and denser core seen in continuum emission.

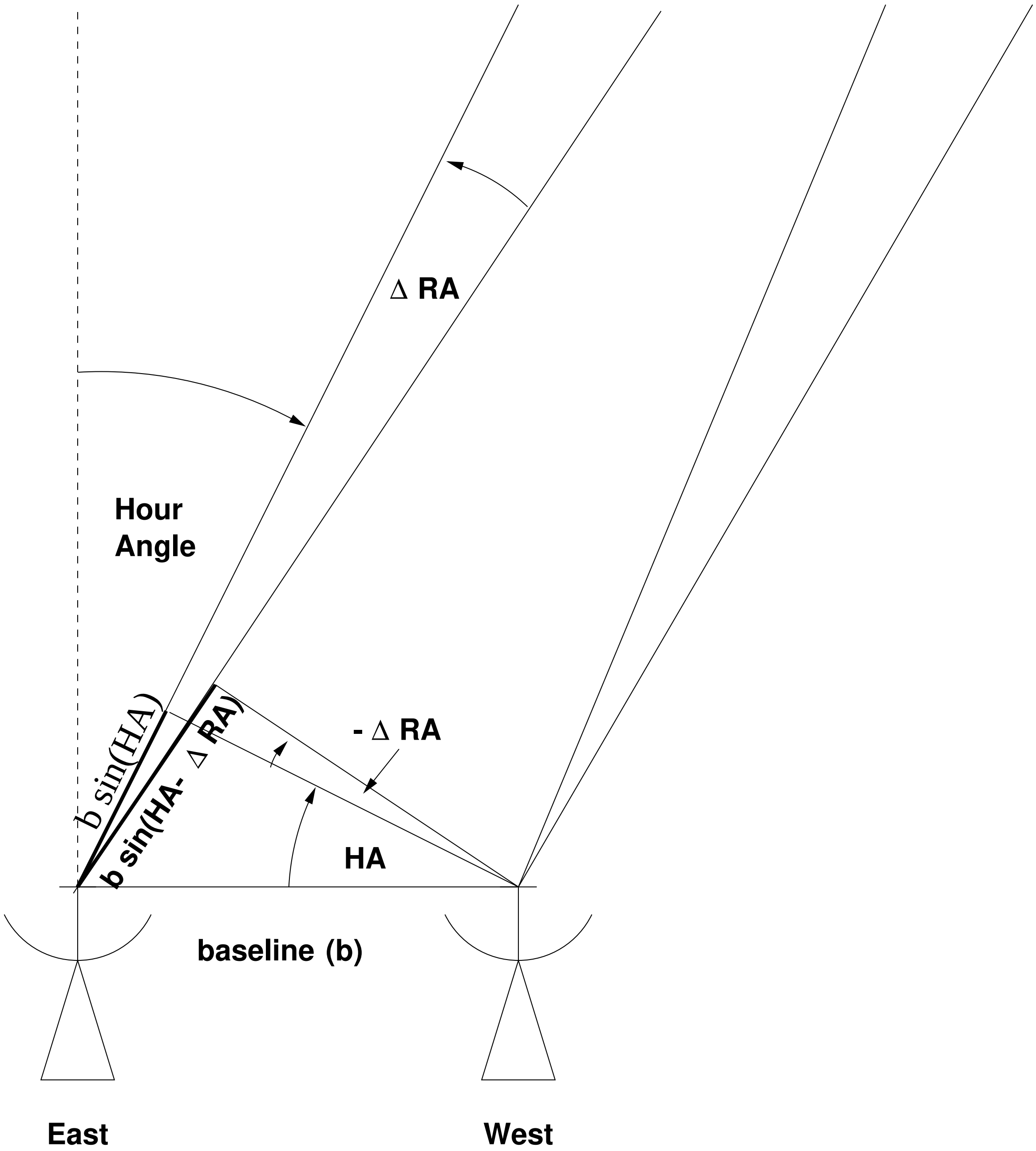


Figure 7.14. Extra optical path in an east-west interferometer for two sources at different right ascension.

7.5.3 Angular separation of nuclei deduced from phase data

Information of position offsets from phase

The interferometer measures the difference in path lengths from the source to the two antennas in terms of phase. Obviously, this phase contains information about the source's position in the sky. Unfortunately, due to instrumental noise and the atmosphere it is not possible to determine an absolute position of a source from its phase with the JCMT-CSO interferometer. However, it is possible to calculate the position offset between two sources by measuring their phase difference. As illustrated in Fig. 7.14 the phase difference between two source of different RA[7] (but both at the same declination) is

$$\begin{aligned}
\Delta\phi &= \frac{2\pi}{\lambda}[b\ \sin(HA) - b\ \sin(HA - \Delta RA)] \\
&\approx \frac{2\pi}{\lambda}[b\ \sin(HA) - b\ (\sin(HA) + \cos(HA)\ \Delta RA)] \\
&= \frac{2\pi}{\lambda} b\ \cos(HA)\ \Delta RA \\
&= 2\pi \frac{b_{proj}}{\lambda} \Delta RA. \qquad (7.2)
\end{aligned}$$

where λ is the wavelength observed at,
HA is the hour angle of the source used as phase reference,
ΔRA is the difference in right ascension between the sources,
b is the baseline length of an east-west interferometer and
b_{proj} is the projected baseline length.

According to Equ. 7.2 the phase difference plotted against time will take the shape of a cosine, whose period is 24 hours and whose amplitude is proportional to ΔRA.

In general the baseline will have three components: a difference in height of the antennas (b_x), an east-west component (b_y) and a north-south component (b_z). The projection of the baseline in the u-v plane is:

$$u = \frac{1}{\lambda} \quad (b_x \sin(HA) + b_y \cos(HA)) \qquad (7.3)$$

$$\begin{aligned}
v = \frac{1}{\lambda} \quad &(-b_x \sin(dec) \cos(HA) \\
&+ b_y \sin(dec) \sin(HA) + b_z \cos(dec)), \qquad (7.4)
\end{aligned}$$

[7]Right ascension is measured towards the east, but the hour angle increases towards the west

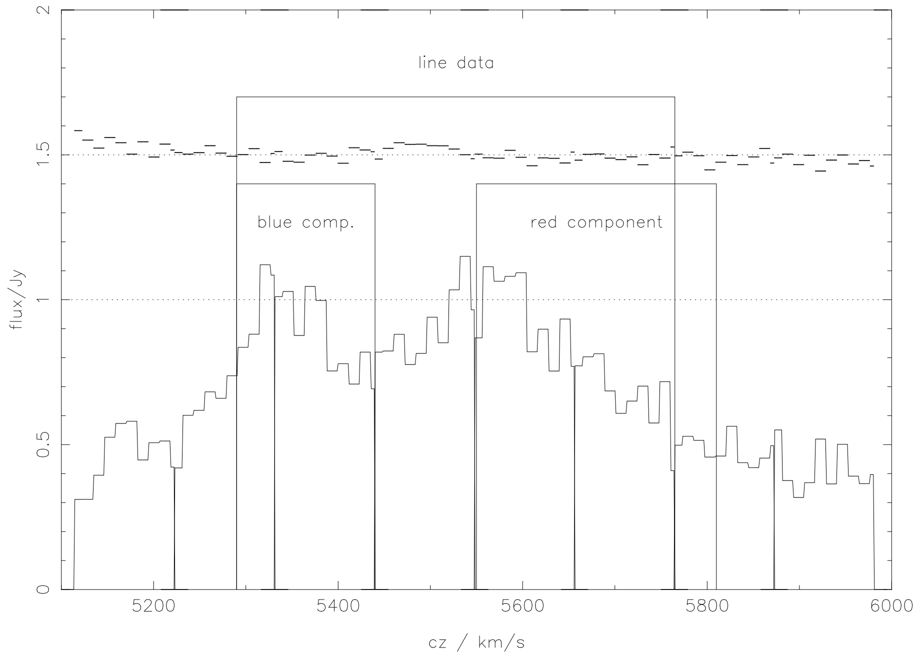


Figure 7.15. Velocity range chosen for blue and red-shifted component of Arp 220.

where *dec* is the declination of the source.

The phase difference becomes

$$\Delta\phi = 2\pi(-u\Delta RA - v\Delta dec). \tag{7.5}$$

Separation of blue and red components in Arp 220

In the case of Arp 220 we wanted to determine the offset of the source of blue-shifted emission from that of red-shifted emission. To separate the blue and red-shifted emission the spectrum (Fig. 7.15) suggests a range of 5290 - 5440 km s^{-1} for the blue shifted emission and 5550 - 5810 km s^{-1} for the red-shifted emission. First the phase in the blue and the red part of the spectrum was measured. The phase in the red part was used as a phase reference and subtracted. Fig. 7.16 shows the phases of the red- and blue-shifted spectra after subtraction. We determined the right ascension and declination offsets which are most likely to fit the data using Equ. sa7.5. The best fit gives $-1''.20 \pm 0''.14$ offset in RA and $0''.06 \pm 0''.14$ in declination (Tab.7.4) and is shown in Fig. 7.16 with asterisks.

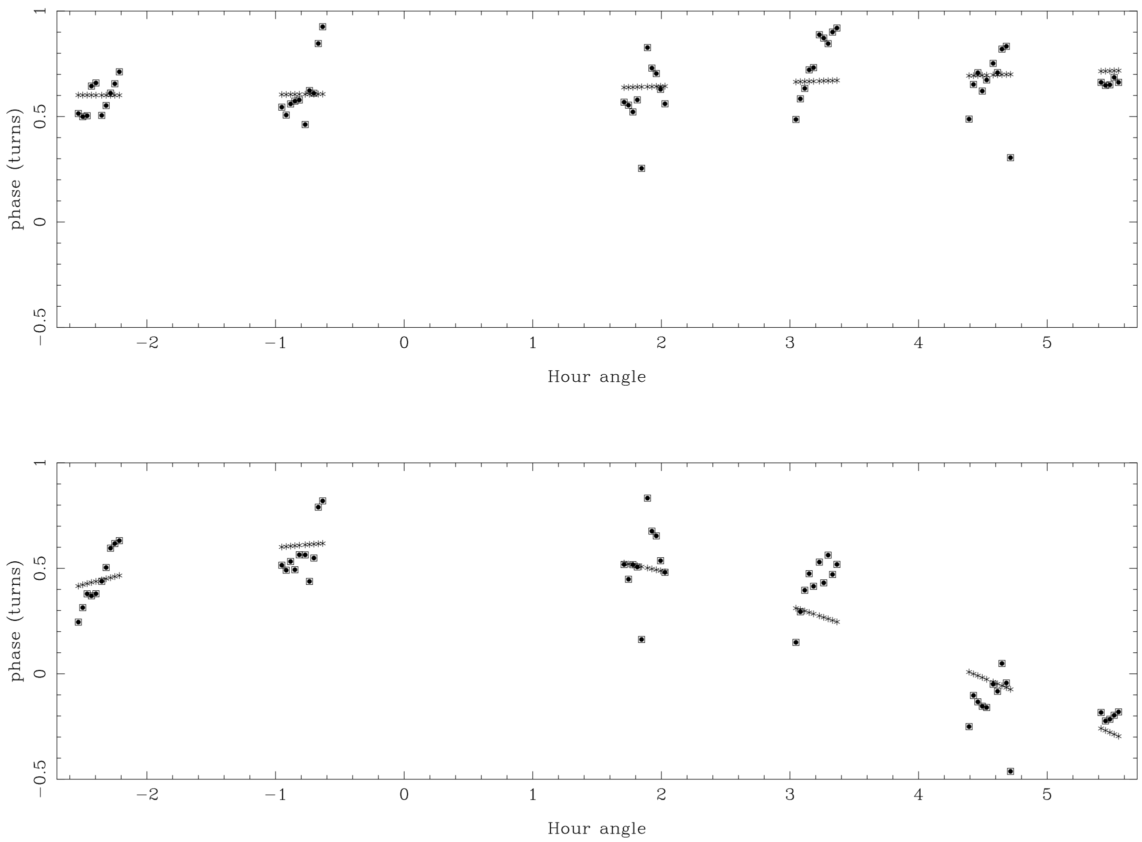


Figure 7.16. Phase of the red (top) and blue (bottom) shifted components of Arp 220. A smooth curve was fitted to the phase of the red component and subtracted from both the red and blue components in order to take out any instrumental phase variations. The scatter on the red phase therefore represents the noise in the data. The asterisks are the models fitting the data best. The model fitting the red component suggests an offset of $0\farcs08$ and $-0\farcs12$, i.e. approximately zero as expected. The model of the blue component suggests an offset of $-1\farcs28$ and $0\farcs18$.

Right ascension	Declination
$-1\farcs20 \pm 0\farcs14$	$0\farcs06 \pm 0\farcs14$

Table 7.4. Offset of blue component of Arp 220 from red component

The offset measured with the phase is slightly bigger than that determined from the amplitude data ($1\farcs0 \pm 0\farcs04$). The next section will show that the blue as well as the red component emit at least some radiation over the whole frequency range. Therefore it is not possible to completely separate the nuclei by choosing different frequency ranges. The $1\farcs2$ separation determined here is the average offset of the origin of blue-shifted emission relative to that of red-shifted emission. The $1\farcs0$ separation determined from the amplitude data is the separation between the peaks of emission independent of their redshift. The separations are not very different though and to first approximation the two emission peaks therefore contribute to different parts of the spectrum.

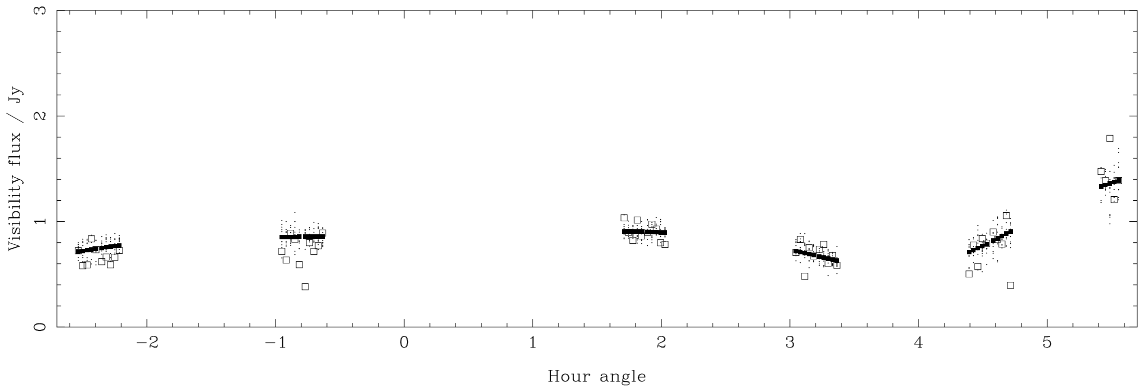
7.5.4 Models for sources of red and blue-shifted emission

Similar to the phase data in the previous section, the amplitude data was divided into a blue (cz=5290 - 5440 km s^{-1}) and a red (5550 - 5810) part (Fig. 7.15). Then models (as described in 7.4.2) were fitted to find the morphology of the blue and red sources of emission respectively. Fig. 7.17 shows the amplitude data (squares) of the red and blue part of the spectrum as well as the binary models which fit the data best (thick line). The parameters of the model are listed in Tab. 7.5.

First, one should notice that each component is best fitted by a binary model rather than by a single elliptical source. If the two nuclei would emit at distinctly different frequencies, one would expect to see a single source in the blue part of the spectrum and the other source in the red part, rather than observing a binary in each component.

However, in both binary models the emission is dominated by one source indicating that most of the flux is coming from this source. The eastern nucleus for example contributes 73% (1.1 Jy/(1.1 Jy+0.4 Jy)) to the red-shifted flux, but only 32% to the blue-shifted flux.

Both models predict a separation of around $1''$ in the east-west direction. If all the emission originated in the two nuclei, one would expect to get the same separation and position angle



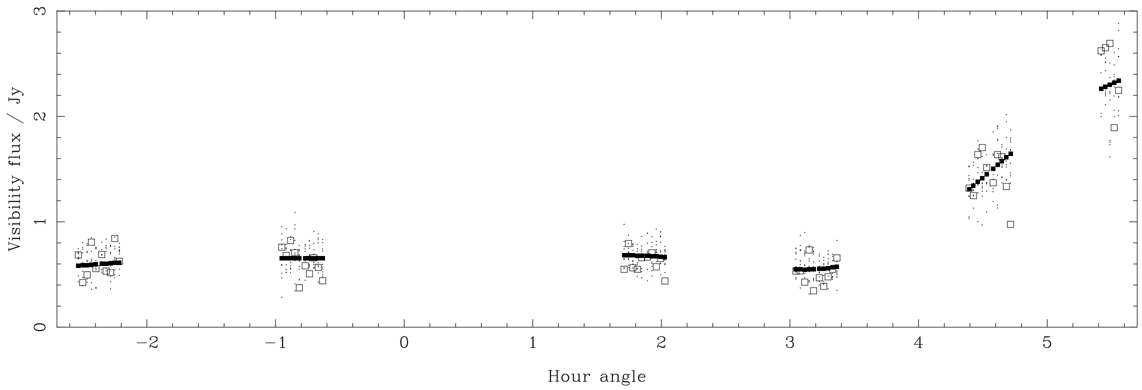


Figure 7.17. In the upper graph the amplitude of the red component of Arp 220 is plotted versus hour angle. The thick line is the best fitting model and the dots represent simulated data given the model. The lower graph show data, model and simulated data for the blue component of Arp 220.

	These data **CO3 $\rightarrow$ 2**	**SYB** **CO2 $\rightarrow$ 1**	**D&S** **CO2 $\rightarrow$ 1**
Red component			
(5450 - 5650 km s^{-1})			
Separation	$1''.1 \pm 0''.05$	$1''$	$1''.3$ (NE-W)
Position angle	$90° \pm 5°$	$86°$	$85°$ (NE-W)
North-eastern nucleus			
Flux	1.1±0.1 Jy	—	—
Size	$(0''.5 \pm 0''.2)^2$	$0''.25$ (theory)	$(0''.9)^2$
Western nucleus			
Flux	0.4±0.1 Jy	—	—
Size	$(0''.3 \pm 0''.2)^2$	$0''.25$ (theory)	—
Blue component			
(5175 - 5325 km s^{-1})			
Separation	$0''.9 \pm 0''.04$	$1''$	$0''.85$ (SE-W)
Position angle	$100° \pm 5°$	$86°$	$110°$ (SE-W)
South-eastern nucleus			
Flux	0.8±0.1 Jy	—	—
Size	$(0''.7 \pm 0''.1)^2$	$0''.25$ (theory)	—
Western nucleus			
Flux	1.7±0.1 Jy	—	—
Size	$(0''.7 \pm 0''.1)^2$	$0''.25$ (theory)	—

Table 7.5. Parameters of models fitting the red and blue components of the CO 3 $\rightarrow$ 2 data best. For comparison the parameters of CO 2 $\rightarrow$ 1 data taken by SYB (Scoville et al., 1997) and D&S (Downes and Solomon, 1998) are also listed.

Nucleus	North-eastern	South-eastern	Western
Velocity range [km s^{-1}]	5580 - 5740	5220 - 5420	5140 - 5660
Flux peaks at [km s^{-1}]	5620	5300	5340
Separation from western n.	$1\farcs3$	$0\farcs85$	$0\farcs0$
Position angle	85°	110°	—

Table 7.6. Downes and Solomon (1998) distinguish between these three sources of emission in the CO $2 \rightarrow 1$ line.

for the models of the red and blue part of the spectrum.

Although the 3 sigma errors of the model parameters overlap, the data indicate that the distances and position angles are slightly different. This can be explained by D&S's model which subdivides the eastern nucleus into two parts (Tab. 7.6), so that the emission originates from three nuclei, a north-eastern component and a south-eastern component and the western component. The western nucleus predominately emits in the blues part of the spectrum, but also exhibits some emission in the red part. Therefore both binary models will contain the western nucleus. The north-eastern component emits quite strongly in the red part of the spectrum between 5580 km s^{-1} and 5740 km s^{-1}. The distance between the western and north-eastern components is $1\farcs3$ at P.A. 85°. In contrast to the north-eastern component, the south-eastern component emits in the blue part of the spectrum between 5220 km s^{-1} and 5420 km s^{-1}. It is separated from the western component by $0\farcs8$ along P.A. 110°. This separation and position angle agree well with that of the binary model for the blue part of the spectrum. In conclusion, the data suggest that there are three sources present. The model on the red part of the spectrum measures the separation between the western and north-eastern component, whereas the model for the blue part measures the distance between the western and the south-eastern component. Therefore it is not surprising that the separations and position angles are different.

7.6 Summary

The interferometer maps of 229 GHz continuum emission and of CO $2 \rightarrow 1$ suggest that most of the emission in Arp 220 comes from two nuclei separated by $\sim 1''$. Therefore we modeled our 342 GHz continuum and the CO $3 \rightarrow 2$ transition data with binary models. Binary models are consistent with the data, except for the CO $3 \rightarrow 2$ data integrated over the whole spectrum, which can not be reproduced well by a simple binary model (see deviation of the model from the data at HA = –1 in Fig. 7.13). A model with three components would probably be more appropriate, but would increase the number of free parameters from 10 to 16. This is clearly too large a number of parameters to fit to so few data points. Instead we have divided the CO $3 \rightarrow 2$ emission into a blue and a red-shifted component. Each component can be well represented by a binary model.

The models fitting the submillimetre data best suggest the following:

- <u>Continuum emission at 342 GHz</u>
  - The continuum emission comes from two nuclei separated by $1''$ at P.A. 80°.
  - Similar physical processes give rise to the continuum flux in both nuclei.
  - The masses of the two nuclei are of the same order of magnitude.
- <u>CO $3 \rightarrow 2$</u>
  - The CO $3 \rightarrow 2$ spectra suggest that 50% of the line emission comes from a more extended source ($\geq 3''$).
  - The physical conditions giving rise to the CO emission are similar in both nuclei.
  - The CO emission is subthermally excited, suggesting densities below 1000 cm^3 to 10 000 cm^3.
  - The line emission originates from a more extended region than the continuum, suggesting that the CO emission originates in an envelope around a denser core.
  - CO emission from the western nucleus is predominately blue-shifted and that of the eastern nucleus is mainly red shifted.
  - However each nucleus emits over a large velocity range (≥ 300 km s^{-1}), consistent with SYB suggestion that each nucleus has its own accretion disk.
  - The eastern nucleus can be divided into a red shifted northern component at $1\farcs1$ separation and P.A. 90° from the western nucleus, and a blue-shifted southern component at $0\farcs9$ separation, P.A. 100° from the western nucleus.

Not surprisingly, SYB and D&S millimetre data and our submillimetre data suggest a similar morphology of Arp 220. In particular, our data is consistent with D&S's model.

During the data reduction I was however most fascinated to learn how much information is contained in only 12 data points, 6 continuum data points and 6 CO $3 \rightarrow 2$ transition spectra. It is amazing how much information about a galaxy 250 million light years away a few voltage vectors can convey.

Appendix A

Additional phase correction

This appendix contains some more water vapour and phase data, as presented in §6.4.2.

On November 8th 1997 we observed the hydrogen recombination maser MWC349 with the interferometer and simultaneously monitored the atmosphere with the radiometers. The weather was relatively poor with 3.3 mm pwv at 1 airmass. The graphs are for consecutive 6 minute intervals. The measured interferometer phase is plotted in all four graphs (dashed line). The JCMT WVM data was subtracted from the CSO WVM data. The resultant brightness temperature was converted into phase, choosing the conversion factor which gives the best fit. The predicted phase of channel I is plotted in the first graph, that of channel II in the second, and that of channel III in the third (solid lines). The solid line in the lowest graph shows the interferometry data after correction with the phase predicted by channel II. Though the corrected phase still has rms errors between 47° and 60°, the rms is significantly reduced in all cases.

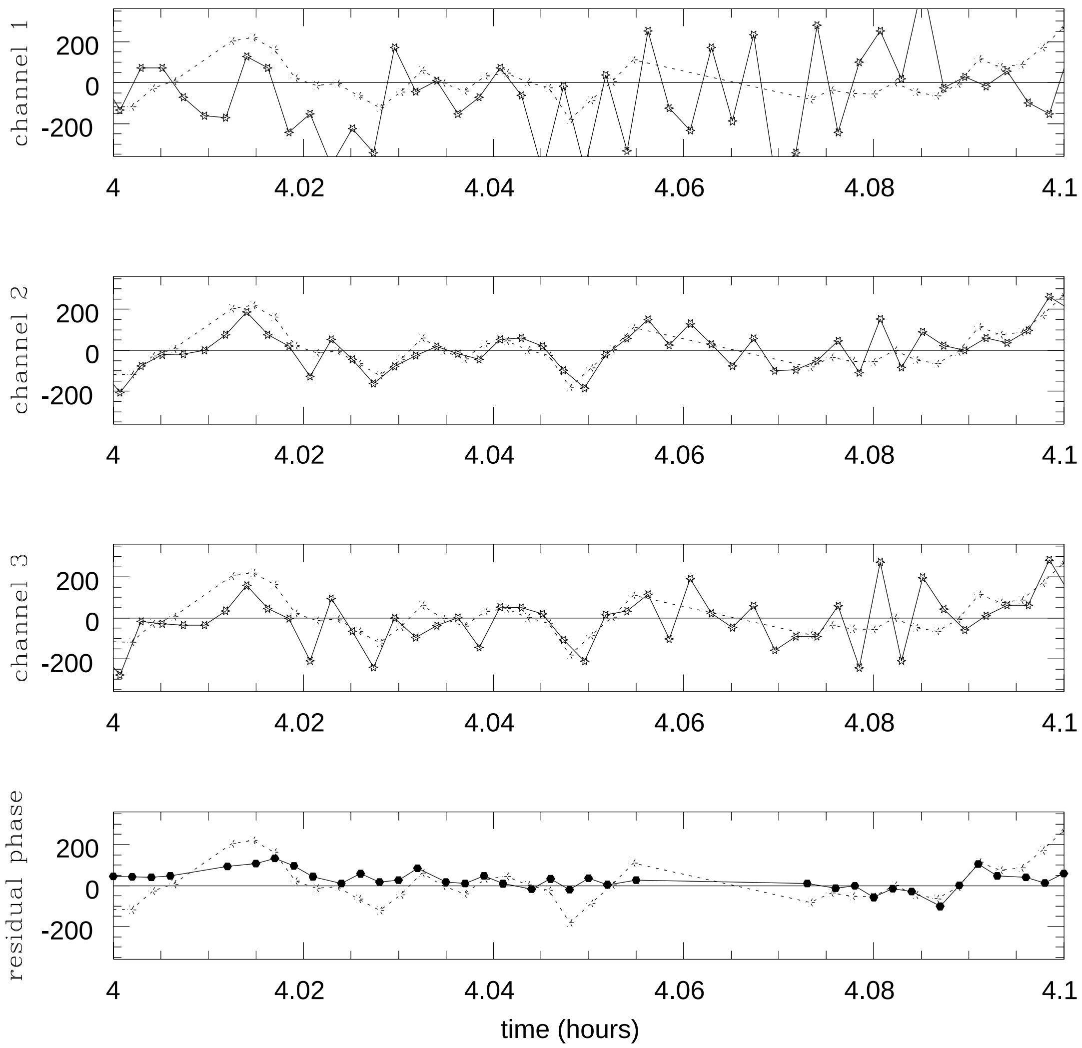


Figure A.1. Phase of MWC349 measured by the interferometer (dashed line) overlaid on phase predicted by channel 1 to 3 of the water vapour monitors (solid line in first, second and third plot, respectively). The rms of 79° of the original interferometer phase (dashed line) is reduced to 47° after correction with channel 2 (solid line).

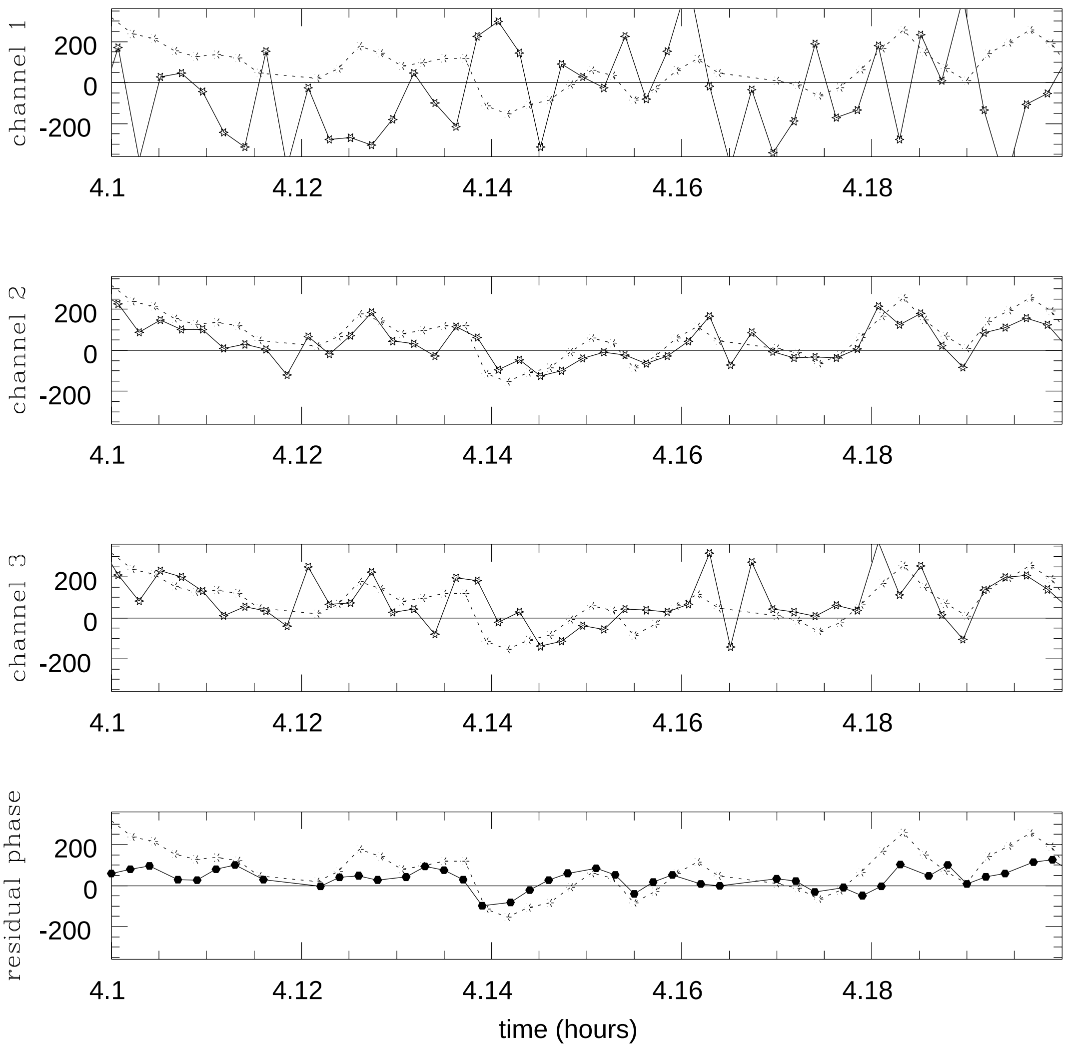


Figure A.2. Phase of MWC349 measured by the interferometer (dashed line) overlaid on phase predicted by channel 1 to 3 of the water vapour monitors (solid line in first, second and third plot, respectively). The rms of 103° of the original interferometer phase (dashed line) is reduced to 55° after correction with channel 2 (solid line).

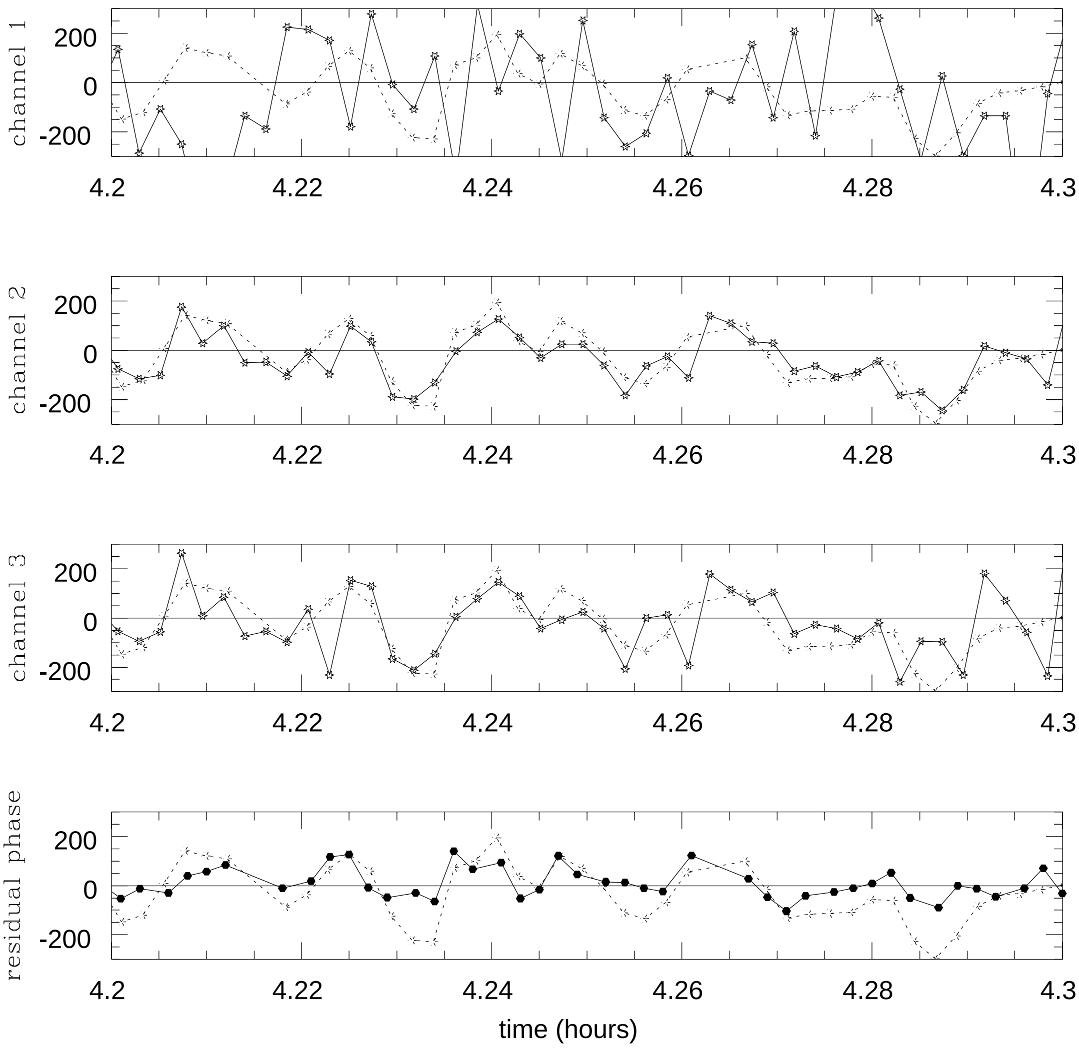


Figure A.3. Phase of MWC349 measured by the interferometer (dashed line) overlaid on phase predicted by channel 1 to 3 of the water vapour monitors (solid line in first, second and third plot, respectively). The rms of 110° of the original interferometer phase (dashed line) is reduced to 60° after correction with channel 2 (solid line).

List of Symbols

$A_{line}(p, T, \nu)$	$=: k_{\nu,line}/\rho$
$A_{cor}(p, T, \nu)$	$=: k_{\nu,cor}/\rho$
$B(\boldsymbol{b})$	correlation function for phase differences as a function of baseline $\boldsymbol{b}$
B_ν	specific intensity (brightness)
$C_j(t+\tau)$	sky temperature of CSO wvm channel j at time $t+\tau$.
$D(\boldsymbol{b})$	phase structure function for baseline $\boldsymbol{b}$
E_l	energy levels of quantum state l
H	scale height of exponential decrease of atmospheric pressure
HA	is the hour angle
H_0	Hubble constant, in this thesis $H_0 = 75kms^{-1}Mpc^{-1}$
$J_i(t)$	sky temperature of JCMT WVM channel i at time t,
$J_iC_j(\tau)$	cross correlation between channel i of the JCMT WVM
	and channel j of the CSO WVM with time delay τ
K	constant in formula of phase structure function
L	excess optical path length
$L_\odot$	solar luminosity ($L_\odot = 3.9 \times 10^{40}$ J/s)
M_r	molecular weight of air ($28.8g/mol$)
$M_\odot$	solar mass ($M_\odot \approx 2 \times 10^{30}$ kg)
R	molar gas constant ($8.31J/(mol\ K)$)
R	radius of curvature of Gaussian beam
RA	right ascension
$T(s)$	temperature (in Kelvin) at position s

T_0	brightness temperature of the radiation entering the layer
$T_{absorber}$	brightness temperature of absorber in optical path of WVM
T_{amb}	brightness temperature of ambient load
T_{bb}	black body temperature of the layer
T_{bri}	brightness temperature
T_{ground}	temperature at ground level
T_{load}	brightness temperature of the calibration load
$T_{measured}$	temperature measured by WVM
T_{N_2}	brightness temperature of load dipped in liquid nitrogen (80 K).
T_{sky}	brightness temperature of sky
T_{sys}	system temperature
V	voltage
a	constant of proportionality
b	baseline length
b_{proj}	the projected baseline length
$\boldsymbol{b}$	baseline vector
c	speed of light $(3.00 \times 10^8 m/s)$
c	coupling factor
dec	declination of astronomical source
f	focal length of a mirror
$\boldsymbol{f}$	the spatial frequency vector with dimensions 1/m
$f(\nu, \nu_{lm})$	lineshape as a function of frequency ν
	and transition frequency ν_{lm}
g	the gravitational constant $(9.80 m/s^2)$
g_l	degeneracy of quantum state l
k	Boltzmann constant $(1.68 \times 10^{-23} J/K)$
h	height above the telescope
i	half the angle between incident and reflected beam
k	Boltzmann constant
$k = 2\pi f/c$	wave number
$k_{\nu,correction}$	correction term of absorption coefficient
$k_{\nu,line}$	absorption coefficient from line emission

n_{moist}	refractive indices of moist air
n_{dry}	refractive indices of dry air
p	pressure (in millibar)
pwv	precipitable water vapour
r	shortest distance to the centre of a Gaussian beam perpendicular to the direction of propagation
$\boldsymbol{s}$	position vector from the telescope
s	distance from the telescope
t	time
v_{cz}	recession velocity due to redshift z
v_{rad}	radial velocity of the interferometer
x	linewidth parameter, see table 3.1
$\boldsymbol{x}$	position vector
z	distance to the minimum beam waist
$\Delta L/\Delta T$	conversion factor
Δs	difference in positions, i.e. length
ΔT	thermal noise of error in temperature measurement
$\Delta T/\Delta L$	sensitivity
ΔT_{gv}	noise due to gain variations
$\Delta T_{thermal}$	thermal noise
Δt	integration time
Δv	velocity range
$\Delta \nu$	linewidth or bandwidth of channel
$\Delta \nu_{lm}^{0}$	linewidth parameter, see table 3.1
$\Delta \nu_{lm}\,(H_2O)$	linewidth parameter, see table 3.1
$\Phi(\boldsymbol{f})$	spatial power spectrum for the spatial frequency vector $\boldsymbol{f}$
Ψ	scalar radiation field
α	elevation angle (straight up is defined as 90°
α	0.5 times the exponent of phase structure function
β	exponent describing the power law of the detectors
λ	observation wavelength
ν	frequency

ν_{amb}	frequency counted in WVM computer
	when looking at load at ambient temperature
ν_{load}	frequency counted in WVM computer
	when WVM is looking at load
ν_{lm}	frequency of transition l $\rightarrow$ m
ν_{N_2}	frequency counted in WVM computer
	when looking at load dipped in liquid nitrogen
ν_{obs}	frequency observed at with the interferometer
ν_{rest}	line frequency as measured in the lab (v=0)
ρ	density of water vapour (g/m^3)
σ_{cont}	sensitivity of interferometer for continuum radiation
σ_{line}	sensitivity of interferometer for line emission
$\tau_\nu(s)$	optical depth at the frequency ν and postion s
ϕ	phase
ϕ_{lm}	probability for the transition l $rightarrow$ m
ω	beam (waist) radius
ω_0	minimum beam waist radius

List of Abbreviations

BIMA	Berkeley Illinois Maryland Array
CSO	Caltech Submillimeter Observatory
DAS	Dutch Autocorrelation Spectrometer
DRO	dielectric resonance oscillator
IF	intermediate frequency
IRAM	Institut de Radio Astronomie Millimétrique
JCMT	James Clark Maxwell Telescope
LSA	Large Southern Array
LO	local oscillator
MMA	Millimeter Array
OVRO	Owens Valley Radio Observatory
PLL	phase lock loop
PD	phase detector
pwv	precipitable water vapour
RAL	Rutherford Appleton Laboratory
RF	receiver frequency
rms	root mean square
Rx	receiver
SMA	Submillimeter Array
V/F	Voltage to frequency converter
wv	water vapour
WVM	water vapour monitors

Bibliography

Asaki, Y., Saito, M., Kawabe, R., Morita, K.-I., and Sasao, T. (1996). Phase compensation experiments with the paired antennas method. *Radio Science*, 31:1615.

Baan, W. A. and Haschick, A. D. (1995). Nuclear antics of IC 4553 (Arp 220). *Astrophysical Journal*, 454:745.

Becklin, E. E. and Wynn-Williams, C. G. (1987). In Longsdale, C. J., editor, *Star Formation in Galaxies*, page 643.

Bremer, M. (1995). The phase project: Observations on quasars. Technical report, Insitut de Radio Astronomie Millémetrique.

Carilli, C. L. and Holdaway, M. A. (1997). Application of fast switching phase calibration at mm wavelengths on 33 km baselines. *VLA Scientific Memo*, 173.

Carlstrom, J., Hills, R., and Lay, O. (1995). Interferometry with the JCMT and CSO. *http://www.jach.hawaii.edu/ hem/userguide/guide/node45.html*.

Conan, J.-M., Rousset, G., and Madec, P. (1995). Wave-front temporal spectra in high-resolution imaging through turbulence. *Journal of the Optical Society of America*, 12(7):1559.

Coulman, C. E. (1989). Atmospheric structure, turbulence and radioastronomical seeing. In Baldwin, J. E. and Shouguan, W., editors, *Radio astronomical seeing*.

Downes, D. and Solomon, P. M. (1998). Rotating nuclear disks in ultraluminous galaxies. *in preparation*.

Dragone, C. (1978). Offset multireflector antennas with perfect pattern symmetry and polarisation discrimination. *The Bell System Technical Journal*, 57:2663.

Focus Software Incorporated (1998). *ZEMAX*.

Gaut, N. E. and Reifenstein, E. C. (1971). *Environmental Research and Technical Report*, volume No 13. Lexington, Massachusetts.

Goldsmith, P. F., Young, J. S., and Langer, W. D. (1983). Curves for analysis of the two lowest rotational transitions of carbon monoxide using the large velocity gradient radiative transfer model. *Astrophysical Journal Supplement*, 51:203.

Harrison, A. (1998). private communication.

Heckman, T. M., Dahlem, M., Eales, S. A., and Weaver, K. (1996). ROSAT observations of the X-ray nebula around Arp 220. *Astrophysical Journal*, 457:616.

Hildebrand, R. H. (1983). The determination of cloud masses and dust characteristics from submillimeter thermal emission. *Quarterly Journal of the royal astronomical Society*, 24:267.

Holdaway, M. A. (1992). Possible phase calibration schemes for the MMA. Technical Report 84, MMA memo.

Holdaway, M. A. and Owen, F. N. (1995). A test of fast switching phase calibration with the VLA at 22 GHz. Technical Report 126, MMA memo.

Houghton, J. T. (1986). *The Physics of Atmospheres*. Cambridge University Press.

JCMT (1998). τ-meter. *http://www.jach.hawaii.edu*.

Joseph, R. D. and Wright, G. S. (1985). Recent star formation in interacting galaxies - II.Super starbursts in merging galaxies. *Monthly Notices of the Royal Astronomical Society*, 214:87.

Keen, N., Spurret, R., and Wylde, R. (1995). An absolute blackbody source for radiometer calibration at submillimeter wavelengths. In *ESA Workshop on Millimeter Wave Technology and Applications*, pages 4.2.1–4.2.10.

Lay, O. P. (1994). *Astronomical Interferometry at Submillimetre Wavelengths.* PhD thesis, University of Cambridge.

Lay, O. P. (1997a). The temporal power spectrum of atmospheric fluctuations due to water vapor. *Astronomy and Astrophysics Supplement Series*, 122(3):535.

Lay, O. P. (1997b). Phase calibration and water vapor radiometry for millimeterwave arrays. *Astronomy and Astrophysics Supplement Series*, 122(3):547.

Lay, O. P. (1998). private communication.

Lay, O. P., Carlstrom, J. E., and Hills, R. E. (1995). NGC1333 IRAS4: Further multiplicity revealed with the CSO-JCMT interferometer. *Astrophysical Journal*, 452:L73.

Lesurf, J. (1990). *Millimetre-wave optics, devises & systems.* Adam Hilger.

Lutz, D., Genzel, G., Sternberg, A., Netzer, H., Kunze, D., Rigopoulou, D., Sturm, E., Egami, E., Feuchtgruber, H., Moorwood, A. F. M., and Degraauw, T. (1996). What powers luminous infrared galaxies? *Astronomy and Astrophysics*, 315:L137.

Masson, C. R. (1992). Phase fluctuations. Technical report, Center for Astrophysics.

Murphy, J. A. (1986). *A heterodyne array receiver for submillimetre wave astronomy.* PhD thesis, University of Cambridge.

Padman, R. (1998). private communication.

Plambeck, R., Forster, R., Frye, B., Lugten, J., Welch, W., and Wright, M. (1996). Interferometer phase correction from receiver total power measurements. In *URSI, National Radio Science Meeting.*

Prado-Carrión, J. R. (1996). *Études de l'atmosphère terrestre au moyen d'observations dans les longueurs d'onde millimétriques et submillimétriques.* PhD thesis, Université Paris VI et Universidad Complutense de Madrid.

Richer, J. S. (1991). *Molecular line and continuum studies of high-mass star formation.* PhD thesis, University of Cambridge.

Richer, J. S. (1998). private communication.

Rohlfs, K. (1986). *Tools of Radio Astronomy.* Springer-Verlag.

Sanders, D. B., Scoville, N. Z., and Soifer, B. T. (1991). Molecular gas in luminous infrared galaxies. *Astrophysical Journal*, 370:158.

Sanders, D. B., Soifer, B. T., Elias, J. H., Madore, B. F., Matthews, K., Neugebauer, G., and Scoville, N. Z. (1988). Ultraluminous infrared galaxies and the origin of quasars. *Astrophysical Journal*, 325:74.

Schilke, P., Bremer, M., Cernicharo, J., Dutrey, A., Guilloteau, S., Hall, P., Hills, R., Lucas, R., Morita, K.-I., Pardo, J., and Wiedner, M. (1998). Calibration working group report. Technical report, LSA.

Scoville, N. Z., Evans, A. S., Thompson, R., Rieke, M., Schneider, G., Low, F. J., Hines, D., and Stobbie, B. (1998). NICMOS imaging of the nuclei of Arp 220. *Astrophysical Journal*, 492:L107.

Scoville, N. Z., Sargent, A. I., Sanders, D. B., and Soifer, B. T. (1991). Dust and gas in the core of ARP 220 (IC 4553). *Astrophysical Journal*, 366:L5.

Scoville, N. Z., Yun, M. S., and Bryant, P. M. (1997). Arcsecond imaging of CO emission in the nucleus of Arp 220. *Astrophysical Journal*, 484:702.

Skinner, C. J., Smith, H. A., Sturm, E., Barlow, M. J., Cohen, R. J., and Stacey, G. J. (1997). A starburst origin of the OH-megamaser emission from the galaxy Arp 220. *Nature*, 386(6624):472.

Smith, H. E., Lonsdale, C. J., Lonsdale, C. J., and Diamond, P. J. (1998). A starburst revealed - luminous radio supernovae in the nuclei of Arp 220. *Astrophysical Journal*, 493:L17.

St-Jacques, D. (1998). *Seeing Compensation in High Resolution Astronomy.* PhD thesis, University of Cambridge.

Sturm, E., Lutz, D., Genzel, R., Sternberg, A., Egami, E., Kunze, D., Rigopoulou, D., Bauer, O. H., Feuchtgruber, H., Moorwood, A. F. M., and Degraauw, T. (1996). ISO-SWS spectroscopy of Arp 220 - a highly obscured starburst galaxy. *Astronomy and Astrophysics*, 315:L133.

Sutton, E. C. and Hueckstaedt, R. M. (1996). Radiometric monitoring of atmospheric water vapour as it pertains to phase correction in millimeter interferometry. *Astronomy and Astrophysics Supplement Series*, 119:559.

Tatarskii, V. I. (1961). *Wave Propagation in a Turbulent Medium*. McGraw-Hill.

Thompson, A. R., Moran, J. M., and Swenson, G. W. J. (1994). *Interferometry and Synthesis in Radio Astronomy*. Krieger Publishing Company.

Waters, J. (1976). Absorption and emission by atmospheric gases. In Meeks, M., editor, *Methods of Experimental Physics, Vol. 12B*, page 142.

Westwater, E. R. (1993). Ground-based microwave remote sensing of meteorological variables. In Janssen, M. A., editor, *Atmospheric remote sensing by microwave radiometry*, page 179.

Woody, D., Holdaway, M., Lay, O., Masson, C., Owen, F., Plambeck, D., Radford, S., and Sutton, E. (1995). Report from the phase calibration working group. Technical Report 144, MMA memo.

Woody, D. P. (1996). A 22 GHz water line monitor for interferometer phase correction. In *URSI, National Radio Science Meeting*.

Zammit, C. C. and Ade, P. A. R. (1981). Zenith atmospheric attenuation measurements at millimetre and sub-millimetre wavelengths. *Nature*, 293:550.