Radiometry of the Atmosphere of Mars

T.J.Vellacott
Oriel College

A thesis submitted to the Faculty of Physical Sciences
for the degree of Doctor of Philosophy
in the University of Oxford
Trinity Term 1992

Abstract

This thesis is concerned with a project to apply the method of pressure modulation radiometry to measurements of the temperature and water vapour distributions in the atmosphere of Mars. The technique has already been used successfully to measure temperature (and occasionally composition) in the atmospheres of Earth and Venus. Two pressure modulators, containing carbon dioxide and water vapour respectively, are used in the Pressure Modulator Infrared Radiometer (PMIRR), on the payload of *Mars Observer*, scheduled for launch in September 1992.

The composition of the Martian atmosphere is almost pure CO_2, so the emission lines are unusually broad, so that the mean pressure in the CO_2 modulator has to be much larger than for previous modulators, thus increasing the power consumption. The limited power available in a Martian spacecraft requires that the power consumption of the pressure modulator and drive be minimized. As a result a high-efficiency drive circuit was designed and developed.

Water vapour is of major importance to the Martian environment and there are several outstanding questions about its behaviour. Measurements using the pressure modulator technique for terrestrial water vapour measurements have had limited success, raising questions about its suitability for condensible, affinitive molecules. A new model of the thermodynamics of pressure modulators was developed to predict their mechanical and spectroscopic behaviour, and spectroscopic measurements of the pressure modulator cell transmission, with high spectral and temporal resolution, using a tunable diode laser spectrometer (TDLS), were performed. The measured transmission agrees well with the predictions of the model, indicating that pressure modulation radiometry can be applied to measurements of water vapour, giving confidence in the successful study of Martian climatology by the PMIRR instrument.

Acknowledgements

Many people have contributed to this work, and it would be impossible to thank all of them. However, I should like especially to thank my supervisor, Professor Fred Taylor, for all his sustained interest, help and encouragement over the last six years.

Most of the measurements described in this thesis were made at the Jet Propulsion Laboratory in Pasadena, California, and thanks are due to most of the PMIRR team for different things; supremely, however, to Dr Dan McCleese, the Principal Investigator, by whose kindness on both scientific and personal levels I was able to carry out the work, and Dr Randy May, who spent a lot of time setting up his spectrometer for me to use. Dr Tim Schofield and Dr Chris Webster also have my gratitude for their helpful suggestions and support.

The assembly of the pressure modulators and other experimental equipment would have been impossible without all the Department's technicians, especially Martin Clarke and Nick O'Donnell. Dr Simon Calcutt has been a constant source of inspiration and energy and has greatly facilitated the work on a practical level.

I am very grateful to my wife, Anna, for all her patience and tolerance over the time this work has taken; and to my friends, without whom it would have been completed very much sooner.

The Science and Engineering Research Council provided financial support during the first three years of this research.

Commonly-used acronyms and symbols

AC	Alternating current
AFGL	Air Force Geophysics Laboratory
CMRR	Common-mode rejection ratio
DAC	Digital-to-analogue converter
DC	Direct Current
DVM	Digital Voltmeter
E''	Ground state energy
EMF	Electromotive force
EPROM	Electrically-programmable read-only memory
FET	Field-effect transistor
FOV	Field of view
FSR	Free spectral range
FWHM	Full (line) width at half maximum (absorption)
HITRAN	Database of spectral line parameters
IC	Integrated circuit
IR	Infrared
IRIS	Infrared Interferometer Spectrometer
IRTM	Infrared Thermal Mapper
ISAMS	Improved Stratospheric and Mesospheric Sounder
JFET	Junction field-effect transistor
JPL	Jet Propulsion Laboratory
k	Spring constant
k_B	Boltzmann's constant
MAWD	Mars Atmospheric Water Detector
MCT	Mercury Cadmium Telluride
NASA	National Aeronautics and Space Administration
NEΔR	Noise-equivalent change in radiance
NEΔT	Noise-equivalent change in temperature
NEP	Noise-equivalent power
PA	Power amplifier
PLL	Phase-locked loop
PMDC	Pressure modulator drive circuit
PMIRR	Pressure Modulator Infrared Radiometer
PMR	Pressure Modulator Radiometer
PMU	Pressure Modulator Unit (mechanism)
PPO	Position pick-off
PRT	Platinum resistance thermometer
PSD	Phase-sensitive detector or detection
PSU	Power supply unit
PWM	Pulse-width modulation
RMS	Root mean square
SAMS	Stratospheric and Mesospheric Sounder
SBC	Self-broadening coefficient

SNR	Signal-to-noise ratio
SVP	Saturated vapour pressure
TDL	Tunable diode laser
TDLS	Tunable diode laser spectrometer
UV	Ultra-violet
VCO	Voltage-controlled oscillator
VL	Viking Lander
VORTEX	Venus Orbiter Radiometric Temperature Experiment
WB	Wideband

Other symbols, used less frequently, are defined in the text.

Contents

Chapter 1:
Observations of the Martian Atmosphere

1.1 Overview

The Martian atmosphere has been a major source of interest ever since the planet was first observed through telescopes. Mars is still thought the most likely planet in the Solar System (other than Earth) to sustain life, since its orbital radius and size, like those for Venus, are similar to those of Earth, while unlike Venus, its atmosphere, though cold, thin, and mostly carbon dioxide, is not totally inhospitable. Many hypotheses have been made as to the composition of the Martian atmosphere, and estimates of the surface pressure have decreased over the years, but since the arrival of spacecraft to make extensive measurements at close range, certain aspects are now well-documented.

1.2 Earth-based measurements of the Atmosphere

Observations of Mars from Earth have often proved misleading. The most notable example is that of the 'canals', first reported by Schiaparelli in 1877 and subsequently popularized extensively by Lowell in 1905. Before the space age the atmosphere was generally held to consist mostly of nitrogen, with a surface pressure of about 80 mb. In fact it is almost entirely carbon dioxide, and the surface pressure is about 6–7mb.

However, major dust storms have often been observed from Earth, and the fact that they generally occur shortly after the Martian perihelion noted. Measurements of the wind speeds (up to 22 m/s [de Vaucouleurs 1954]) have also been made, by tracking cloud features. Water vapour spectroscopy has been performed from Earth with some accuracy, and radar surface mapping has also been successful, although it lacks the resolution to discriminate the steep Martian relief.

1.3 Space missions to Mars

Table 1.I gives a summary of the spacecraft sent to Mars up to 1976. In general, the US missions to Mars have been highly productive; of the eight spacecraft sent to Mars to date, only two have failed. The Soviet Union has, by comparison, had little success with space missions to Mars; the more successful ones (*Mars 2–7*) have measured the surface temperature, atmospheric composition, cloud heights, surface pressure, magnetic field, ionospheric chemistry and relief.

1.3.1 *Mariner*

Four spacecraft of the *Mariner* series have visited Mars. *Mariner 4* took 22 low-resolution images of 1% of Mars showing more than 70 craters. It measured the surface pressure (6

	Name	launch	arrival	comments
USSR	1960A	10 Oct 60	—	failed to reach earth parking orbit
USSR	1960B	14 Oct 60	—	failed to reach earth parking orbit
USSR	Sputnik 22	24 Oct 62	—	failed to leave earth orbit
USSR	Mars 1	1 Nov 62	June 63	flyby – orientation failed; lost contact
USSR	Sputnik 24	4 Nov 62	—	flyby – failed to leave earth orbit
USA	Mariner 3	5 Nov 64	—	launch failure
USA	Mariner 4	28 Nov 64	14 Jul 65	flyby at 9844 km – images returned
USSR	Zond 2	30 Nov 64	Aug 65	flyby+?lander; contact lost April 65
USSR	Zond 3	18 July 65	—	communication test
USA	Mariner 6	25 Feb 69	31 Jul	flyby at 3215km
USA	Mariner 7	27 Mar 69	5 Aug	flyby at 3516km
USSR	1969A	27 Mar 69	—	failed to reach earth orbit
USSR	1969B	14 Apr 69	—	failed to reach earth orbit
USA	Mariner 8	8 May 71	—	Booster failed
USSR	Cosmos 419	10 May 71	—	Orbiter/lander: failed to leave earth orbit
USSR	Mars 2	19 May 71	27 Nov	orbiter successful; lander failed
USSR	Mars 3	28 May 71	2 Dec	orbiter successful; lander failed
USA	Mariner 9	30 May 71	14 Nov	orbiter – worked till 27 Oct 72
USSR	Mars 4	21 Jul 73	10 Feb 74	orbiter mission; retro-rocket failure caused flyby
USSR	Mars 5	25 Jul 73	12 Feb 74	orbiter
USSR	Mars 6	5 Aug 73	12 Mar 74	flyby/lander: contact with lander lost
USSR	Mars 7	9 Aug 73	9 Mar 74	flyby/lander: lander missed planet by 1300 km
USA	Viking 1	20 Aug 75	19 Jun 76	orbiter/lander; lander worked to 13 Nov 82
USA	Viking 2	9 Sep 75	7 Aug 76	orbiter/lander; lander worked to 12 April 80

Table 1.I: Space missions to Mars to 1976. Six of the eight US missions have succeeded, whereas only five of the (former) USSR's sixteen missions have sent back useful results. More recently its *Phobos* missions have also failed. Data mainly from NASA [1984].

mb) by radio occultation, and determined that most of the atmosphere consists of carbon dioxide, although some argon was also detected. *Mariner 6* and *Mariner 7* together mapped 20% of the planet with a higher resolution imaging system, returning 193 images in total. Other instruments carried were a UV spectrometer, an IR spectrometer and temperature sensors. The composition of the atmosphere was determined as being greater than 95% CO_2, and between 0.5% and 4% N_2 with some H_2O. Traces of CO_2^+, CO, atomic hydrogen and some O_2 were detected. The equatorial temperature was found to vary between 200 and 290K, and the minimum polar temperature to be around 150K.

Mariner 9's payload included a low and high-resolution imaging system for surface mapping, an infrared spectrometer for surface temperature, a UV spectrometer for atmospheric composition and structure, an IR interferometer spectrometer similar to the *Nimbus* IRIS, for the surface and atmospheric radiation, and an S-band radio occultation experiment to measure the pressure and structure of the atmosphere. Its orbit was 17140×1387 km with a period of 11.98 hours and an orbital inclination of 65°. The low periapsis altitude gave an overlap between wide-angle camera images, and the sub-spacecraft track repeated after 17 days and 35 revolutions. The mission had a total useful lifetime of 515 days, of which 349 were in Martian orbit, and returned 7329 images of Mars and its moons.

Mariner 9 arrived at Mars during a major global dust storm which obscured most of the

surface. Initially only the south polar ice cap and *Olympus Mons* were visible. After the main storm had cleared, several localized dust storms were observed. The images, which covered all the planet, revealed not only craters (mostly in the northern hemisphere), some of which had been observed even through the dust, but also huge volcanoes and canyons, intricate channels and soil deposits at the poles. There are many volcanoes on Mars, mostly in the northern hemisphere. Four are gigantic; three of these lie on the *Tharsis* bulge. These are geologically young 'shield'-type volcanoes, with a broad caldera and gently-sloping sides. The largest, *Olympus Mons*, is 27 km high, and 600 km broad at its base, with a caldera 64 km across, much larger than the largest equivalent feature on Earth (*Mauna Kea* in Hawaii). *Olympus* was known to be a high region from terrestrial radar measurements, but was not proved to be a mountain until observed by the spacecraft. The most notable canyon, *Valles Marineris*, named after *Mariner 9*, is a huge chaotic valley 4000 km long, 700 km wide at its widest, and 7 km deep, a little east of *Tharsis*, with many interconnecting valleys, and evidence of landslides and faulting.

1.3.2 *Viking* Landers

The *Viking* landers were equipped to make measurements in the atmosphere during their descent, and on the surface after landing. The entry science experiments included a Retarding Potential Analyzer (RPA) which measured the solar wind and ionospheric electrons, to determine the temperatures of the electrons and the composition, concentration and temperatures of positive ions, and an Upper Atmosphere Mass Spectrometer (UAMS), to measure the composition of the atmosphere, including some isotope ratios, between 230 and 100 km. Apart from CO_2, the spectrometers on both landers detected 2.5%N_2, 1.5%Ar, and some CO, O_2, and HNO_3. Number densities were derived from the mass peaks for CO_2, Ar, N_2, CO, O_2 and NO, and temperatures from masses 44, 22 and 12. Generally the temperatures measured by the UAMS instruments were 10% lower than those measured by the Retarding Potential Analyzers. Nearer the landing, the density and temperature of the lowest 90 km of the atmosphere were measured by accelerometers, temperature sensors, and an aeroshell stagnation-pressure instrument.

The lander sites were chosen from the orbiter images and Earth radar to provide, in addition to the safest, the wettest landing terrain, thereby increasing the probability of detecting life. *Viking Lander 1* (VL1) landed in *Chryse Planitia* (22.46°N, 48.01°W) in the morning of 20th July 1976, and VL2 at *Utopia Planitia* (47.96°N, 225.77°W) in the afternoon of 3rd September. They measured the pressure, temperature, and wind speed and direction for more than one Martian year, and analyzed the soil, finding that the most abundant elements were silicon (15–30%), and iron (12-16%). About a quarter of the iron is probably in the form of magnetite, so that the soil is magnetic.

1.3.3 *Viking* Orbiters

The orbiters were solar-powered, 3-axis-stabilized spacecraft. The orbital periapsis was 300 km, and the initial period such that the orbiters passed above the landers on each orbit. The imaging system had faster and higher-resolution cameras than *Mariner 9*, allowing high-resolution mapping of contiguous areas. Images taken during the approach showed clouds up to 19 km up the sides of *Olympus Mons*, and layered haze at around 30 km altitude in *Argyre*. From orbit, dendritic channels and large flood features were observed, particularly in the equatorial regions east of *Valles Marineris*. It is generally accepted from this evidence that liquid water flowed at one time on Mars, implying that the surface

pressure must once have been much higher than it is now.

Some of the most interesting atmospheric measurements were made by the *Mars Atmospheric Water Detector* (MAWD), a five channel grating spectrometer with a spectral resolution of about 1.2cm^{-1}, and a field of view of approximately 1° by 0.1°. It used five PbS detectors cooled to 200K, of which two measured radiation from at the centres of two water lines of different strengths near 1.4μm, and the other three looked at the wings on either side and between the two. From these, monthly water vapour maps of all the sunlit planet were generated. The minimum detectable water vapour amount was 0.1 precipitable microns and the average accuracy 20%. The retrievals were based on 81 lines in the 1.38 μm water vapour band, with an allowance made for the effect of the measured solar spectrum. The line-of-sight water amount is converted to the vertical column amount by dividing by η, the sum of the secants of the solar incidence and atmospheric emission angles. Particulate scattering leads to an underestimation of the water amount but this effect can be ignored for low values of η (less than 4). MAWD's measurements are further discussed later in this chapter.

1.4 Atmospheric characteristics

The atmosphere has a troposphere extending from the surface to about 30 km in altitude, a stratosphere to about 120 km and a thermosphere above that. CO_2 is the major constituent (95%) below 160km. The troposphere and stratosphere are well-mixed; above the stratopause the volume mixing ratios of N_2, Ar and O_2 increase rapidly. Mars has an ionosphere with a lower boundary at about 100km, which is enhanced as a result of the significant penetration of the atmosphere by the solar wind, because of Mars's weak magnetic field. In this region CO_2 is photolysed, and O_2^+ is the major constituent, with a peak in concentration at about 130 km, where the ionic temperature is about 160 K. CO_2^+ is about one ninth as abundant. Photon-induced ion exchange also takes place, with CO and O_2^+ being formed from CO_2^+ and O.

The Martian atmosphere exhibits daily, seasonal, geographic and topographic variations in wind, dust, cloud and water content. Clouds, hazes, fronts and airglow have also all been seen; in short, the Martian atmosphere seems to undergo many of the processes observed in the terrestrial atmosphere. However, there are important differences. There are no oceans acting as sources and sinks of convectable heat, and large diurnal variations in surface weather are observed. Large-scale dust storms dominate the low-atmosphere behaviour for months at a time, and in the polar regions, where condensation and evaporation are important on a large scale, there are dramatic seasonal variations in pressure, temperature and water vapour content.

The 24° inclination of its axis of rotation leads to distinct seasons on Mars, and because of the eccentricity of its orbit, the length of the seasons varies; as the southern summer occurs near perihelion, it is hotter (and shorter) than the northern summer.

1.4.1 Composition

Table 1.II shows the presently-accepted composition of the Martian atmosphere below 120km. The scale height (the distance over which the pressure changes by a factor of *e*) was measured during the *Viking* lander entries as 11.697 and 11.363km for VL1 and VL2 respectively, indicating effective atmospheric molecular masses of 44.36 (±0.41) and 43.36 (±0.35). The surface mass spectrometer indicated 43.486 (±0.066). This is consistent with

gas	molecular weight	by weight %	by volume %	abundance atm-cm
CO_2	44.01	96.63	95.32	10109.3
N_2	28.02	1.74	2.7	286.35
Ar	39.944	1.47	1.6	169.69
O_2	32.00	0.096	0.13	13.79
CO	28.01	0.045	0.07	7.42
H_2O	18.016	0.012	0.03	3.18
Ne	20.18	1.16×10^{-4}	2.5×10^{-4}	0.0265
Kr	83.80	5.8×10^{-5}	3×10^{-5}	0.0032
Xe	131.30	2.4×10^{-5}	8×10^{-6}	0.0008
unknown		0.005	0.15	

Table 1.II: Principal gases in the Martian atmosphere [from NASA 1984]. The data is collated from measurements by many different techniques. The water vapour distribution varies strongly with temperature and height; this is discussed in the text.

an atmosphere consisting mainly of CO_2, which has a molecular weight of 44. Optical absorption by this carbon dioxide enhances the temperature of the surface by about 10K.

Figure 1.1: Low altitude zonal mean temperatures, measured by the *Mariner 9* IRIS instrument [from Conrath 1981]. The altitude at the 0.1 mb level is approximately 40 km. These northern hemisphere data show a sharp gradient around 60°, implying a large vertical (thermal) wind shear in that region.

1.4.2 Temperature and pressure

The IRIS instrument on *Mariner 9* mapped the global temperature profile in January–February 1972; the zonal mean cross section for the northern hemisphere, derived from

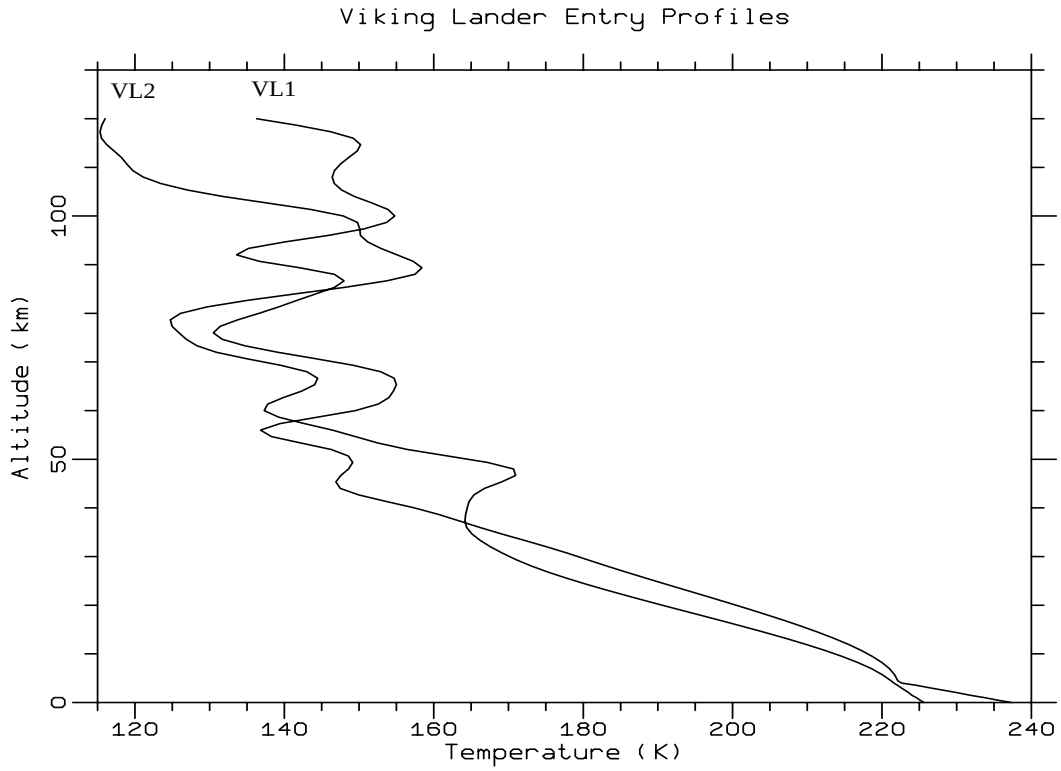

Figure 1.2: *Viking* Lander entry temperatures [data from NASA 1984]. The data above 28km were obtained by the accelerometer; the pressure was measured directly below 5km and extrapolated below 1 km. There are no data between 5km and 28 km.

the emission spectrum of CO_2 at 667 cm^{-1}, is shown in figure 1.1. The instrument had a bandpass from 200–2000cm^{-1}, at 2.4 cm^{-1} resolution. The thermal structure was resolved to an accuracy of 2 K between the surface and 0.1 mb (approximately 40 km) in 30 levels, with a vertical resolution of approximately half a scale height (over which the atmospheric pressure drops by a factor of e, about 5 km for Mars) near the surface, and 2 scale heights near the top of its vertical range.

The vertical temperature and pressure profile of the atmosphere was also measured by the *Viking* landers on entry (figure 1.2). The pressure shows an exponential relationship with height, in accordance with the hydrostatic equation, and the surface pressure shows concomitant variations because of the large range of altitude. Throughout the descents, waves were seen in the temperature profiles, with a wavelength of 4–31 km below 100 km, and 50–60 km above. The amplitudes were a few K below 40 km, rising to 15–20K at 50–60 km, and 20–40K above 100 km. Such amplitudes can arise from vertical excursions of atmospheric packets of only a few km, or from much larger horizontal excursions [NASA 1984].

The surface pressure measured by VL1 when it landed was 7 mb, and this was observed to decrease at the rate of 0.012 mb/sol (a sol is a Martian day – 24 hours, 37 minutes) as winter approached. VL2, which was at a lower altitude, measured a higher pressure of 7.7mb. The *Viking* landers observed diurnal and semi-diurnal oscillations in surface pressure, possibly arising from solar tides. There is a phase lag, similar in magnitude to that observed on Earth, between the solar heating and the temperature cycles, indicating that the circulation is predominantly convective.

Typical surface temperatures range from 220K at the equator to 120K at the winter pole,

whilst noon temperatures in equatorial regions can reach 300K in summer. The dark areas on the planet (*Maria*) are in general warmer than the light areas. The temperatures measured by the *Viking* landers were in the range of 150K (VL2) to more than 250K (VL1) with a typical diurnal variation of 35–50K, from 240K at 2 p.m. to 190K just after dawn, with a maximum at about 2 p.m. The *Viking* Infrared Thermal Mapper (IRTM) measured daily temperature ranges of 140–270K, much greater than the typical diurnal temperature variation of 5–10K on Earth, where the atmosphere is thicker, there are more clouds, and the oceans act as a heat reservoir. The temperatures decreased in the afternoon, while the atmospheric temperatures did not. It has therefore been suggested [NASA 1984] that the cooling was caused by clouds. The measured temperatures dropped rapidly after dusk, indicating a low heat capacity.

1.4.3 Winds

There are tidal wind systems on Mars, arising from the expansion and contraction of the atmosphere with solar heating. *Mariner 9* measured a large temperature gradient near 60°N, which implies a correspondingly large circumpolar thermal wind vortex [Conrath 1981]. The thermal wind increases with height up to about 150 m/s.

The large equatorial-polar temperature difference leads to westerlies and depressions, as on Earth. Because the atmosphere is so thin, it absorbs less radiation from the sun at high summer latitudes than Earth's does, so the summers are very stable with light easterlies less than about 20 m/s in speed. Winds of up to 110 m/s have been observed at the edge of the southern polar ice cap as it thaws. Violent winds also occur in the *Valles Marineris*; the solar heating of one end while the other is still in darkness leads to large temperature and pressure gradients.

The surface winds measured by the *Viking* landers were generally light, reaching 7 m/s at about 10 a.m., dropping to about 2 m/s in the afternoon, with the direction changing from easterly to southwesterly. This pattern was common from day to day. On the surface, the wind leads to the formation of enormous dune-like features; Martian weathering is dominated by wind, although there is evidence of water abrasion in larger-scale features.

1.4.4 Clouds, fogs and frosts

Thin clouds are observed on Mars; the total cloud cover can reach 5%. At mid-latitudes the clouds are of water ice, but at the poles, some CO_2 ice clouds also form. Convective clouds often appear over the equator at noon, especially around the slopes of the huge volcanoes in *Tharsis*; the atmospheric warming causes air parcels to rise, so that they saturate and form ice crystals. Orographic clouds have been observed, and are another particularly prominent feature in the *Tharsis* region. Air parcels forced over the topography can take hundreds of kilometers to relax, so buoyancy waves are formed in the lee of large land masses such as craters and mountains, and clouds may form at the cold extreme of the motion.

Fogs were observed by the *Viking* Orbiters at sunrise in the Canyons, and VL2 observed frosts in the morning. The fogs are believed to arise from water ice condensing onto the surface in the evening and being released into the atmosphere after warming of the ground by the morning sunlight; the ice then recondenses in the cooler air above the low-level inversion layer.

1.4.5 Aerosols

Aerosols are important in the Martian atmosphere as their optical depth is typically about 0.5 [McCleese *et al.* 1986]. There are frequent dust storms on Mars, mostly local, but occasionally global, during which the optical depth rises well above unity, so that aerosols are responsible for much and variable absorption. The aerosol particles, with radii (inferred from scattering and atmospheric opacity) of about one micron, are distributed hydrostatically up to 45 km in altitude. The near-surface dust contributes to large and very rapid changes in the low-altitude temperature, partly because the radiative feedback mechanisms speed the dust propagation. The dust makes the atmosphere more isothermal, thereby altering gas density levels by about a scale height. It also makes the Martian sky pink.

Three types of particles were inferred from measurements by the *Viking* lander imagers, which had 3 infrared channels in addition to their visible ones. Firstly, water ice forms ground fogs, which were present in all the dawn imager pictures. Their average depth was 400m, and they were formed at 2 a.m. local time, growing until sunrise and dissipating in the late morning. Secondly, suspended soil particles were present throughout the *Viking* mission, and formed most of the diurnally-varying component of atmospheric opacity. The average particle radius was 0.4 μm, although the shape is rough and non-spherical. These particles are distributed smoothly from the surface up to heights of more than 30 km, having an exponential height dependence with a scale height of 10 km. They are introduced into the atmosphere by wind mixing, and removed both by planet-wide eddy mixing processes and by acting (in conjunction with water ice particles) as condensation nuclei for CO_2 ice in winter. This process produces a preferential deposition of dust and water ice in the polar regions, which may explain the low albedo of the ice deposits and the clearness of the atmosphere in the winter hemisphere. Thirdly, a 'polar hood' is formed in winter by high-level ice clouds. It is inhibited in summer because the suspended soil particles absorb solar radiation, thus warming the upper troposphere.

Dust storms, which obscure surface features, are commonly observed; 35 were seen from Earth in 1977. Localized dust storms seem to occur in response to seasonal changes in temperature generated by solar radiation. In a local (restricted-area) dust storm the opacity in the visible region can be increased by a factor of 5, although the effect in the infrared is not as marked. Occasionally, local dust storms grow and become global, especially around perihelion, when the solar constant has about 1.4 times its aphelion value. During a global storm the dust reaches up to 50 km in altitude, causing the temperature of the atmosphere in that region to rise by several tens of degrees.

The 1971 dust storm was particularly severe and well-observed. It started on the 22nd September, seen from Earth as a brilliant white cloud streak 2400 km long which covered the whole of the *Noachis* region in a few hours. On 24th September the cloud expanded westwards, obscuring a path around two thirds of the planet in a few days. On 28th September a new cloud was observed in *Eos*. When *Mariner 9* reached Mars (on 14th November) only the south polar cap and the tops of the giant volcanoes were visible, and the planet did not clear until February 1972. The opacity of the dust clouds, which were 8–10 km high, caused an enhancement in atmospheric temperature and a decrease of 20–30K in surface temperature; the atmosphere was observed by IRIS to cool down in the period between January and February 1972 [Conrath 1981] as the dust settled.

1.4.6 Influence of the poles

Both poles have permanent caps which extend a few degrees in latitude, the northern polar permanent cap being generally larger than the southern one. Temperatures around 200K have been measured at the northern pole in summer. Under Martian pressure conditions atmospheric water vapour forms a frost at this temperature, so the northern polar cap is assumed to contain a combination of water ice and CO_2 ice. At the southern pole the summer temperature is typically 160K; this is nearer the frost point of CO_2 and so the southern polar cap is assumed to be almost entirely CO_2 ice. The amount of water stored at the poles greatly exceeds that in the atmosphere.

About a third of the atmosphere condenses out at the poles in winter, lowering the pressure by 20–30% and removing most of the water vapour in the winter hemisphere. This leaves a higher concentration of other gases such as N_2 and Ar. Complete condensation of the carbon dioxide onto the poles is prevented by these non-condensibles, through which the atmospheric CO_2 has to diffuse. The south polar night temperatures measured by the *Viking* IRTM instruments were lower than the CO_2 condensation point at the surface pressure (130 K, compared to 150K), indicating that supercooling takes place.

The condensation causes the poles to grow in both vertical thickness and horizontal extent. The southern cap has been seen to extend up to 60°S, and the northern cap down to 65°N with isolated frost areas further south. The polar caps show spiral patterning as a result of radial winds, and appear to be layered, from the continual erosion and redeposition of ice and soil. The layers are regular in form, implying the existence of cyclic climatic variations on a long timescale. The reflectance of the poles is lower than that for terrestrial ice, because of the admixture of soil and dust, and the northern cap has a lower albedo than the southern one. This is assumed to occur because the global dust storms tend to occur in the southern summer, so that the northern hemisphere is condensing when there is more dust in the air which is drawn onto the pole, making it dirty. Perhaps for this reason, the southern polar cap experiences a greater range of temperatures and therefore shows the larger variation in size.

1.4.7 Water Vapour

Because of the high opacity of water in the thermal infrared, H_2O has a significant impact on the climate and thermal transport mechanisms of the Martian atmosphere, in spite of its small abundance (typically 0.012%).

Around the solstices, the maximum H_2O amount is about 10 precipitable microns at high latitudes in the winter hemisphere, and about 100 precipitable microns at high latitudes in the summer hemisphere [Farmer *et al.* 1977]. Both of these values are near the saturated vapour pressure of water for the measured surface and atmospheric temperatures. (Earth water vapour column amounts are much larger, typically 2–3 precipitable cm.) As the equinox approaches, the water vapour moves from the summer hemisphere to the winter hemisphere, and the greatest concentration is at mid-latitudes. The column abundance measured around the time of the 1969 opposition decreased to less than 5 precipitable microns at the autumn equinox [Owen and Mason 1969, Schorn *et al.* 1969]. The following southern summer was similar to the northern summer.

The diurnal variation of water vapour column amount observed by MAWD followed the pattern of an increase from zero at dawn to a maximum shortly after noon. The magnitude of the variation was smaller than that inferred from terrestrial measurements, although the qualitative pattern had been identified from Earth.

There is small-scale structure in the water vapour concentration in regions with steep relief, thought to be caused by orographic adiabatic cooling. Although adiabatic cooling also occurs in flatter areas, its effects are reduced by radiative processes, because the radiative relaxation time for Mars is short (about 1 day). This corresponds to a distance of only about 200 km [Goody and Belton 1967]. This may explain why no correlation has been seen between shallow relief and water vapour abundance, whilst over areas of rough terrain, there are marked daily variations in water vapour abundance, and shifting wind patterns, dust and dawn haze.

MAWD measured a boundary in water vapour concentration at 30°N, between 20 and 70 °W; a steep rise is observed from 10 precipitable microns from 20–30°N to 26–28 precipitable microns at 35°N. This boundary feature was stationary for at least six weeks, and Farmer *et al.* [1977] could not explain it in terms of topography or any other cause, although Wells [1979] suggests that it may arise from a spur from *Tempe*. There is occasionally a stationary white cloud in this region.

The sources and sinks of water vapour are not yet well known. In the 1960s it was widely held that the atmospheric water vapour came from melting of the polar cap, and that there could be a surface adsorption reservoir. The ionospheric measurements of atomic hydrogen (by *Mariner 9*) and O^+ (by *Viking*) suggest that one of the loss mechanisms is as atomic ions at high altitude; it is also widely held that a large amount of water is held in a sub-surface permafrost.

1.5 Current topics of scientific interest

The composition and overall structure of the atmosphere are now well-known, and the most significant questions involve water vapour, dust and dynamics. The detailed form of the general circulation, and the interactions between the dust, water vapour and the surface, climate and atmosphere are largely unknown. In particular, the different sizes of the winter polar caps, and the consistency of the annual pressure cycle considering the variable dust amount cannot be explained by current simple models [Leovy 1990].

The three-dimensional temperature structure is still to be resolved, as previous measurements have had insufficient signal-to-noise ratio to measure the longitudinal variations, required for understanding of the circulation and its forcing mechanisms, such as the distribution of heating and cooling, and the size, gravity, rotation rate, topography, albedo and surface heat capacity of the planet, to enable comparison with other planetary atmospheres. The reason for the steep temperature gradient at 60°N (measured by the *Mariner 9* IRIS) also needs to be explored.

The MAWD measurements have left several unresolved questions concerning water vapour. These include the mechanism for the seasonal migration of water, the behaviour near the pole, the nature and cause of the small-scale structure observed by MAWD, and the possible existence of other surface reservoirs. Furthermore, the supposed much greater water abundance in past epochs requires a search for the cause of its subsequent demise. Current theories include early outgassing or cataclysmic impact, which reduced the atmospheric density, causing saturation, and thereby removal of the water.

The contribution of the dust to the overall radiative balance, the manner in which it is transported, and the causes of dust storms, are all necessary for full understanding of Martian atmospheric dynamics. At present, dust transport occurs predominantly during northern winter and the reason for this asymmetry is not understood. The dramatic events

expected in winter and spring, when the rate of exchange of CO_2, H_2O and dust between the atmosphere and poles is greatest, with significant short-term effects on the global circulation, also warrant intensive study, including three-dimensional measurements of temperature, dust amount and water vapour.

The Pressure Modulator Infrared Radiometer (PMIRR), described in chapter 2, will address these problems by making simultaneous, co-located, three-dimensional measurements of temperature, water vapour concentration, and longwave radiative balance. These measurements will in turn provide more information for the design of future, possibly international, exploratory missions to Mars.

This page intentionally left blank.

Chapter 2:
The *Mars Observer* programme and PMIRR

2.1 *Mars Observer*

Mars Observer, to be launched in September 1992, will be the first US spacecraft to be sent to Mars since the *Vikings* in 1976. The orbit will be a low (361-km) solar-synchronous polar orbit, with a velocity of 3.38 km/s and a a period of 117 minutes, completing about 12.7 orbits in a Martian day. It will cross the sunlit equator in a southward direction at 2 pm local time, and successive passes will be separated by 28.4 degrees of longitude. The nominal mission lifetime is 1 Martian year (687 days).

The primary objectives are: to identify and map surface elements and minerals, to measure the surface topography and features, to measure the gravitational and magnetic fields, to study the structure and circulation of the atmosphere, and to determine the distribution, abundance, sources, and destinations of volatile material and dust over a seasonal cycle.

The spacecraft has a three-axis stabilized platform on which the instruments are mounted; there is no scan platform. The instruments in the payload are: a high-resolution Gamma Ray Spectrometer (GRS) to measure the abundance of K, Fe, U, Si, Th, and other elements on and under the surface, including the permafrost; a magnetometer (MAG), to measure the very weak Martian magnetic field; the Mars Observer Camera (MOC), with wide and narrow fields of view with surface resolution up to 1.4 m; a Laser Altimeter (MOLA) to map the surface topography; the Radio Science experiments (RS), which as part of the telecommunications link will measure gravitational anomalies and atmospheric absorption; a Thermal Emission Spectrometer (TES), an infrared Michelson interferometer to measure the thermal IR emissions from the surface rocks and frosts and the atmospheric clouds, and the Pressure Modulator Infrared Radiometer (PMIRR), to measure atmospheric temperature, water vapour density and dust and condensate distribution.

2.2 The Pressure Modulator Infrared Radiometer (PMIRR)

The PMIRR instrument has been built by the Jet Propulsion Laboratory in Pasadena, California; the Department of Atmospheric Physics (now the sub-department of Atmospheric, Oceanic and Planetary Physics) at Oxford has supplied detectors, optics and pressure modulators.

PMIRR is designed to determine the abundance, distribution, sources and sinks of volatiles and dust over one Martian year. It will measure the state of the atmosphere for each Mars Observer orbit, mapping the temperature to 80 km altitude, and water vapour abundance to 35 km, on a daily basis, in a similar manner to terrestrial satellite observations. From the temperature, the thermal wind, transport and atmospheric stability will be inferred. PMIRR will also make radiometric measurements of other condensates and dust abundance. It

will monitor the polar radiative balance and the variability of atmospheric pressure.

PMIRR uses the high effective resolution of the pressure modulator radiometer (PMR) technique [Taylor 1984], to reduce the effect of the atmospheric dust, with its large and variable optical depth, on the measurement of temperature and water vapour. This method, which involves the insertion of cells of gas in the optical path inside the instrument, will be described in outline later in this chapter and more fully in the rest of this thesis.

Channel number	Bandpass cm^{-1}	central wavelength μm	channel type	measurement function
1	1226–1420	7.6	PMR (CO_2) + WB	temperature 0–20km
2	719–739	13.7	PMR (CO_2) + WB	temperature 20–50km
3	678–702	14.5	PMR (CO_2) + WB	temperature 50–80km and pressure 30–50km
4	1408–1500	6.9	PMR (H_2O) + WB	water vapour 0–35km
5	821–873	11.8	WB	) dust and condensates
6	470–506	20.5	WB	) 0–80 km;
7	280–334	32.6	WB	) radiative balance
8	184–230	48.3	WB	)
9	2116–25510	0.7	WB	)

Table 2.I: PMIRR channels, as built [J.T.Schofield, private communication]. In channels 1–4, PMR denotes a subchannel in which the optical path includes a pressure modulator and the signal is synchronously demodulated at the pressure modulation frequency. In the wideband (WB) channels the optical path still includes the PMU but the pressure-modulated sidebands are ignored.

The filter bandpasses of PMIRR's channels are given in table 2.I. Most view fairly narrow IR bands, while channel 9 has a broad bandpass which extends into the visible region. Channels 1–4, the temperature and water vapour sounding channels, (see figure 2.1) are located in the minima of dust opacity at 15 and 7 μm [Pollack *et al.* 1979], where the opacity is around one tenth of its visible-region value.

2.2.1 Scanning and Resolution

PMIRR mostly scans the limb and occasionally the nadir. Scanning is performed by a two-axis-mounted mirror providing nadir, off-nadir and limb views of the planet and atmosphere, and a space reference view. The normal scanning mode is arranged to produce alternate along-track limb and nadir views. The limb scans are broken into twelve 5 km steps in vertical tangent height, each observed for 2 s. The detector fields-of-view are in three groups scanned together across the limb, to provide a total vertical coverage of -20 to 80km in tangent height. The temperature-sounding channels are arranged so that channel 1, which uses a band of an uncommon isotope of CO_2 to enhance the low-altitude measurements, scans from -20 to +30 km, while channel 2 is scanned from about 0 to 60 km; channel 3 is still higher in altitude. After the limb scan, a single sounding is performed for a nadir view. The entire cycle takes 28 seconds, which, when combined with the 3.38 km/s spacecraft velocity, provides a 95 km horizontal separation along the track between sounding cycles.

The advantages of limb viewing are threefold. For nadir sounding the vertical resolution is limited essentially to about one scale height by the weighting function; for limb viewing

Figure 2.1: PMIRR spectral channel locations compared with cumulative Planck functions, the spectral characteristics of Mars atmospheric components, and a typical Mariner 9 IRIS spectrum [J.T.Schofield, private communication].

the resolution is determined by the field of view (if the field of view is greater than about 3 km), and a resolution of half a scale height can be obtained. Secondly, because the atmospheric path is longer for a limb view the atmospheric emission reaching the spacecraft from an optically thin emitter is much greater. Thirdly, the instrument does not view the surface so that the small changes in radiance in the atmosphere are not masked by large surface emission variations.

The PMIRR detectors have a field of view of 3.1 by 15.6 milliradians, which maps to a 5km vertical by 25km horizontal atmospheric region for a limb view and a 1km (along track) by 5 km (across track) footprint for a nadir view. Along the line of sight for the limb view the resolution is about 240 km. Some vertical responses (weighting functions) of the PMIRR temperature-sounding channels for different limb-scanning tangent heights are shown in figure 2.2.

Figure 2.2: Typical PMIRR limb-scanning weighting functions for the CO_2 channels, each for four different scan mirror positions [from McCleese *et al.* 1986]

Figure 2.3: Optical layout of PMIRR (schematic) [from McCleese *et al.* 1986]. The chopper is reflective on both sides, and the CO_2 modulator is used for three channels simultaneously. This diagram does not show the 300 K blackbody calibration target for the IR channels, or the diffusely scattering calibration target for channel 9.

2.2.2 Optics, detectors and cooler

The optical layout is shown schematically in figure 2.3. The telescope is a 6cm aperture $f/11.9$ Gregorian design. Radiation from Mars reaches first the flat scan mirror mounted on a short boom, then the primary, and then the chopper alternately to one of two secondary mirrors. The scan mirror has orthogonal elevation and azimuth stepper motors, each with a resolution of 0.045 degrees, which corresponds to half of the field of view of each detector. A two-point radiometric calibration is provided by a 300K blackbody occasionally introduced in front of the prime focus of the telescope, in combination with the space view against which the atmospheric radiation is chopped.

All the optical components are maintained at 300K, except the channel 1–5 bandpass filters (at 140K), and the HgCdTe detectors and condensing optics which are mounted on a single focal plane assembly, and cooled to 80K. Because of the low heat load of the detectors on the cold focal plane assembly and the need to reduce the power consumption of the instrument a passive (radiative) cooler is used in preference to an active (closed-cycle) one. It has a protective cover that is used to prevent contamination and to act as a planet shield when at Mars. The cooler has been built by the Santa Barbara Research Center (SBRC).

Channels 1–5 use HgCdTe (MCT) detectors mounted on the cold focal plane. These were made by Philips in the UK and packaged, focussed and aligned at Oxford. The detectors for channels 1 and 4 are photovoltaic; in these channels the limiting noise contribution is from the instrument background radiation. The others are photoconductive, because of their long wavelength, and the detector noise dominates. The focussing lenses are integral with the detector mount, and the detector focussing and pointing are set by means of shims whose thickness is calculated by an iterative technique. The angular responsivity map of the packaged detectors was measured at Oxford using a Globar and pinhole which could be positioned independently in three dimensions. Channels 6–9 use deuterated triglycine sulphate (DTGS) detectors at 300K. The noise in these channels is again dominated by the detectors.

2.2.3 Electronics and signal processing

The signal chain contains phase-sensitive-amplifiers to resolve the 800Hz (chopper frequency) component of the detector signal and (in the case of the pressure-modulator channels) its sidebands at about $\pm$45Hz. The output from either is digitized by a counting analogue-to-digital converter (ADC) using a linear voltage-to-frequency convertor and a 16-bit counter. This technique has good linearity, and because it integrates over time, it reduces the noise compared with an instantaneous measurement. The output from the counter is read out every 2 seconds, and the counter is then reset; this defines the effective integration time.

2.3 Pressure Modulators

Pressure Modulation Radiometry is an example of 'gas correlation spectroscopy' – the use of the spectral lines of a gas under investigation as a species-specific filter. It was developed at Oxford and has been used extensively in balloon- and space-borne remote sounders. These include the PMR on *Nimbus 6*, SAMS on *Nimbus 7*, and VORTEX on the *Pioneer Venus* Orbiter. The Upper Atmosphere Research Satellite (UARS) includes the Improved Stratospheric and Mesospheric Sounder (ISAMS), which has 8 pressure modulators, in its payload.

Figure 2.4: Pressure modulated lineshape [from McCleese *et al.* 1986]. The PMIRR H_2O pressure modulator transmission is compared with the atmospheric emission (for a path with 4 km tangent height, using the COSPAR standard model temperature profile and a 0.05% mixing ratio of H_2O) for the water vapour line at 1375.0879 cm^{-1}.

Figure 2.5: Spectral selectivity of the PMIRR CO_2 PMR [from McCleese *et al.* 1986]. The pressure modulator transmission is compared with the atmospheric emission (calculated for 5 km tangent height, using the COSPAR model temperature profile and a uniform H_2O mixing ratio of 0.1%. This figure shows that the CO_2 PMR largely rejects radiation from atmospheric water vapour.

2.3.1 Principle of operation

The pressure modulator unit (PMU) causes the pressure in an optical cell in the instrument to oscillate, modulating the amount of optical absorption by the spectral lines of the gas in the cell. The major effect at typical pressures is periodic variation of the linewidth, so that the transmission of the gas in the cell varies between two extremes, as shown in figure 2.4, which shows two typical calculated transmission extremes, T_{P_1} and T_{P_2}, and their difference, $T_{P_1} - T_{P_2}$. A signal is obtained by demodulation of the detector signal synchronously with the pressure-modulation. The resulting transmission forms a filter with passbands in the wings of every spectral line of the gas. The pressure modulator mechanism is described in chapter 3, and a computer model of its thermodynamic behaviour in chapter 5.

2.3.2 Sensitivity and selectivity

Because only the lines of the gas in the modulator cell are transmitted, the modulator channels are insensitive to emission from other species (see figure 2.5), except for the small overlap between the spectral lines of different gases. Thus if the cell contains a gas whose mixing ratio is known to be constant in the atmosphere, so that its emission is mostly a function of the temperature alone, the pressure modulator signal can be used to derive the atmospheric temperature distribution. If the distribution in the atmosphere of the pressure-modulated gas is not known, it may be inferred from its emission (measured by the pressure modulator channel) and the atmospheric temperature in the source locality deduced from a temperature-sounding channel. Thus in PMIRR the CO_2 channels are used to derive the global temperature field and also to give the temperature input into the water vapour retrieval.

The retrieval of pressure depends on knowledge of the spectroscopy of the wideband and pressure-modulated signals, and of the spacecraft attitude. It is expected that PMIRR will be able to retrieve pressure in the 30–50 km region to about 1.5 %; pressure at other levels will be inferred from the temperature profile, using the hydrostatic equation.

In principle the spectral sensitivity of a PMR is limited ultimately by the linewidth of the modulated lines. Because all the lines of the gas under modulation are transmitted, and because the high spectral resolution does not require a small étendue (energy grasp), the radiation incident on the detector, though small, is greater than that obtainable from a conventional spectrometer of similar resolution. This compensates for the loss in signal arising from the absorption in the cell.

Figures 2.4 and 2.5 show that the atmospheric emission lineshape is similar in form to that of the pressure modulator, because at low altitudes where the pressure is high, the spectral lines are broad. At higher altitudes (between the tangent height and the radiometer) the pressure drops so that the spectral lines are narrower. If the temperature at high altitude is lower than that below, the molecules high up are net absorbers over the narrower spectral region of their linewidth; if the temperature increases with height new narrow lines may be seen. In the main therefore the atmospheric emission profile is of a broad line with the line centre removed by upper atmosphere absorption.

2.3.3 Doppler shift

In a limb-scanning instrument, account must be taken of the possibility and effects of Doppler shifting of the incoming atmospheric radiation. This will occur if the spacecraft velocity has a component along the line of sight, but is only important if the magnitude of

the Doppler shift is not much smaller than the spectral resolution of the instrument or of the atmosphere, whichever is the greater. In the case of PMIRR, which views the aft limb, the resolution is determined by the pressure modulators. The spacecraft velocity along the line of sight leads to a Doppler shift for emission at a frequency ν of approximately $1.3 \times 10^{-5}\nu$, or about 10^{-2}cm^{-1} for channels 2 and 3, and $1.64 \times 10^{-2}\text{cm}^{-1}$ for channels 1 and 4.

Because of the large amount of CO_2 in the Martian atmosphere, the atmospheric CO_2 lines are very broad, especially at low altitude. In channel 1 the atmospheric line width (FWHM) for the strongest lines, which are responsible for most of the emission, are typically 0.12 cm^{-1}, much greater than the Doppler shift. In channels 2 and 3, the atmospheric lines are about 0.04cm^{-1} wide at about 40km altitude; the lines in the pressure modulator cell are about 0.02cm^{-1} wide, so again the Doppler shift is a small effect. For water vapour, however, the atmospheric lines are quite narrow, with widths of about 0.02cm^{-1} at 30km. The lines in the water vapour pressure modulator cell are about 0.025cm^{-1} wide, so that the Doppler shift of 0.0164 is significant. Its effect is to add a small high tail onto the PMR weighting function, changing the detector signal by a few percent [J.T.Schofield, private communication].

2.4 PMIRR CO_2 pressure modulator

The carbon dioxide pressure modulator is used simultaneously for temperature sounding in three channels (channels 1–3), each with different altitude ranges. The three channels are separated from each other by field-splitting mirrors placed after the modulator in the optical path, and go to separate detectors. CO_2 is used for temperature sounding because of its uniformity throughout the Martian atmosphere. Its spectroscopy is well known and it has strong bands near the peak of the Planck function appropriate for atmospheric temperatures.

For moderate (20-50 km) altitudes the wing of the 15 micron ν_2 $^{16}O^{12}C^{16}O$ (referred to as '626') band is used, and for high altitude measurements (above 15km), the centre of the ν_2 band. Low altitude sounding is particularly difficult for two reasons. CO_2 forms 95% of the Martian atmosphere, so the CO_2 amounts in a limb path of low tangent height (below 20 km) are very large, and the spectral lines saturate. To overcome this problem PMIRR uses the weak fundamental ν_1 band near 7 microns of $^{16}O^{12}C^{18}O$ ('628'), which has an isotopic abundance of about 0.4 percent in the atmosphere. Thus the optical depth for each spectral line is much less than that for a normal ('626') isotope band. Because this band is forbidden for symmetric isotopes, it is not obscured by emission from the more abundant isotopes.

The choice of this passband also addresses the other problem at low altitude, that of dust blown up from the surface, particularly during global dust storms. The dust extinction is smallest around the 7 μm channel 1 passband (see figure 2.1), and the spectral discrimination of the PMR will also reduce the sensitivity to dust. Because of the variations in the dust amount, PMIRR retrievals of temperature are designed to have as small a dependence on the dust as possible.

The CO_2 pressure modulator unit (PMU) contains carbon dioxide at 80mb mean pressure, with a 2:1 compression ratio in a 4 cm long optical cell with coated germanium windows. The high pressure is required because of the breadth of the emission lines from the Martian atmosphere; the high compression ratio increases the pressure-modulated transmission in

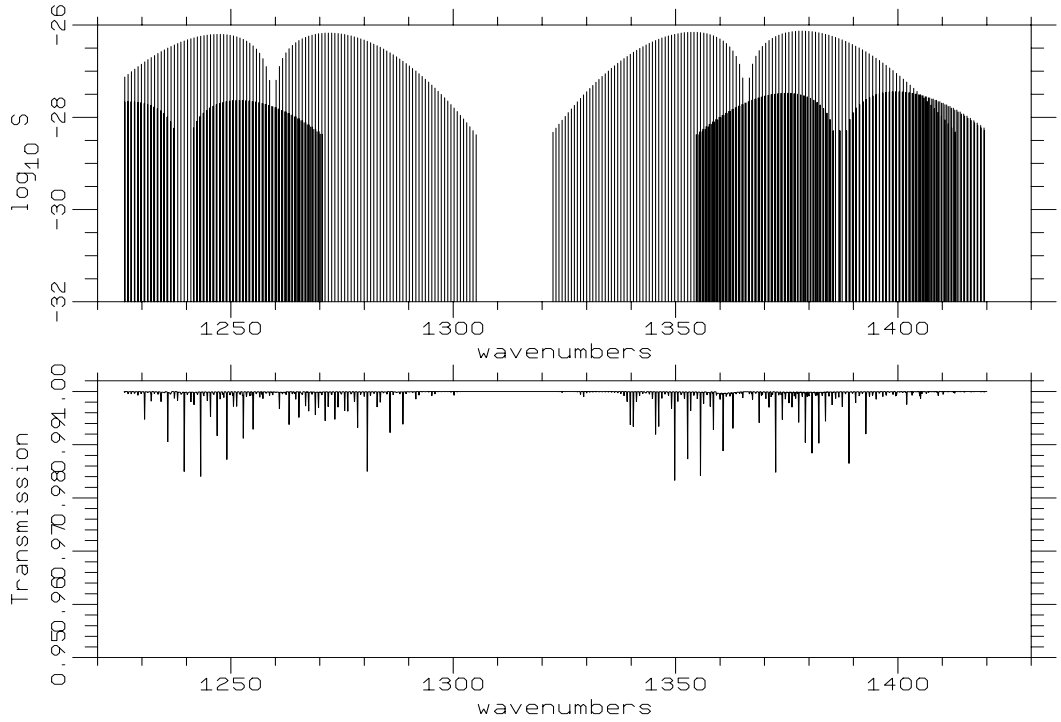

Figure 2.6: Wideband transmission of the PMIRR CO_2 modulator in the channel 1 passband (1226–1420 cm^{-1}). The '628' isotopic lines are very weak, so that the peak absorption is small. Line strengths, plotted here in units of $cm^{-1}/(molecule\ m^2)$, have been corrected for the isotopic abundance in the cell.

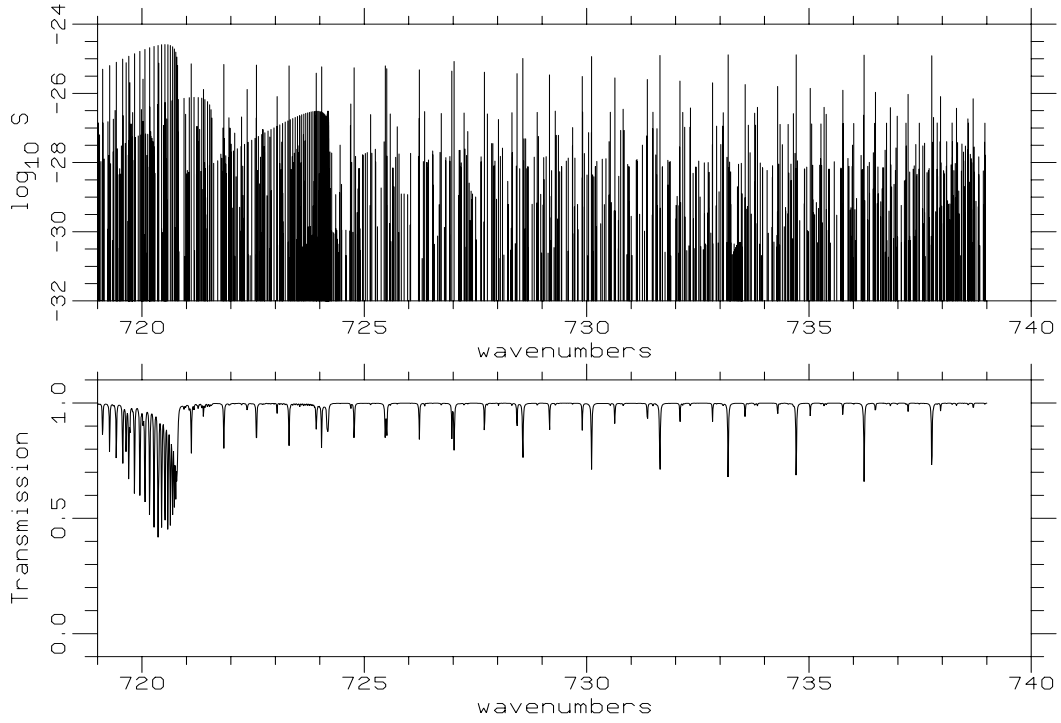

Figure 2.7: Wideband transmission of the PMIRR CO_2 modulator in the channel 2 passband (719–739 cm^{-1}). Line strengths, plotted in units of $cm^{-1}/(molecule\ m^2)$, have been corrected for the isotopic abundance in the cell.

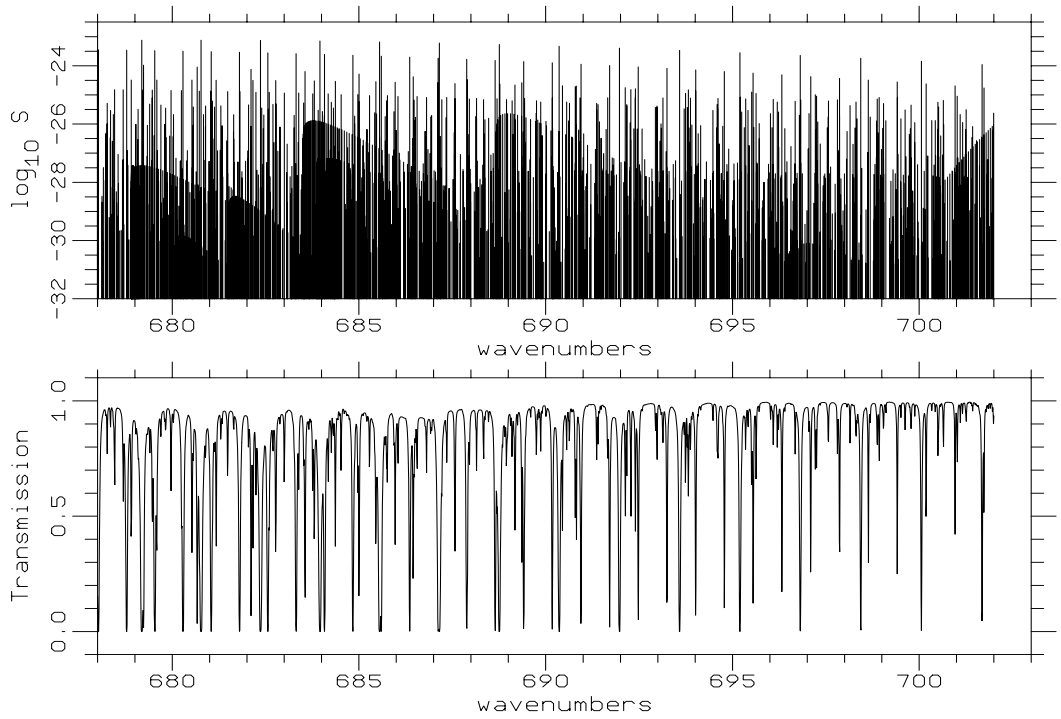

Figure 2.8: Wideband transmission of the PMIRR CO_2 modulator in the channel 3 passband (678–702 cm^{-1}). Line strengths, plotted in units of cm^{-1}/(molecule m^2), have been corrected for the isotopic abundance in the cell.

the near wings of spectral lines, and thus signal-to-noise ratio of the modulator channels. The radiance emitted by the atmosphere in channel 1 is small, as a result of the low atmospheric abundance of the asymmetric isotope. In the radiometer, the signal component from this band is enhanced by increasing the abundance of the '628' isotope in the pressure modulator cell. $^{18}O^{12}C^{18}O$ ('828') is also present as the result of oxygen exchange, as well as traces of other isotopes. The relative abundances for 626:628:828 isotopes were intended to be in the ratio 4:4:1. The transmission of the cell is shown in figures 2.6–2.8.

2.5 PMIRR H_2O pressure modulator

The major reasons for using a PMR to measure H_2O are to discriminate it from the large CO_2 signal and to overcome the measurement problems associated with the aerosol constituents of the atmosphere (which adversely affected the *Viking* MAWD), as for CO_2.

The water modulator channel (channel 4) at 6.9 microns, in the near wing of the ν_1 vibration-rotation band, and near the dust extinction minimum, like channel 1, has the same spatial resolution as the temperature measuring channels. Knowledge of the temperature of the source region is necessary to determine the water vapour abundance from the emission measured in the water vapour channel.

The water vapour pressure modulator, built to the same mechanical design, has a 10cm long optical cell, also with coated germanium windows. The mean pressure is 16.8 mb, and the compression ratio is about 1.3. The pressure was chosen to be as large as possible without saturation occurring, thus broadening the spectral lines in the PMU as much as possible, yet preventing the formation of liquid water which might cause an unknown

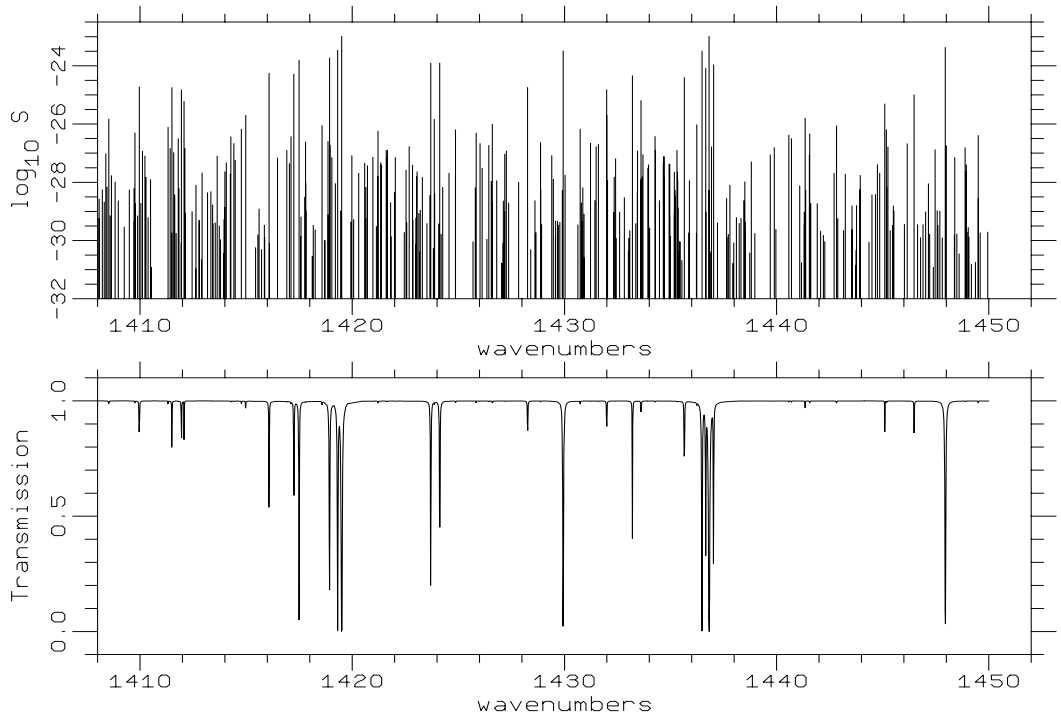

Figure 2.9: Spectral line intensities (in units of $cm^{-1}/(molecule\ m^2)$) and wideband transmission of the PMIRR H_2O modulator in the first half of the channel 4 passband (1408-1450 cm^{-1}).

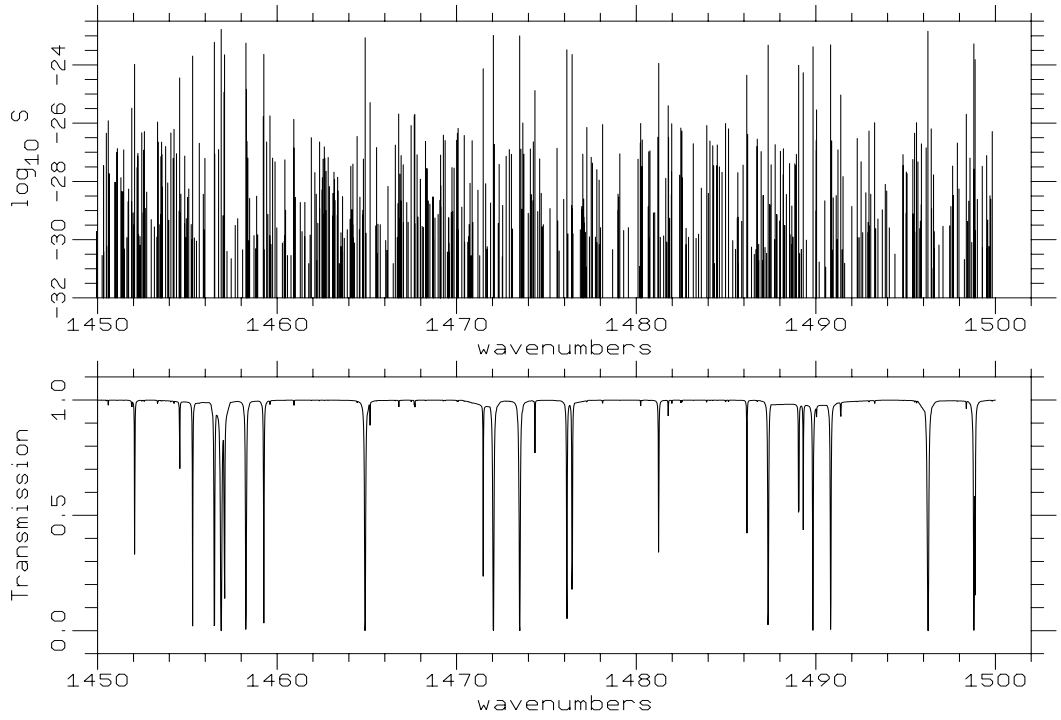

Figure 2.10: Spectral line intensities (in units of $cm^{-1}/(molecule\ m^2)$), and wideband transmission of the PMIRR H_2O modulator in the second half of the channel 4 passband (1450-1500 cm^{-1}).

absorption in the cell. The physical processes involved in the cycling are discussed in chapter 5. There is a trade-off between cell length and compression ratio to provide the best signal-to-noise ratio, because a longer cell has a larger volume and thus decreases the compression ratio, although the amount of gas is increased. For PMIRR, use of the long cell increases the width of the lines in the cell, so that they are closer to the widths of the atmospheric emission lines. The transmission of the cell is shown in figures 2.9 and 2.10.

2.6 PMIRR aerosol and surface measurements

PMIRR has eight spectral channels between 6.5 and 50 microns to measure aerosols, with the same spatial resolution as the temperature channels. The radiance detected in these channels comes from thermal emission by the atmosphere, and both thermal emission and scattering by aerosol particles.

At 11.8 μm (channel 5) and 20.6 μm (channel 6), dust and water ice both have absorption bands (see figure 2.1). At 31.7 μm (channel 7) gaseous absorption is minimal and so the signal is determined almost entirely by the aerosols; this channel is used as a reference for total line-of-sight opacity. A long-wave channel at 46.5 μm (channel 8), having a low sensitivity to dust and condensate particles with the distribution expected in the Martian atmosphere, is used as a zero-reference for the other aerosol channels. The aerosol channels, and the intercomparison between the pressure modulated and wideband signals (channels 1–4), provide, in combination, discrimination between dust and condensate aerosols.

PMIRR is also used to return information about the surface of Mars. The nadir and off-nadir brightness temperatures in the 6.8 to 50 μm region are measured simultaneously with the broadband solar reflectance from a 0.3–3 μm channel (channel 9), which accepts over 97% of the solar flux and will be used to construct daily polar albedo maps. The brightness temperatures at 31.7 μm (channel 7) and 46.5 μm (channel 8) are used to detect surface frost by comparison with the solid-vapour phase transition temperature at the surface pressure.

2.7 Expected performance

A general expression for the energy F incident on a detector in a radiometer has been given in Houghton, Taylor and Rodgers [1984]. An equivalent form is

$$F = A\Omega \int f_\nu \left(B(\nu, T_0)\bar{\tau}(p_0) + \int B(\nu, T)\frac{d\tau_\nu}{dx}dx \right) d\nu \tag{2.1}$$

in which $A\Omega$ is the energy grasp or étendue of the instrument, f_ν is the transmission at frequency ν of the optical path inside the instrument for light reaching the detector, and B is the Planck function. The first term in the bracket represents black-body radiation from the ground at temperature T_0, transmitted by the atmosphere with transmission $\bar{\tau}_0$. This is zero for a limb path. The second term represents radiation from the atmosphere itself, according to Schwarzchild's equation of transfer,

$$\frac{dI}{d\chi} = I - B \tag{2.2}$$

(in which χ is the optical path (see Appendix A) and I is the local intensity) integrated along the path of light (line of sight) reaching the instrument; dx is an infinitesimal section of that path.

As the Planck function varies little with frequency across the bandwidth of a channel, it may be approximated by a constant and taken outside the integral. Furthermore the values of τ_0 and $d\tau_\nu/dx$ may be averaged over the channel (weighted by the transmission profile f_ν), to separate the atmospheric and instrument behaviour.

Table 2.II: PMIRR radiometric performance for channels 1–4 [from McCleese *et al.* 1986]. Raised numbers denote exponents. PMR denotes the pressure modulator channel, WB the wideband channel. Other terms are defined in the text.

2.7.1 Signal to noise ratio

The expected radiometric performance of the four pressure-modulator channels is summarized in table 2.II. The 'noise equivalent radiance', NEΔR is the spectral radiance change which would produce a signal equivalent to the detector noise equivalent power (NEP):

$$\text{NE}\Delta\text{R} = \text{NEP} \times 2^{\frac{3}{2}}/(4/\pi\, A\Omega\, W\bar{f}\sqrt{2t}) \tag{2.3}$$

Similarly NEΔT is the temperature change which would produce the same signal:

$$\text{NE}\Delta\text{T} = \text{NE}\Delta\text{R}/(dB(\nu_0, T_0)/dT_0) \tag{2.4}$$

the signal-to-noise ratio (SNR) is

$$\text{SNR} = B(\nu_0, T_0)/\text{NE}\Delta\text{R} \tag{2.5}$$

The minimum detectable temperature is that temperature for which the signal-to-noise ratio is unity:

$$T_{\min} = B^{-1}(\nu_0, \mathrm{NE}\Delta\mathrm{R}) \tag{2.6}$$

In these expressions, W is the equivalent width, *i.e.,* the integrated transmission, $\bar{f}$ is the mean transmission of the channel, and t is the signal integration time. ν_0 is the central frequency of the channel; the instrument is assumed to be viewing a blackbody at temperature T_0.

2.7.2 Calibration

PMIRR's pre-flight and onboard calibration is intended to provide an absolute radiometric accuracy of 0.5% in the thermal infrared channels and 3% in the broadband (albedo) channels. Extensive modelling of the response of PMIRR to Martian conditions considered typical in the light of the Viking results and dust models shows [McCleese *et al.* 1986] that PMIRR should be able to measure temperature to 2K below 80km, dust extinction to 20% below 50km, and dust column abundances to better than 5%. The vertical water vapour structure will be resolved below 35 km. Column abundances of water will have an accuracy generally better than 10%. The high vertical resolution of PMIRR is necessary for these degrees of accuracy, and since the temperature is required for the retrieval of abundances, the accuracy is greatest at about 30km, (near the lower end of the altitude range for channel 3) where the pressure is about 0.32 mb, the temperature about 180K and the water vapour mixing ratio 1 ppm [McCleese *et al.* 1986].

The sensitivity of a channel to temperature or mixing ratio change is determined by modelling the atmosphere, for example using the COSPAR model, calculating the transmission of the channel, and using equation 2.1 to give the detector output. The atmosphere may then be perturbed by 1K in temperature or the mixing ratio by 10% and the change in radiance noted. Variations in the radiance will also occur when different values of spectral line parameters (Appendix A) and instrument transmission are used. The latter arise from uncertainties in optical components, and in the line data for pressure modulator channels. By comparing these radiance changes, the instrument calibration requirements may be determined.

For a limb path, a number of simplifying assumptions may be made. The atmosphere may be treated as isothermal, so that the path-integral of $d\tau/dx$ becomes the emissivity ε, and the mixing ratio of a trace gas may similarly be treated as constant. With these assumptions, the radiance change in channel 4 arising from a 1K increase in temperature or a 10% increase in water vapour mixing ratio is about 9%. Uncertainties in the pressure modulated transmission (calculated from the model which will be described in chapter 5), neglecting errors in the line data, cause a change of about 1% in the wideband signal and 0.5% in the PMR signal, demonstrating that the uncertainties from the pressure modulator cycle itself do not limit the quality of the measurements. (The most significant limiting factor is expected to be the retrieval technique [J.T.Schofield, private communication].) However, a systematic error in the line strengths of 15%, as suggested by J.J.Remedios [private communication] causes a change in atmospheric emitted radiance of about 12%, which is transmitted through to both wideband and pressure modulated signals.

2.7.3 Experimental verification of H_2O transmission

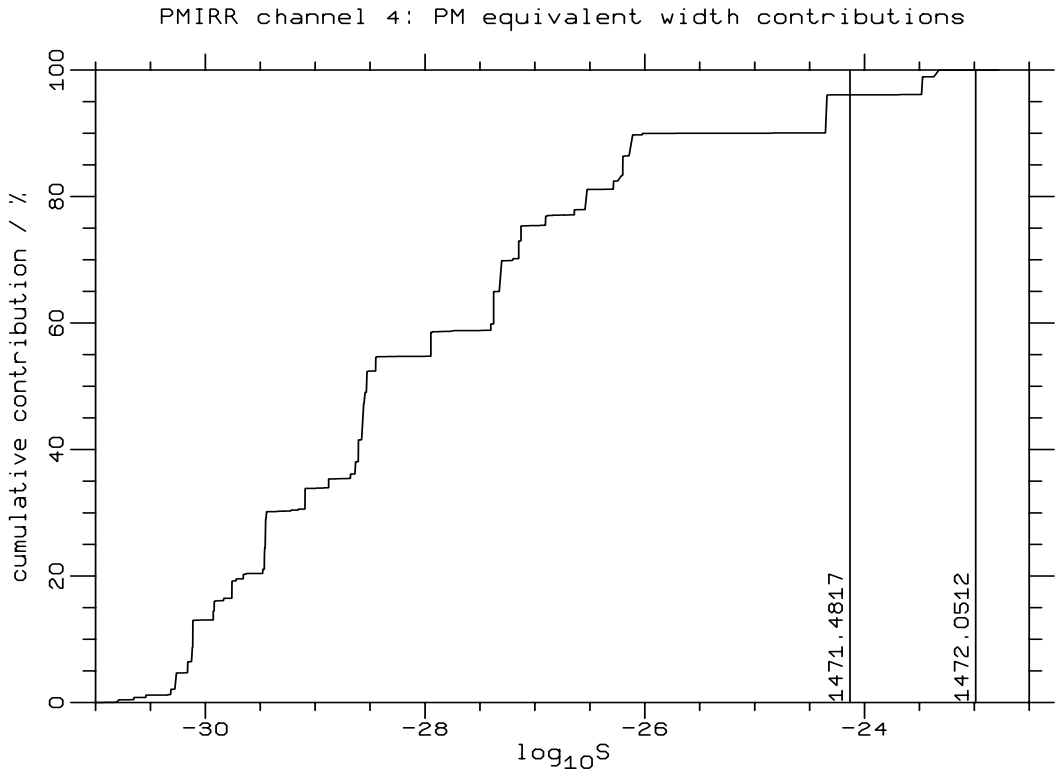

Figure 2.11: Relative contribution to the total equivalent width for the channel 4 pressure modulator channel, from lines of different strengths (in units of $cm^{-1}/(molecule\ m^2)$). The two strongest lines measured in the tunable diode laser experiment are marked.

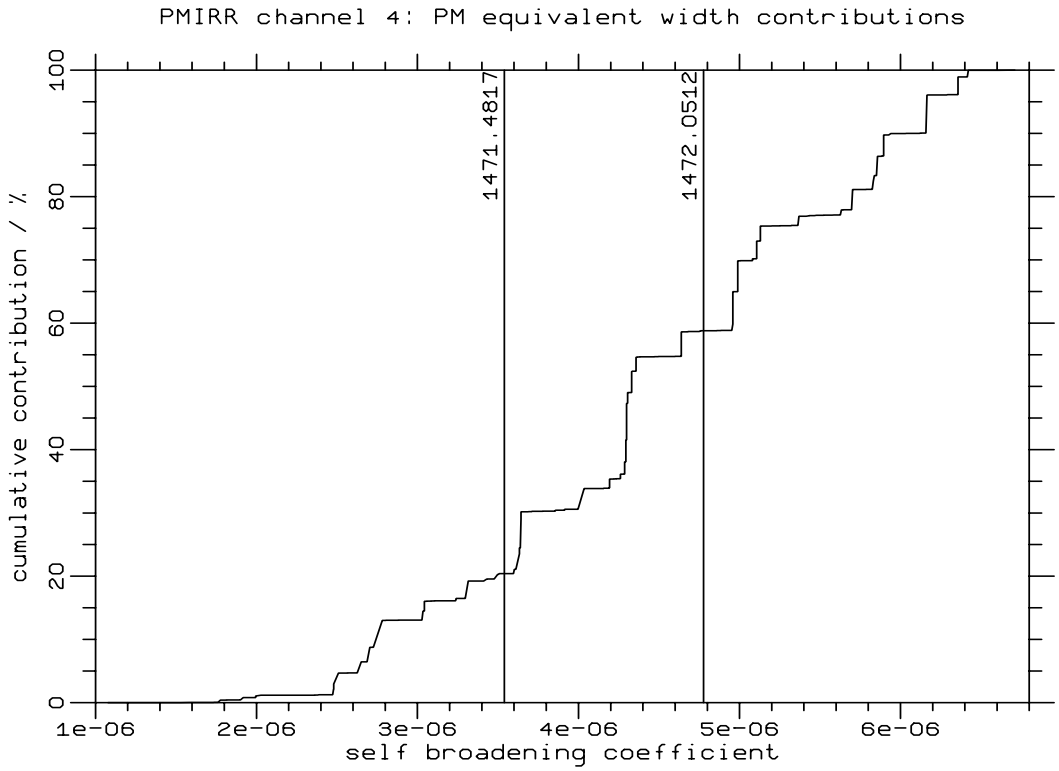

Figure 2.12: Relative contribution to the total equivalent width for the channel 4 pressure modulator channel, from lines of different self-broadening coefficient (in units of cm^{-1}/Pa). The two strongest lines measured in the tunable diode laser experiment are marked.

Data from the water vapour on the *Nimbus 7* SAMS instrument have been difficult to interpret [Mutlow 1984, Munro 1991]. The present thesis is particularly directed towards characterization of the PMIRR H_2O modulator by thermodynamic modelling, pressure cycle measurements, and high-resolution lineshape measurements, to ensure that similar difficulties in interpretation are not encountered in the PMIRR data.

Figures 2.11 and 2.12 show how the equivalent width contributions are divided among the lines in the passband of channel 4. These contributions were derived by calculating the pressure modulated transmission of the whole channel, and then removing each line in turn and recording the change in transmission.

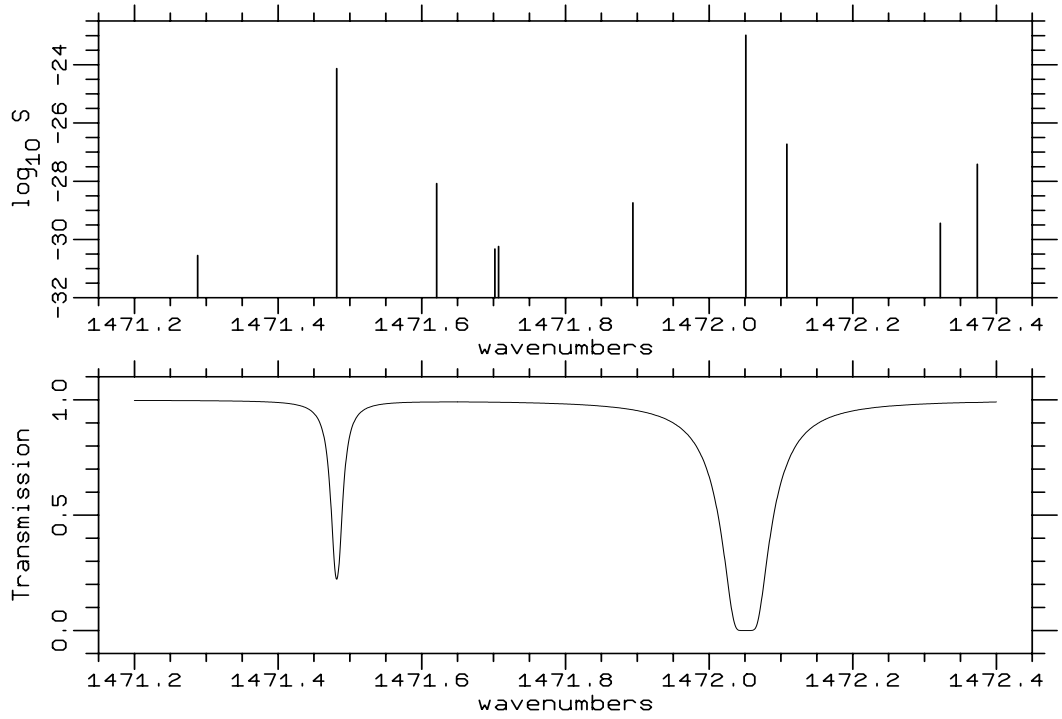

Figure 2.13: Spectral line intensities and wideband transmission of the PMIRR H_2O modulator, over the spectral interval used for high-resolution lineshape measurements. It is evident that the weaker lines make little contribution to the absorption of the cell in this short region.

The region from 1471.2 to 1472.4 cm^{-1} was used for the spectroscopic measurements described in chapter 6. There are two strong lines in this region, at 1471.4817 and 1472.0512cm^{-1}, which dominate its wideband absorption (figure 2.13), and pressure modulated transmission, and therefore had a good SNR in the tunable diode laser experiment. However, it is evident from figure 2.11 that weaker lines make a significant contribution to the equivalent width, and thus to the radiometric signal in this channel. This implies that broadband transmission measurements, over the whole passband of the channel, are also required. Such measurements have been made for PMIRR, using both CO_2 and H_2O in a pressure modulator, by Irwin [1991].

Chapter 3:
PMIRR Pressure Modulators

This chapter describes the mechanical design, filling with gas, and characterization of the pressure modulator units (PMUs) used in the PMIRR instrument. The design is the result of many years of development of pressure modulators since their conception in the 1960s. It is very similar to the design used on ISAMS, since the PMUs for the two instruments were built almost simultaneously. However, due to different functional requirements, and to remove some of the difficulties experienced in assembling early ISAMS modulators, several important changes were made which are recorded here. One aspect of this concerns the springs. Since both ISAMS and PMIRR use double-ended modulators to remove almost all of the momentum imbalance, the springs used to suspend the two pistons must be matched in strength. Another important change described is the increasing of the efficiency of the linear motors in order to reduce the power consumption of the two pressure modulators.

The filling procedure, including the method by which the isotopically-enhanced CO_2 was made, is recorded, and the mechanical characteristics of the PMIRR flight-model pressure modulators, measured before delivery to JPL, are described.

3.1 Design and Construction

The mechanism used to effect the modulation of cell pressure in the PMIRR instrument is similar to that used in ISAMS (figure 3.1), which has eight pressure modulators containing seven different gases to measure temperature and species abundances for Earth. Two shaped pistons, mounted on spindles suspended on diaphragm springs, are made to oscillate inside a titanium cylinder by electric linear motors which have static coils and moving permanent magnets; one magnet is mounted on each spindle, and the coils are fixed to the outside of the cylinder. On the end of each spindle is a soft iron slug, which runs inside a thin tube on a flange which forms the end of the cylinder. Outside the tube a differential transformer, which uses the slug as a movable core, is used as a position-pick-off (PPO), described later in this chapter, to sense the position of each piston. The PPO output is used to control the motor coil current by a drive circuit which will be described in chapter 4.

Gas from a reservoir (which in the spacecraft takes the form of a thermostatted aluminium cylinder containing silicalite, used as a molecular sieve) is introduced into the pressure modulator through a flange which communicates with the region behind each piston; it passes around each piston through a small gap about 75 μm wide. The viscosity of the gas effectively seals a gap of this width at the operating frequency. From the region between the pistons, a tube allows easy passage of gas to an optical cell. Oscillatory motion of the pistons causes the pressure in the region between them to oscillate, and gas to move to and from the cell, so that the pressure in the cell is modulated.

Early pressure modulators used a single piston to compress and rarefy the gas; because

Figure 3.1: The balanced pressure modulator, from a drawing produced at the Rutherford Appleton Laboratory. This figure depicts an ISAMS PMU, which differs from the PMIRR design in a few respects described in the text.

PMIRR, like ISAMS, is limb-scanning, the momentum compensation provided by a balanced modulator (with two opposed pistons) was essential to reduce motional chopping. For this purpose both the masses and the spring strengths in the two halves of the modulator must be matched.

As described in chapter 2, two pressure modulators are used on PMIRR, so that two flight model PMUs have been supplied, and two flight spares PMUs prepared. A prototype (non-flight) PMIRR PMU ('brassboard') was also built to lower specification, from surplus ISAMS parts, but using PMIRR CO_2 coils. It has been used to evaluate the drive circuit and for most of the spectroscopic measurements which are described later in this thesis.

3.1.1 Detail changes from ISAMS design

Plate II is a photograph of a disassembled (ISAMS) modulator. The cell is smaller than either of those used for PMIRR, and the two end flanges are different (one has two ports on it), whereas the PMIRR PMUs have identical end flanges and a single gas port in the middle of the body.

The clearance between the pistons and the liner was increased from 50 µm to 75 µm in order to reduce the probability of the piston rubbing against the liner, which leads to an

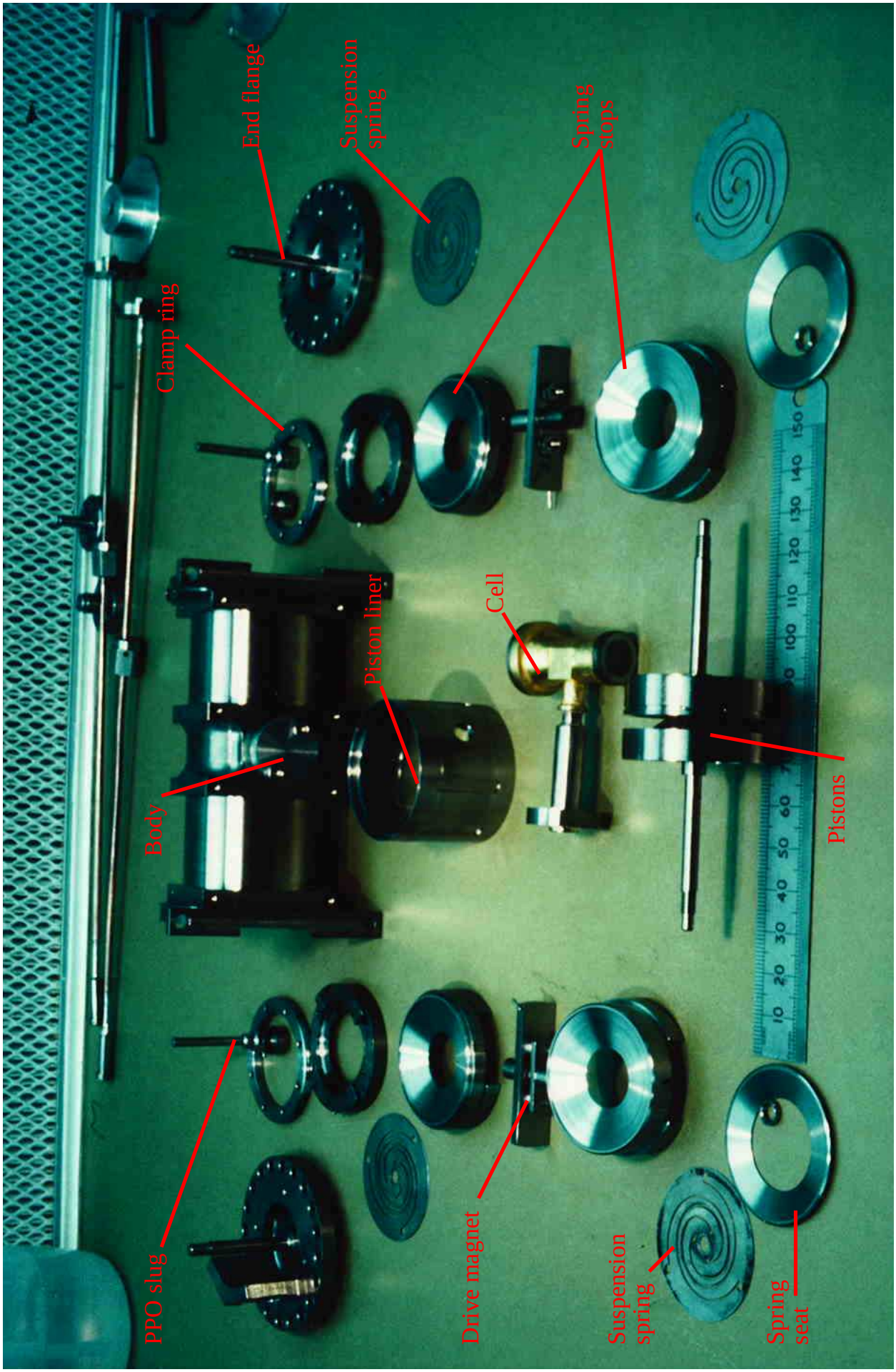

Plate I: Pressure Modulator Unit parts

increase in power consumption, an imbalance in momentum, and the possible failure of the PMU. At the higher pressures in PMIRR (80 mb for CO_2, 16.8 mb for H_2O), the sealing effect of the viscosity of the gas is more pronounced, so that the gap width could be increased without significantly increasing the leak past the pistons.

The spring material and manufacturing technique was slightly different. Both of these last changes are further described below. The plane of the drive coils and magnets was made perpendicular to that of the cell, to simplify mounting; the ISAMS optical bench had to be slotted to accommodate the tops of the coils.

3.1.2 Improvement of drive motor efficiency

Because of the low power available to drive the PMIRR PMUs, it was necessary to increase the efficiency of the motor for the CO_2 modulator, which because of its high gas pressure has a large power consumption. Measurements of the DC and dynamic resistance of the drive coils using a sine-wave (analogue) drive showed that the overall efficiency of the ISAMS pressure modulator motors was 13.8%.

The motor efficiency could have been greatly increased by the use of a moving-coil motor, with tight magnetic coupling between the stator and motor. However it was apparent after some preliminary work that not enough time would be available to re-design the PMU to incorporate a moving-coil motor. (A moving-coil design was rejected for ISAMS because of the risk of chemical reactions between the reactive gases and the coil insulation, and of a possible gas leak into the modulator through the coil feedthrough seal). Therefore the efficiency of the moving-magnet motor was improved by increasing the strength of the moving magnets, by using two ISAMS-type magnets on each spindle, and by increasing the size of the drive coil windings. Space constraints in the instrument limited this to a certain extent, and prohibited an increase for the water vapour modulator, so that, whilst the CO_2 modulator coils were larger, the coils on the H_2O PMU were eventually the same length as ISAMS coils. However, the power consumption of the H_2O PMU is very small so that the extra efficiency would have been of little benefit.

3.1.2.1 Efficiency by coil design

It may be shown that the efficiency of the drive coils is dependent only upon the core geometry, and not upon the wire diameter. The force F acting on the piston is proportional to the flux density B produced by the drive coils and thus to the current I and number of turns N. It is constant in magnitude for a given gas pressure and piston amplitude, because the mechanical system experiences frictional and gas-cycling power losses which do not depend on the drive coil efficiency. Therefore

$$F = \kappa B = \lambda N I \tag{3.1}$$

where λ, the coupling coefficient, is a geometrical constant.

The coil resistance is

$$R = \rho N \bar{c} / a \tag{3.2}$$

in which ρ is the resistivity, $\bar{c}$ the mean winding circumference, and a the cross-sectional area of the wire. The number of turns is determined by the geometry:

$$N = g \times A / a \tag{3.3}$$

in which A is the area available for winding (perpendicular to the wires), and g is a the proportion of that (rectangular) area occupied by the conductors. So

$$< I^2R >= \frac{F^2}{\lambda^2 N^2}\rho N\frac{\bar{c}}{a} = \frac{F^2\rho\bar{c}}{\lambda^2 gA}, \tag{3.4}$$

which is independent of a. To reduce the mean dissipated power, therefore, it is necessary to increase A, g or λ or decrease ρ or $\bar{c}$. In the case of the PMU drive coils, g, $\bar{c}$ and λ are fixed by the geometry. ρ cannot reasonably be decreased, so only A can be altered; thus the legs of the core must be made as long as possible.

3.1.2.2 Core saturation

Concern about whether the efficiency might be decreased further by saturation in the cores, which is more likely for the larger coils, may be dismissed on the following argument. The saturation flux density of typical ferrites is around 0.3 T [Kaye and Laby 1973]. Consideration of the core geometry gives a gap length l of about 10 mm. The number of turns on a core is about 3900. Therefore

$$\frac{B}{I} = \frac{N\mu_0}{l} \approx 0.5 \text{ T/A} \tag{3.5}$$

Thus to achieve a flux density of 0.3 Tesla, a current of 0.8 A (peak) must flow. This would dissipate 8 W (mean) per core, or 32 W (mean) for the whole modulator, assuming that the DC component of the drive current is small in comparison with the AC component. As the power budget is much less than 32 W, the coils cannot reach saturation.

3.1.2.3 Ratio of inductance to resistance

Increasing the size of the drive coils also increases the ratio of inductance to resistance, which affects the drive circuit design. The phase shift between the voltage and current in the drive coils is

$$\Delta\phi = \tan^{-1}\left(\frac{\omega L}{R}\right) \tag{3.6}$$

The resistance of the coils is

$$R = \rho\bar{c}gA/a^2 \tag{3.7}$$

and the inductance is

$$L \approx N^2\mu_0 A_2/l_g \tag{3.8}$$

where A_2 is the cross-sectional area of the core. Substituting the value of N from the previous section,

$$L/R = \frac{g^2A^2}{a^2}\frac{\mu_0 A_2}{l_g}\frac{a^2}{\rho\bar{c}gA} = \frac{gAA_2\mu_0}{l_g\rho barc} \tag{3.9}$$

This increases with A, so that the phase-shift between the voltage and current in the drive coils increases with coil size. For two ISAMS coils the ratio is 3.08 ms, whereas for two PMIRR CO_2 coils it is 5.87 ms. At a typical PMU frequency of 40 Hz, therefore, the phase shift in the H_2O coils is 37.7°, while for the longer CO_2 coils it is 55.9°. This phase shift caused a loss in driving efficiency from running off resonance in a voltage-driven configuration (such as in the ISAMS drive circuit), and affected the drive circuit power amplifier design (see chapter 4) because of the large back EMF generated by the high inductance.

3.1.3 Springs

The pressure modulator springs must fulfil two objectives: firstly they must not break over the mission lifetime, and secondly they must be uniform, so that pairs of equal spring strength can be chosen. Because the springs are the only part of the pressure modulator assembly subjected to repetitive stress, they are expected to be the limiting factor on the mechanical life of the PMU.

Differences in either the masses or the spring strengths on the two sides of the PMU can cause momentum imbalance when the pressure modulator is running. This oscillating net momentum can rock the instrument or bend the optical bench at the PMU frequency, at which the detector signal is demodulated. Either of these kinds of 'motional chopping' lead to systematic errors in the measured response of the instrument and so must be reduced to a minimum. No explicit balance criterion was imposed by JPL, so I specified that the difference between the amplitudes of the two pistons should be less than 3% of the sum of their amplitudes at the 3mm peak-to-peak nominal piston amplitude. This figure was based on the experience of assembling the ISAMS modulators, for which the imbalance specification was 10%, but which in practice were unbalanced by less than 5%. It is simple to remove the mass contribution to the imbalance, by removing material until the two sides weigh the same, whereas the spring strengths are more difficult to match because of the wide variation in the strength of the diaphragm springs as produced, and the high sensitivity to this form of imbalance. Venters [1991] has shown that the PMU is approximately ten times more sensitive to a spring imbalance than to the same proportional mass imbalance.

The balance was measured using an oscilloscope, dividing the amplitude of the difference between the two PPO traces by the amplitude of the sum trace. However, the difference trace is very noisy as a result of pickup from the drive current, and from the different carrier frequencies in the two PPOs. Furthermore the gains of the oscilloscope probes and input amplifiers vary with time and temperature, and the tolerance of the input resistors in the drive circuit is 1%. As a result, the total uncertainty in the balance as measured is about 2%. A measured imbalance of more than 3%, above which it is just possible to feel the vibration of the pressure modulator, led to the modulator springs being changed.

3.1.3.1 Spring manufacture

The springs for PMIRR modulators were originally chosen from those left over from ISAMS use, but it became apparent that there were not enough to produce the required matched spring pairs, so that new springs had to be made.

The suspension springs were made from 301 stainless steel strip, cold-rolled to a hard surface finish. The cold-rolling increases the life expectancy without affecting the value of Young's modulus for the material, because the plastic deformation produced in the rolling process reduces the cross-sectional area, while the yielding force remains the same, so that the yielding strain is increased. In the rolling process, the steel strip develops a natural curvature. To ensure proper spring seating, the radius of curvature must be much greater than the diameter of the modulator.

The steel for the ISAMS springs was taken from sample sheets with 530 Vickers hardness provided by the Nippon Steel Company. However, the company had ceased production of this thickness of steel when more was required for PMIRR. After much searching, British Steel Stainless Precision Strip were found to produce slightly softer (about 450 Vickers)

stainless steel of the right thickness and composition, and with a curvature radius of about 1.3 m.

The steel was photo-etched by the Physics Photography Unit with the diaphragm spring pattern, and the resulting springs were turned to their final shape and fitted to the spindles and seats as required. Corresponding springs on the two sides of the PMU were matched in spring constant to balance the modulator.

3.1.3.2 Spring life testing

Because of the change of materials for PMIRR and in particular the reduced hardness, which might decrease the expected lifetime of the springs, it was necessary to demonstrate that the new springs would outlive the mission. The number of cycles of repetitive stress causing failure in a ferrous metal becomes infinite if the stress is below a certain threshold [Curtis 1974]. The relationship between the stress and the number of cycles to failure is such that failure occurs after 100 million cycles if the repetitive applied stress is approximately 1.25 times the threshold stress. Thus if some sample springs withstand a repetitive stress (and consequently motional amplitude) 25% greater than the maximum anticipated in the PMUs, for 10^8 cycles, then that maximum anticipated stress is less than the threshold, and the springs used in the real modulators should last indefinitely.

To perform a life test on some sample springs, I built a rig which clamped two springs, connecting their centres with a spindle on which a reflective disc and a magnet were mounted. Two opto-switches were turned on alternately as the disc moved from one extreme of travel to the other. The switches controlled a J-K flip-flop which switched the direction of the current through a coil axially placed around the moving magnet.

The test rig ran at an amplitude of 5 mm, to allow for a 25% over-stress over an upper-limit design amplitude in the PMU of 4 mm. This amplitude was set up by making the reflective disc 5 mm thick and changing the coil current until the top edge of the disc on the downward stroke was seen to coincide with the bottom edge on the upward stroke. The frequency was about 40 Hz, so the test took about a month to complete 10^8 cycles. As neither spring failed, the material was considered adequate for use in the pressure modulators.

3.1.3.3 Measurement and selection

The tension-to-extension ratio (spring constant) k of the springs was measured indirectly. Each spring was clamped in turn to a lead block, and a small brass bob and small magnet were attached to the centre. A coil, driven by a variable oscillator, excited the bob-spring system, and the resonant frequency was measured. The mass of the spring is much smaller than that of the bob, so that variations in mass between springs could be neglected, and the frequency could be used to derive k. The frequencies are measured several times, with the springs mounted each way up, to reduce variations from inconsistent clamping.

Generally, the two spring orientations resonated at different frequencies. I suggest that this is because as the curvature of the spring (produced by the motion) passes the natural curvature of the material, (produced by the rolling process), the spring constant k changes, so that the spring has a different strength on each side of the natural curvature. The change occurs at a small but non-zero spring displacement; therefore the effective spring constant for oscillatory motion is a weighted average of the two values of k, with the weights

determined by the proportion of time that the spring spends each side of its natural curvature.

In the testing rig the weight of the bob introduces an initial displacement or 'set' into the spring and therefore biases the motion to occur more in one k regime than the other (in fact as the amplitude of the motion is so small for this measurement all the motion may be on one side of the natural curvature). When the spring is turned upside-down the weights are altered, so that the effective spring constant changes. In the modulator, however, the motional amplitude is larger, so that both spring constants must be taken into account, and full modelling of the spring behaviour must include this intrinsic non-linearity. As the amplitude increases, the time spent each on side of the natural curvature changes, so that the relative weights when averaging the two spring constants also alters, the result being that the effective spring constant changes with amplitude.

Springs for the flight modulators were chosen so that matching springs on opposite sides of the PMU had similar values of k (derived from the frequency measurements) for both the inward and outward motion of the pistons. The frequencies measured for the PMIRR springs were found to be much more uniform than those made for the ISAMS modulators, tending to very similar values for each orientation. This probably occurred because the ISAMS springs were lightly etched (without a pattern) in order to make them clean enough to take the spiral spring pattern; this etching introduced random variations in the thickness of the spring material. The material used for the PMIRR springs could be satisfactorily etched with the spiral pattern after cleaning with solvent alone. Most of the modulator imbalances were almost certainly caused by mishandling of the springs, such as often occurs when a modulator is disassembled. Any irreversible (non-elastic) bending causes the spring constant to change dramatically, and the spring, and probably its matched pair, must be rejected.

3.1.4 Position Pick-offs (PPOs)

As mentioned above, the position of each piston spindle is sensed from outside the PMU body by a position pick-off. The PPO circuit, consisting of a 3 kHz oscillator, a transformer with multiple secondary windings, and a demodulator, is shown in figure 3.2. The position of the slug on the end of the spindle is nearly proportional to the difference between the amplitudes of the outputs from the two secondary windings of the transformer. The difference and sum signals are detected by synchronous sampling at the carrier frequency. The amplitude of the carrier is kept constant by servoing the sum of the two secondary outputs to a fixed reference voltage (nominally 10 V). The carrier signal breaks through slightly into the output, and is visible on an oscilloscope, but is removed in the pressure modulator drive circuit by the input filter.

The electronics are assembled on two printed circuit boards (PCBs), between which is a fibreglass tube along which the transformer can move. This tube incorporates a mumetal screen which reduces the sensitivity of the PPO to external magnetic fields, which are generated by the pressure modulator drive motors, for example. The differential design of the circuit further reduces this sensitivity. The PPOs are placed in housings on the end flanges of the pressure modulators, and the bobbins are screwed against individually-made shims to give zero output when the pistons are at rest.

The flight PPOs were made to the ISAMS specifications, with 3 kΩ output resistors. The ISAMS PPO output resistance had been changed (without documentation) to 10 kΩ. The

Figure 3.2: PMIRR PPO circuit, designed by S.T.Werrett. Q1, the transformer, and associated components form an oscillator; IC1 detects the sum and difference signals from the transformer secondaries; Q3 and Q4 form sample and hold circuits, triggered by IC2. IC3a controls the amplitude of the carrier frequency oscillations, and IC3b is an output buffer.

gain of the PPOs is set to 0.5 mm/V by adjusting a resistor. As this gain was measured using a circuit with an input resistance of 17 kΩ and a feedback resistor of 27 kΩ, the PMIRR PPO gain was smaller by 1.35 than the ISAMS PPO gain.

3.2 Filling and characterization of flight PMUs

In order to verify that the PMUs were working correctly, to provide a conversion between frequency and pressure, to interpret the data from the PMIRR instrument, measurements of maximum attainable amplitude, amplitude imbalance between the two pistons, and operating frequency at different pressures and amplitudes were made while assembling, filling and operating the flight and brassboard PMUs.

Because of the compressibility of the gas in the pressure modulator, there is an additional element in the 'spring constant' of the modulator motion, which increases with pressure; thus, as the running frequency varies as $\sqrt{k/m}$, the frequency is expected and observed (see below) to increase with pressure. In the PMIRR instrument, the frequency at which the modulator runs will be used to measure the pressure within it, the frequency having been measured during filling, as a function of pressure and piston amplitude. The frequency wanders slightly with time, as a result of sieve pressure variations arising from thermostat excursions, so approximate sieve adsorption time constants were measured to give an indication of the period required for the frequency to stabilize. The power consumption was also measured, to facilitate electronic design.

The composition of the gas in the modulator is required to predict the effective filter response in the instrument, and so a mass spectrometer was used to measure the residual

gases in the flight modulators, and to demonstrate that unwanted molecules formed a negligible fraction of the final gas mixture. The composition of the isotopically-enhanced CO_2 was calculated from the ratio of the isotopes used to make it, and measured spectroscopically.

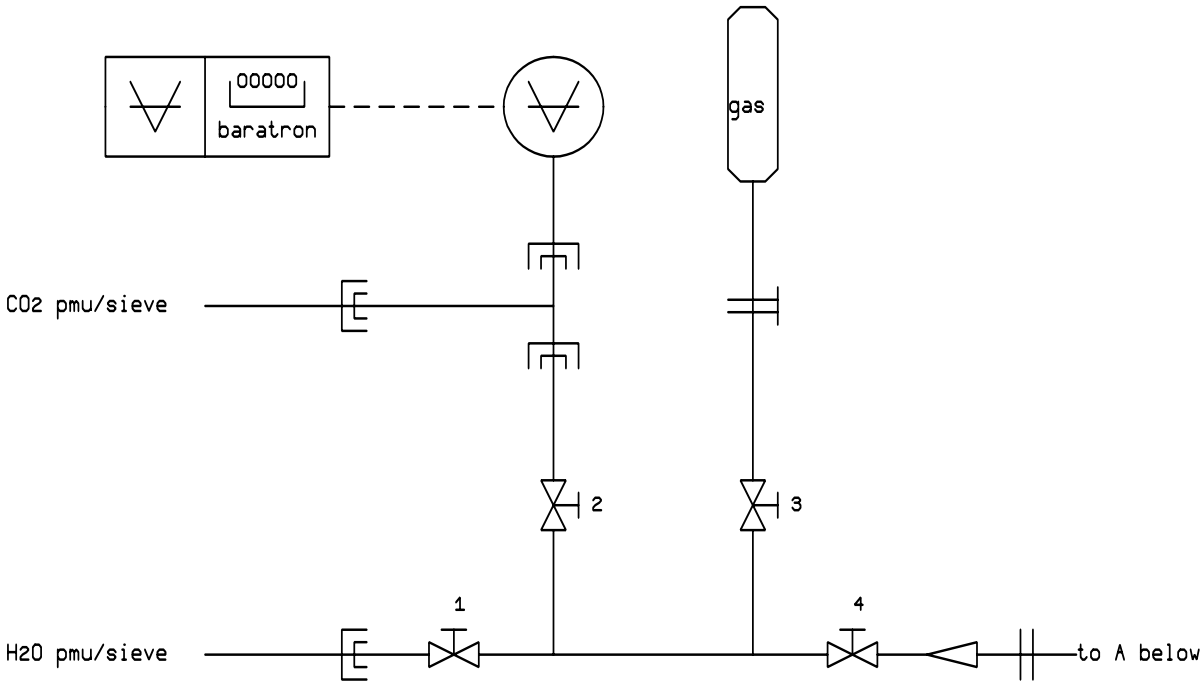

Figure 3.3: PMIRR PMU filling rig: high pressure part. Symbols are as defined in DIN 28 401 Standard specification. The 100 torr baratron has a resolution (determined by the control unit) of 0.01 torr below 20 torr, and 0.1 torr for greater pressures. Control of the amount of admitted gas is achieved by filling the region between all the valves from the bottle, closing V3, and opening V1 or V2 as appropriate.

3.2.1 Filling rig

I designed and built a filling rig (figures 3.3 and 3.4), to bake the flight PMUs and sieves, to manufacture the flight CO_2 isotopic mixture, and to fill and calibrate the flight PMUs. Because only a limited amount of the 4:4:1 isotopic CO_2 was available, the volume of the filling rig was kept to a minimum. The rig uses a small (50 litre/second maximum pumping rate) Pfeiffer turbomolecular pump which is very clean and can produce pressures in

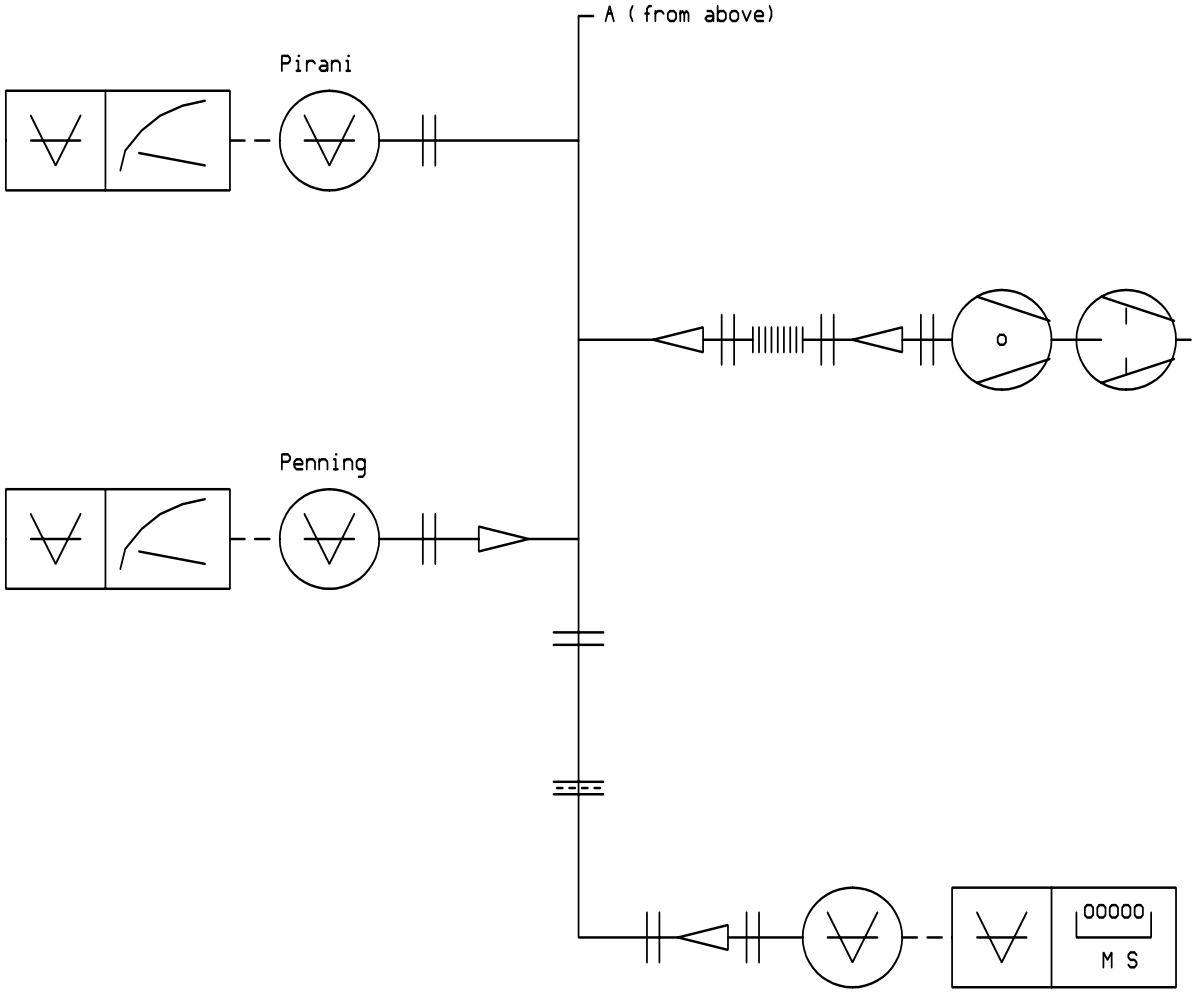

Figure 3.4: PMIRR PMU filling rig: low pressure part. Symbols are as defined in DIN 28 401 Standard specification. MS signifies a quadruple mass spectrometer used for residual gas analysis.

the region of 10^{-8}mb in a small, clean, leak-tight vacuum system. It has four ports which can be isolated, joined at a central cross. These are connected to the pressure modulators, baratron, gas bottle and vacuum pump. The vacuum part of the filling rig has ports for a low-vacuum Pirani gauge, a high-vacuum Penning gauge, and a mass spectrometer head. The rig is connected to the vacuum pump by a 1 metre long by 25 mm diameter stainless steel bellows.

The flanges for the pressure-modulator attachments are the Cajon VCR 2 type. Aluminium gaskets were used to maximize the lifetime of the gland surfaces, and to provide a soft seal which is clean and tight. The flange for the gas bottle is a 6-hole bolted flange, designed at the Rutherford Appleton Laboratory, with a gold 'O' ring seal. The baratron flange is a Cajon VCR size 8 flange, also sealed with an aluminium gasket. The whole rig is wound with Hotfoil heating tape powered from a Variac, adjusted to heat it to 160°C in order to remove the water vapour from the walls.

The CO_2 PMU was filled first, so that the baratron could be used for the H_2O filling, and to reduce the outgassing from water adsorbed onto the rig walls, which would necessitate a long bakeout before filling the second PMU.

3.2.2 Manufacture of the flight CO_2 mixture

The CO_2 in the PMIRR flight PMU is intended to have a [626]:[628]:[828] ratio of 4:4:1 to enhance the amount of the asymmetric isotope for channel 1 (see chapter 2). At equilibrium the CO_2 molecules all have random isotopic labelling so that the probability of any one oxygen atom being ^{18}O is 1/3 and that of two oxygen atoms both being ^{18}O is 1/9. Similarly the probability of 626 is $(2/3)^2$ or 4/9, and that of 628 is $2 \times (2/3) \times (1/3) = 4/9$ as there are two possible orientations of the isotopes. Thus an $[^{16}O]$:$[^{18}O]$ ratio of 2:1 yields the correct 4:4:1 ratio of [626]:[628]:[828]. This equilibrium ratio is obtained slowly by atomic exchange, but can be encouraged by using a molecular sieve, as chemisorption onto the large surface area catalyzes the isotope exchange. The composition of the gas mixture is measured spectroscopically after manufacture (see below).

To ensure that enough gas was made to fill the flight CO_2 PMU, it was attempted to make enough to produce a pressure of 160 mb over a cold sieve, thus providing twice as much gas as is required for one flight fill of 80 mb. The available gases were from a bottle of 99.999% pure research grade CO_2, of which (by normal isotopic abundances) 98.42% is 626 and a negligible amount is 628 and 828, and a bottle of high (but unknown) purity CO_2 having an isotopic abundance of ^{18}O of 93.2%.

The ^{18}O-enriched CO_2 was applied first, and its pressure was allowed to stabilize, and then the normal CO_2 was added in controlled amounts, to achieve a total pressure appropriate for the desired isotope ratio. Because of the isotope ratios and purities already in the source gas bottles, the actual gas amounts from each gas bottle deviate slightly from the 2:1 ratio. If the ^{18}O-enriched CO_2 is added first, to a pressure of p_{18}, then the total pressure after the normal CO_2 is added should be

$$p_{tot} = 2.825 p_{18} \tag{3.10}$$

After this has occurred, the ratio of abundances of the two oxygen isotopes is

$$\frac{^{16}O}{^{18}O} = 1.056 \frac{p_{tot}}{p_{18}} - 0.983 \tag{3.11}$$

and the ratio of the CO_2 isotopic abundances is

$$[626] : [628] : [828] = [16]^2 : 2[16][18] : [18]^2 \tag{3.12}$$

where [16] is the abundance of the ^{16}O isotope as a fraction of the whole.

The sieve was filled with the '828' CO_2 to a pressure of 61.81 mb, and then the ordinary CO_2 was added, to a total pressure (at equilibrium) of 174.55 mb, so that the final ratio of abundances was (from 3.11)

$$^{16}O/^{18}O = 1.999(02) \tag{3.13}$$

$$[626] : [628] : [828] = 0.43519(74) : 0.43549(64) : 0.10895(13) \tag{3.14}$$

$$\text{unaccounted} = 0.0204(15) \tag{3.15}$$

where the figures in brackets are the uncertainties in the last two figures, assuming that the baratron accuracy is 0.027 mb. The real errors may be larger than this because of uncertainty in how the gas has settled onto the sieve.

The gas was mixed on the sieve to speed the equilibrium, heated up to 80 °C, and left to cool overnight. Nine hours later the sieve was heated up again to 80 °C, while the receiving bottle was immersed in liquid nitrogen. 20.03 mb was left on the sieve and PMU when the bottle was isolated. Most of this residual gas was pumped off, leaving 4.09 mb for the first calibration run. The gas remained because of the large effective volume of the sieve compared with that of the 50ml bottle, and the fact that at 77K the temperature of the gas in the bottle is about a quarter of that in the sieve. The vapour pressure above solid CO_2 at 77K is negligible (<0.017 μbar).

A sample of the mixture so made was analyzed spectroscopically at the Rutherford Appleton Laboratory by R.J.Knight and P.G.J.Irwin. The measurements showed that the isotope exchange was not quite complete, but that one more sieve temperature cycle resulted in the random labelling required.

3.2.3 Filling procedure

The PMUs and sieves were pumped while heated to 60°C and 100°C respectively, to remove the water vapour, until the total pressure recorded by the mass spectrometer fell to about 10^{-6} mb. For each modulator a residual gas mass spectrum was recorded after pumping and baking.

The molecular sieves were then thermostatted to 305K, and the gas was gradually introduced into the modulators, in a series of pressure steps. At each pressure the PMUs were run at different amplitude settings, and the frequencies and power consumptions were measured. The modulators were initially overfilled, to obtain frequency calibrations on both sides of the operating pressures. Large pressure increments (about 13 mb for CO_2 and 6 mb for H_2O) were used at low pressures, and smaller steps (to about 1.3 mb for CO_2 and 0.6 mb for H_2O), near the nominal filling pressures, to reduce the amount of gas transfer to or from the sieve, and to provide a finer grid for the final frequency/pressure PMU measurements near the operating pressure. Throughout filling, the pressure and sieve temperatures were registered on a chart recorder. About an hour was allowed between adding (or removing) gas and making frequency measurements, so that the gas reached equilibrium with the molecular sieves, and the pressures stabilized accordingly.

At the filling pressures, the modulators were left running overnight, to ensure that the systems were at equilibrium. Final sets of frequency measurements were made, and the modulators were run at a moderate amplitude for half an hour, until stable frequencies were obtained. The copper tubes between the PMUs and the filling rig were then crimped, sealing the PMU–sieve systems. This process displaced some gas and disturbed the equilibrium distribution of gas between the PMUs, sieves and filling rig, causing an increase in frequency. However, the frequencies subsequently decreased to within 0.01 Hz of the value before crimping, and remained constant for more than thirty minutes, so that there appeared to be no leak in either modulator. A further set of calibration measurements was performed later, when the gases had resettled onto the sieves. Similar measurements have been made periodically at JPL to verify that the modulators continue to function properly and do not leak.

3.2.4 Sieve adsorption time constant

Chart recorder traces of the pressures during filling after an increment showed that the pressure followed a roughly exponential approach to equilibrium, with a time constant of 350–450s for CO_2 and 620–1000s for H_2O. The ranges of these figures are due to the inaccuracy in measuring by this method. However, they show that the PMU–sieve system should be left to reach equilibrium for about an hour after pressure changes, before making measurements. As gas is moved between the sieve and PMU during a drive amplitude or sieve temperature change, it is important to allow adequate time for equilibrium to be reached during filling and calibration.

3.3 Characteristics of PMIRR pressure modulators

Property	CO_2 PMU	H_2O PMU	Brassboard
Filling pressure (mb)	80.3 (0.1)	16.81 (0.01)	N/A
Maximum amplitude (mm)	3.1 (0.05)	3.15 (0.05)	2.68 (0.05)
Measured imbalance at 2.9 mm(%)	0.48 (2)	0.48 (2)	4.0 (2)
Power consumption at 2.9 mm (W)	2.85 (0.1)	0.79 (0.1)	N/A

Table 3.I: PMIRR PMU characteristics. Figures in brackets are uncertainties.

3.3.1 Imbalance and attainable amplitude

The imbalance (defined as the difference of the measured PPO amplitudes divided by their sum), maximum attainable amplitude and final filling pressures of the two flight PMUs, and the brassboard modulator, are shown in table 3.I. The final filling pressure for water vapour (16.81 mb) was lower than specified in the PMIRR proposal (18 mb), as a result of the limited time available for filling the modulator and the long settling time for water in the sieve and modulator.

The imbalances in the flight model modulators were much less than the 3% upper limit described above, and the minimum attainable amplitude requirement (3 mm peak to peak for each piston) was also met. We recommended an operating amplitude of 2.90 mm peak to peak for each piston in flight, providing a safe margin in case of degradation in performance of the modulator or drive circuit, an external impulse, or a power transient.

3.3.2 Frequency versus pressure characteristics

The measured frequency *versus* pressure and amplitude characteristic curves for the flight model H_2O PMU (at 300K) is shown in figure 3.5. Similar curves for the brassboard modulator, at 295 and 298K, are shown in figures 3.6 and 3.7. The variation of the system frequency with pressure at the low body temperature of 295K confirms results described by G.Davis [1987]; the frequency does not increase monotonically with pressure, but increases to a maximum, decreasing at higher pressures. This occurs in the region of 15 mb for the brassboard PMU at 295K, and at about 21 mb at 298K. The cause of this effect will be discussed in chapter 5. It is evident that the use of the system frequency as a measure of mean PMU pressure requires careful consideration at high pressures of

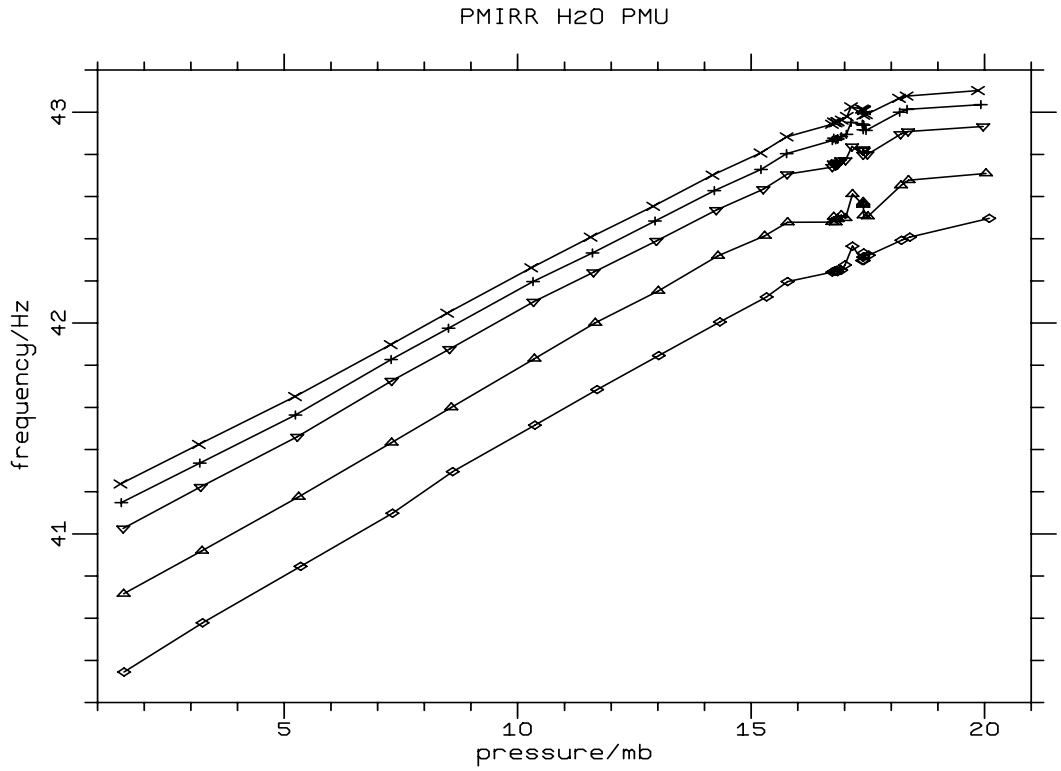

Figure 3.5: Frequency *versus* pressure characteristics measured at five demanded piston amplitudes (1.07, 2.12, 2.76, 2.97 and 3.10 mm) for the PMIRR flight model H_2O PMU at 300K. The frequency increases with amplitude at all pressures. The modulator was initially filled to about 20 mb and then water vapour was removed to achieve the final pressure. These data are collated from measurements taken in both pressure directions. The 'kinks' near 17 mb arise from molecular sieve thermostat cycling.

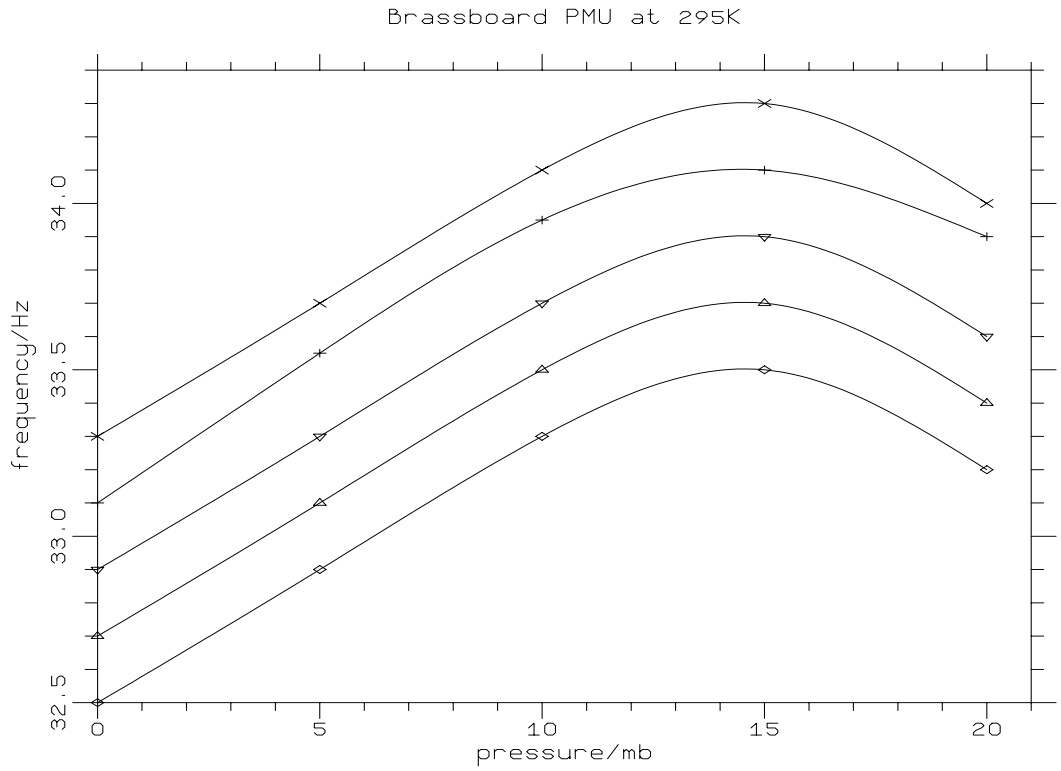

Figure 3.6: Frequency *versus* pressure characteristics for five demanded piston amplitudes (1, 1.5, 2, 2.5 and 2.7 mm) measured for the brassboard PMU, at a body temperature of 295K, containing H_2O. The frequency at each pressure increases with amplitude. The frequencies are lower than in the flight model case, because of the extra mass of the double magnets. The maxima in these curves, absent in figure 3.5, are a result of the low body temperature. The effect is discussed in chapter 5.

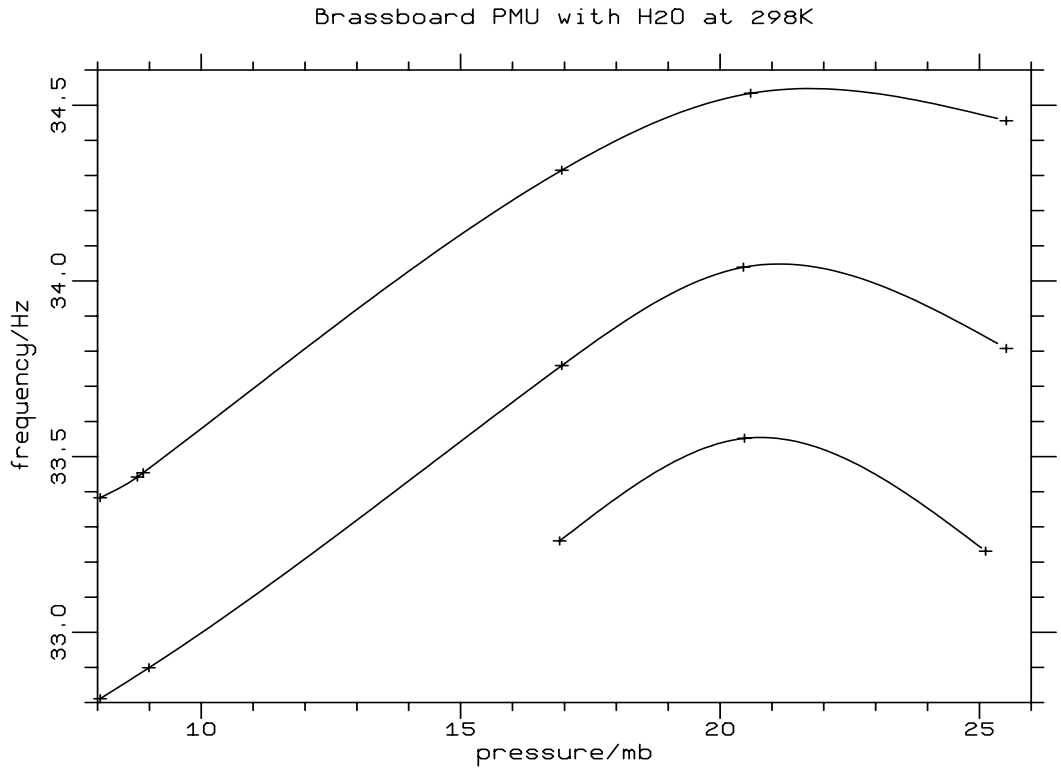

Figure 3.7: Frequency *versus* pressure characteristics for three demanded piston amplitudes (1.44, 2.09 and 2.65 mm) measured for the brassboard PMU, at a body temperature of 298K, containing H_2O. The frequency at each pressure increases with amplitude.

water vapour, since the result could be ambiguous if a characteristic similar to that for the brassboard modulator were obtained. However, as seen in figure 3.5, this curve was not observed with the slightly warmer (300 K) flight model PMU bodies, for which the frequency remains a monotonic function of the pressure with a non-zero gradient, apart from minor fluctuations due to sieve thermostat action, which occur around the final filling pressure. It is, however, evident that at higher pressures (above 18 mb) the variation of frequency with pressure decreases in magnitude.

The frequency of the flight model CO_2 PMU is shown in figure 3.8. A similar set of curves, shown in figure 3.9, was obtained for the brassboard modulator at 295K. As expected there is a small dependence of frequency on amplitude and a large dependence on mean pressure. The frequency increases monotonically with both amplitude and pressure, although the gradient decreases at high pressures.

Slight variations in the operating frequencies of both modulators occur; they were observed during filling at Oxford, and are synchronous with the sieve thermostat cycling, which alters the pressure by a small amount periodically. They produce an uncertainty in the pressure when the PMUs are in the instrument. As the sieve is thermostatted to a constant temperature, it merely acts as a finite (but large) gas reservoir, so that if a slow leak occurs during flight and the pressure falls, the change in frequency can be used to calculate the new pressure for recalculation of the weighting functions. A gradual rise in frequency after crimping indicates a leak, as mentioned above. After more than two years in storage at JPL, neither flight model pressure modulator / sieve system appears to have leaked.

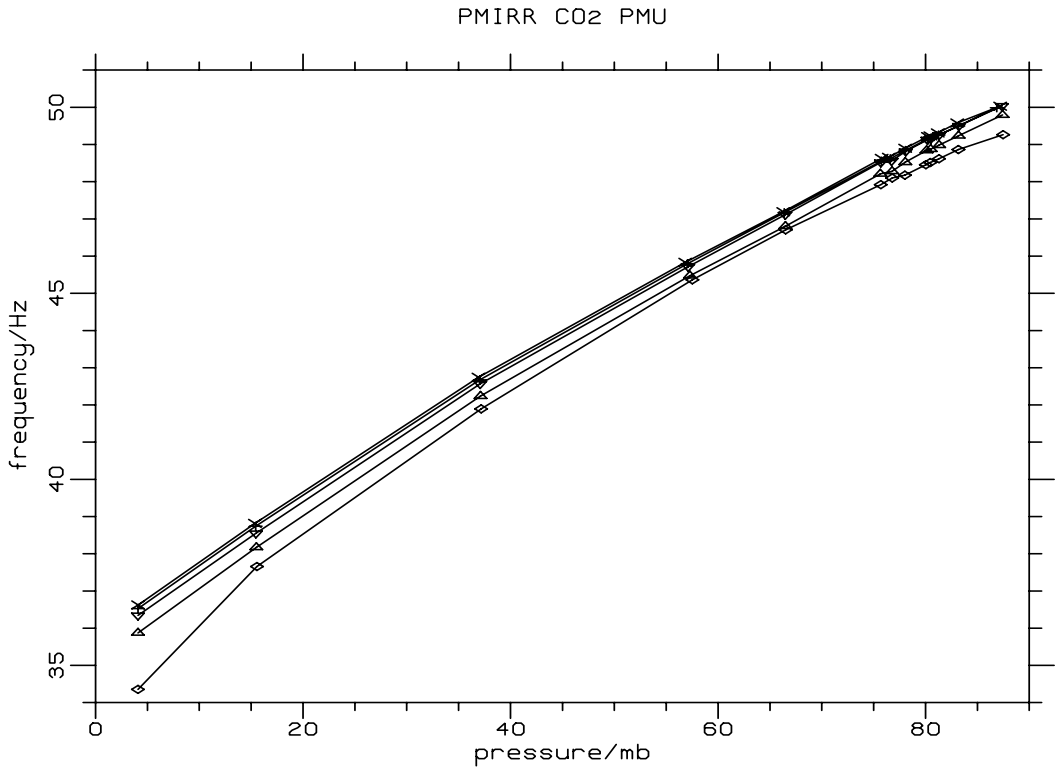

Figure 3.8: Frequency *versus* pressure characteristics measured for the PMIRR flight model CO_2 PMU during filling, at five piston amplitude settings: 1.07, 2.15, 2.77, 3.00 and 3.09 mm. The frequency increases with amplitude for all the pressures.

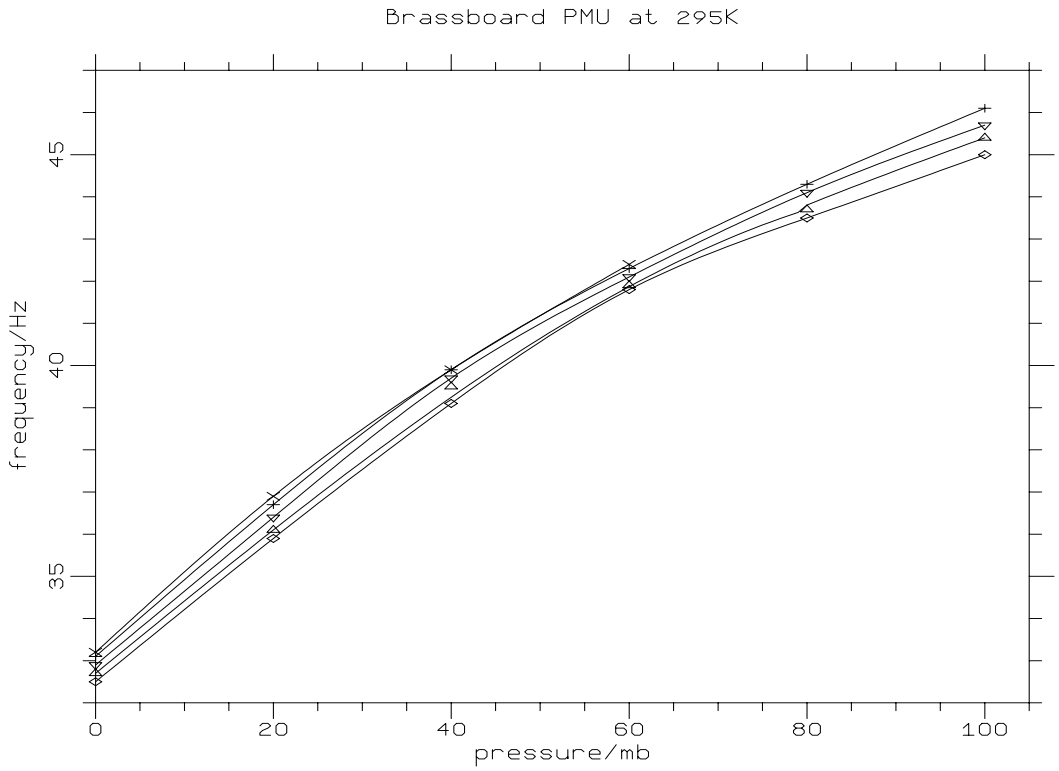

Figure 3.9: Frequency *versus* pressure characteristics measured at five piston amplitude settings (1, 1.5, 2, 2.5 and 2.7 mm) for the brassboard PMU containing CO_2. The frequency at each pressure increases with amplitude.

This page intentionally left blank.

Chapter 4:
PMIRR Pressure Modulator Drive Circuit

In order to take full advantage of the spectroscopic benefits of pressure modulators, the mechanism must be run at a constant amplitude and frequency. A complex drive circuit is required to achieve this. It should control the amplitude of oscillation within a few percent of the desired value, correct the tendency of the pistons to creep together, be insensitive to variations in temperature and supply voltages over the expected ranges, and minimize the power consumption of the modulator.

Planetary spacecraft generally have only a small amount of power available for the instruments in the payload, because of the small amount of mass available to transport either radioisotope thermoelectric generators (RTGs) or solar panels. In the case of *Mars Observer*, the total payload power is about 150 W. PMIRR's budgeted power consumption was 30W, of which 3.6 W was allocated for both modulator systems. Early measurements of the power consumption of a pressure modulator running under the PMIRR CO_2 conditions (80 mb, 2:1 compression ratio) indicated that with an ISAMS-type modulator and drive, the power consumption could exceed 10W. This high power would have caused a significant impact on the scope of PMIRR, and so I undertook an intensive investigation into the power consumption of PMUs under different conditions, to identify the causes of inefficiency. Some of this lay in the motor, as discussed in chapter 3.

I made measurements with a variety of circuits with sine-wave and pulse-width-modulated (PWM) outputs, in the course of which I managed to drive the prototype ('brassboard') PMU with less than 4 watts near the operating conditions of the PMIRR CO_2 modulator, using a sinusoidal drive waveform. Considering the results of the measurements, I designed a new drive circuit for the PMIRR pressure modulators, to run with high efficiency. This chapter describes that circuit in some detail.

4.1 The ISAMS drive circuit

The ISAMS drive circuit (figure 4.1) provides positive- and negative-going pulses, one per half-cycle, at the PMU frequency. The width of the pulses in each direction is altered to control the amplitude of the differential-mode piston motion, determined from the sum of the PPO outputs, and the mean piston position. The phase of the drive coil voltage waveform is very close to 90° different to that of the PPO output, because of the integration produced by the triangle generator. A phase-locked-loop (PLL) is used to provide a stable frequency for the drive waveform, locked in phase to the PPO output (when it is large enough).

In the light of the measurements, and from analysis of the ISAMS drive circuit, I identified several deficiencies in its design which reduce its efficiency. The square output pulses lead to inductive commutation spikes and power loss in unwanted harmonics. The inductance of the drive coils causes the current through them to be out of phase with the voltage across

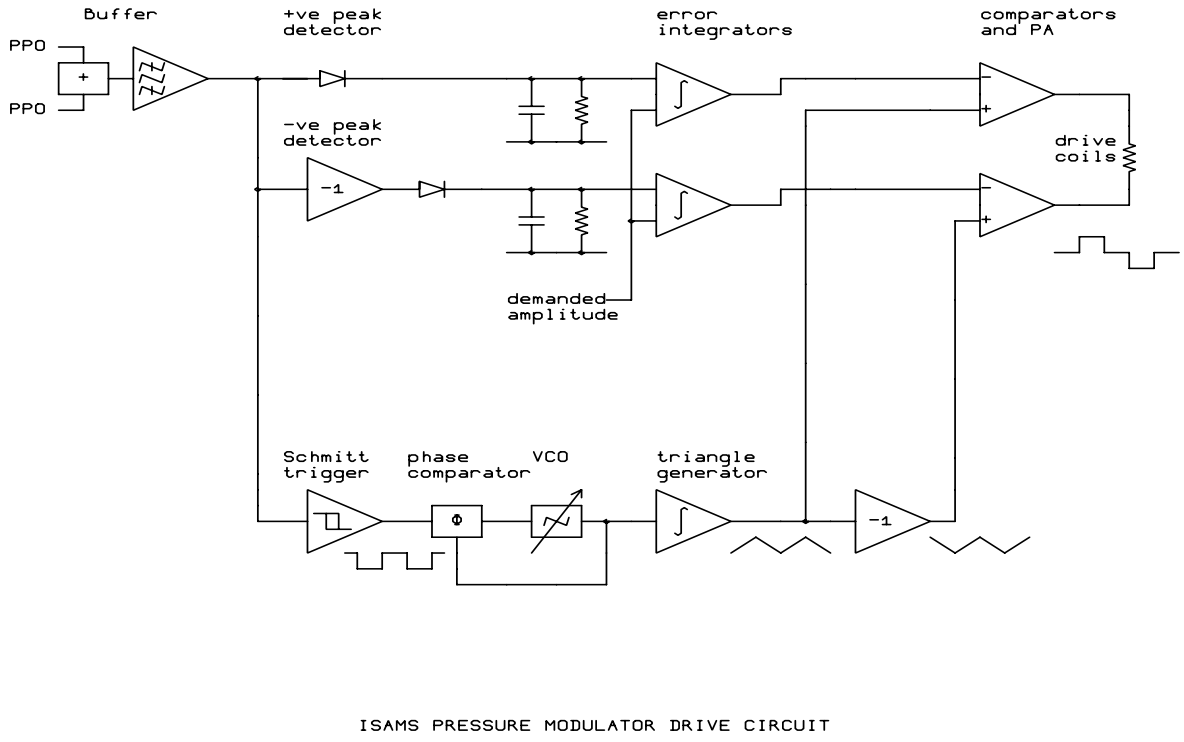

Figure 4.1: ISAMS pressure modulator drive circuit: block diagram (after S.T.Werrett).

them, so that the modulator is run with a phase error between the drive circuit output and the driving force. Also, the phase comparator is susceptible to noise, which leads to jitter in the operating frequency and phase. These phase errors lead to running off-resonance, thus causing a significant inefficiency in the drive/modulator system.

4.2 Design of improved circuit

4.2.1 General approach

An analytical model of the PMU, described in Appendix B, was used to assist in the design of the present high-efficiency circuit in which the coupling between the circuit and modulator is optimized. The mechanical behaviour of the two-piston modulator may be separated into two normal modes, of which only the differential mode is useful for compressing the gas. It is valid to consider only that mode, using the differential-mode PPO output and differential-mode excitation configuration of the drive coils. The system may then be treated (in the absence of imbalance) as a single-piston modulator, and the term 'piston position' is used to represent 'differential-mode displacement' in the case of the balanced modulators used in PMIRR.

To ensure that only the PMU mechanical/gas system defines the running frequency, the drive circuit sets the phase shift between the piston position and the drive current. The system frequency is then a function of the gas pressure inside the PMU, and can be used for diagnostic purposes, if correctly calibrated, as discussed in chapter 3. The analytical model of the modulator (Appendix B) shows that the drive frequency which gives the maximum motional amplitude for minimum power is close, but not identical, to the reso-

nant frequency (the natural frequency of oscillation, determined from the complementary function of the motion solution) and the frequency for which the drive current leads the motion by 90 degrees. The force on the pistons is proportional to the current through the coils, with a small further dependence on the piston position which affects the coupling between the coils and the magnets. I have assumed that this effect may be neglected, because it is very difficult to treat and because the predictions of a more complicated pressure modulator model (described in chapter 5) suggest that the mechanical behaviour of the modulator is insensitive to the nature of the driving waveform.

To ensure that the drive current is used as a true phase reference, and not the drive voltage (which is about 45° out of phase with it), and that phase errors do not accumulate in the drive circuit, the phase-locked-loop (PLL) encompasses the whole system, with inputs from both a drive coil current sensor and the PPO output. Furthermore by redesigning it in the light of the analytical model, very tight phase control is achieved in such a way that the frequency is determined solely by the PMU's frequency/phase relationship.

The power amplifier is a pulse-width-modulated design, described in detail below, which uses a switching ('carrier') frequency in the region of 16 kHz, from which the sinusoidal drive voltage is generated synchronously, so that the number of samples per cycle remains constant. It has a measured efficiency greater than 90%. The static power dissipation of the electronics was reduced to less than half a Watt by using low-power low-bandwidth ICs where possible. The PPOs themselves dissipate 150 mW each; the PPO design was fixed, so this power could not be reduced.

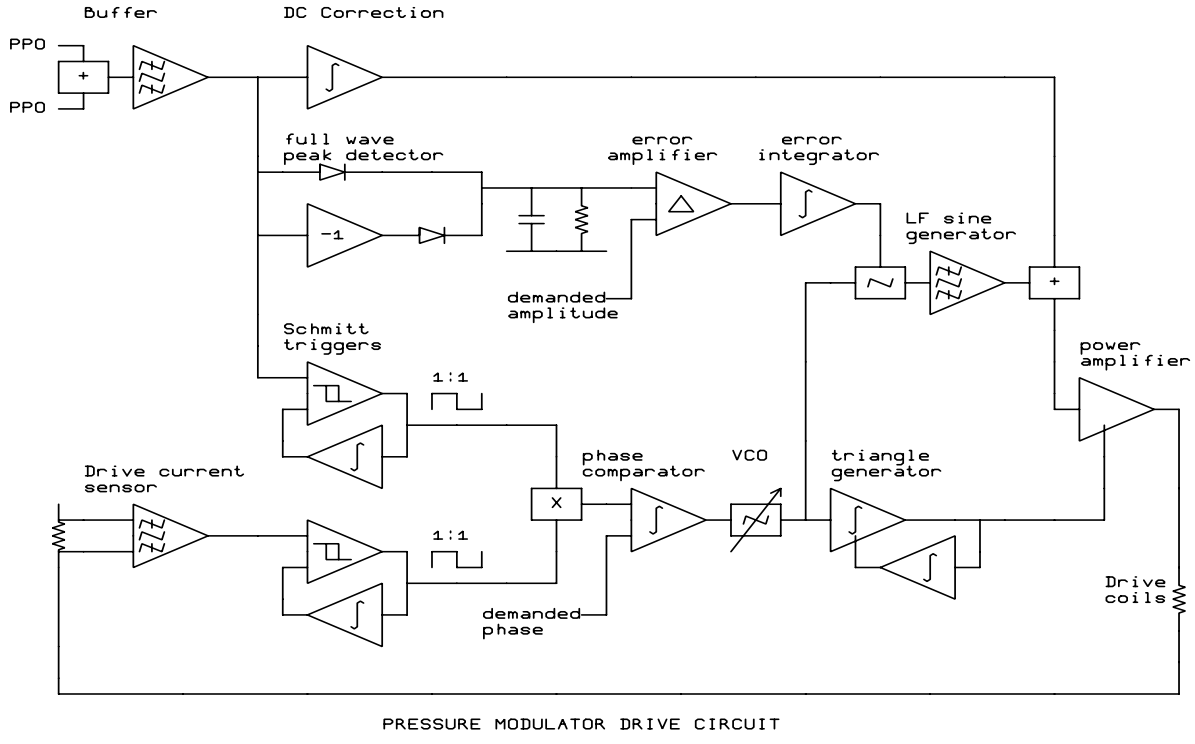

Figure 4.2: Pressure modulator drive circuit: block diagram. Comparison with the ISAMS circuit (figure 4.1) shows the different use of the phase-locked-loop, which here encompasses the whole PMU–drive system.

The block diagram is shown in figure 4.2, and the circuit diagrams in figures 4.3, 4.4, 4.5

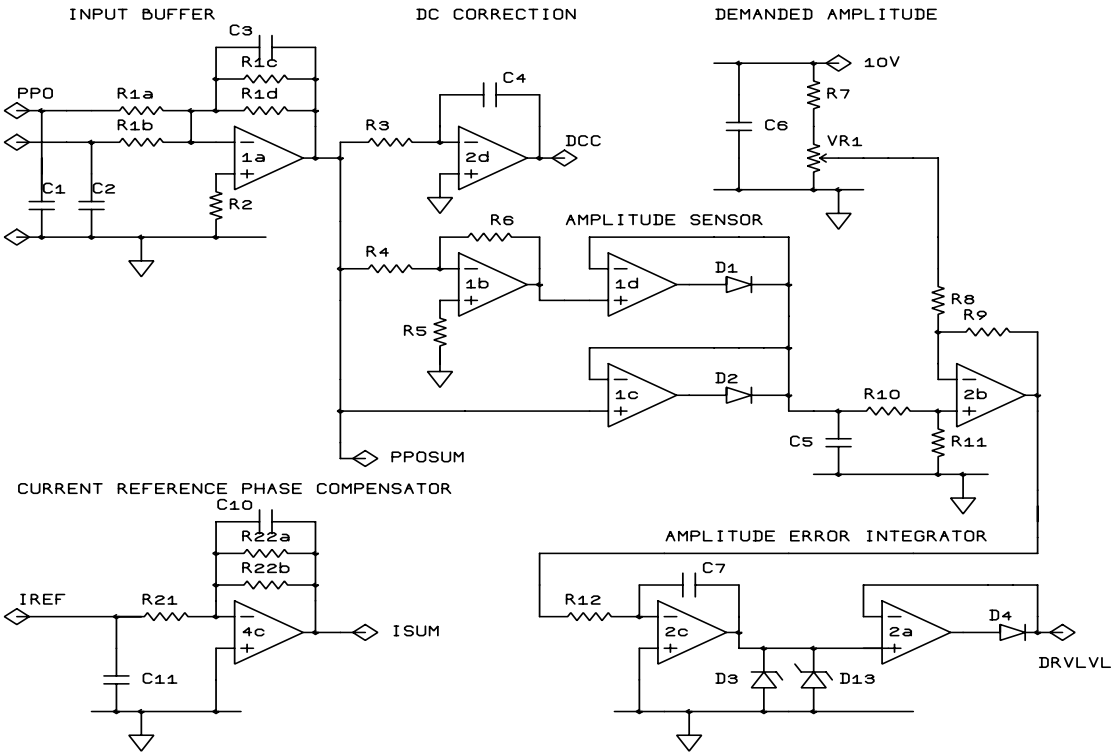

Figure 4.3: Signal Conditioning: the sum of the PPO outputs is used to determine the amplitude, DC offset and phase of the piston motion, which are all controlled by the circuit. VR1 sets the demanded amplitude.

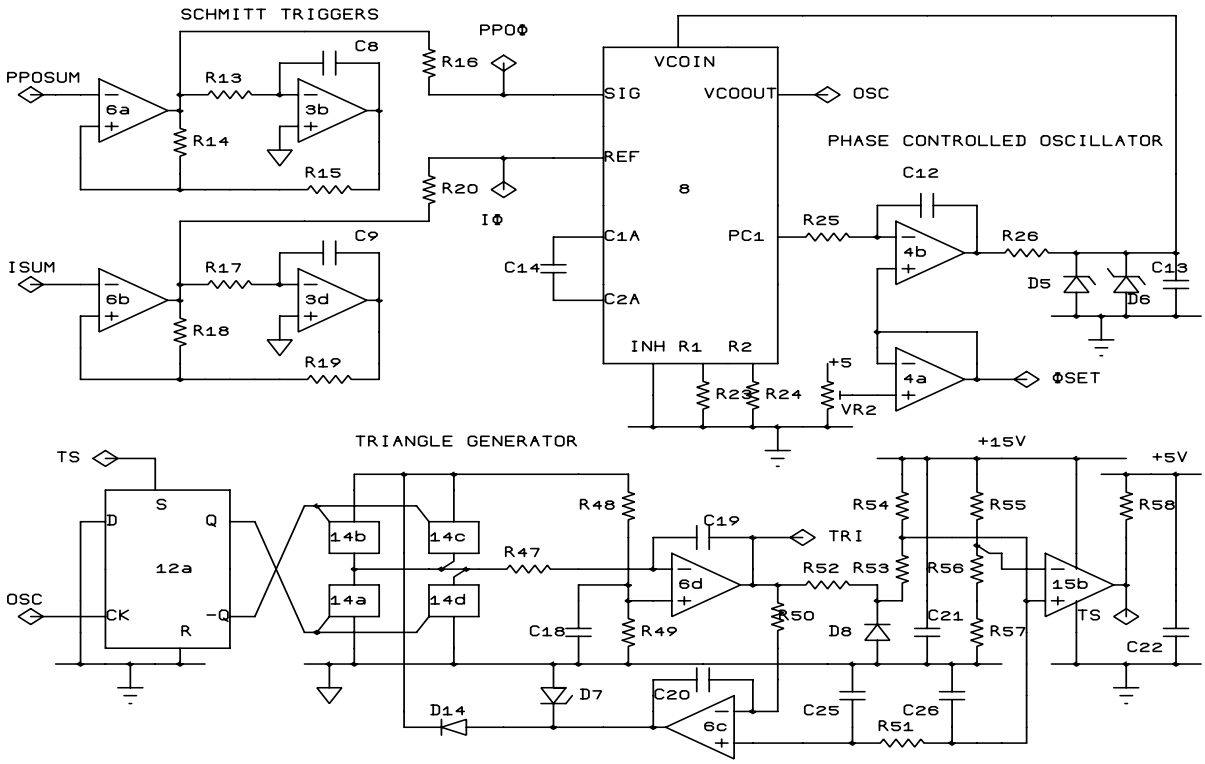

Figure 4.4: PLL and triangle generator. The demanded phase difference between the piston position (PPO output) and drive force (coil current) is set by VR2. The triangle generator is self-adjusting in amplitude, and configured manually so that the minimum voltage in the output waveform is 0V.

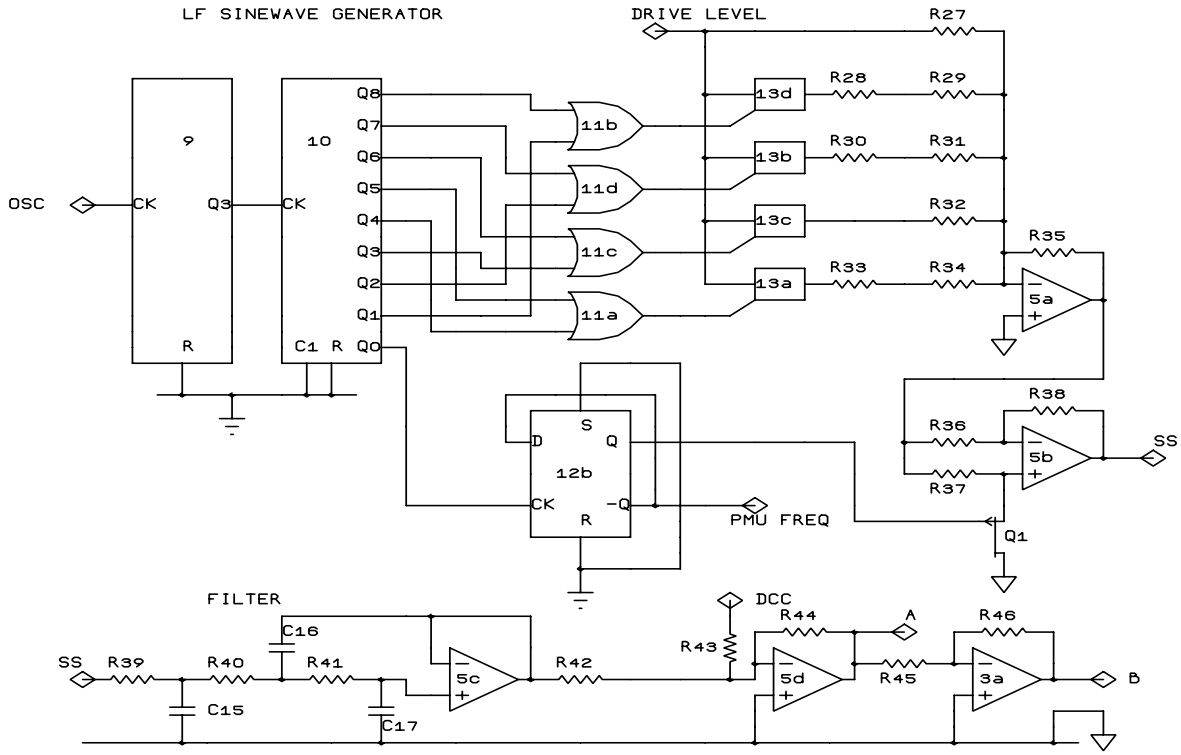

Figure 4.5: Sinewave generator, DC correction combiner and driving inverter. Resistors R27–R35 and IC5a form a switched-gain amplifier which synthesizes a rectified sinewave, inverted on alternate half cycles by IC5b.

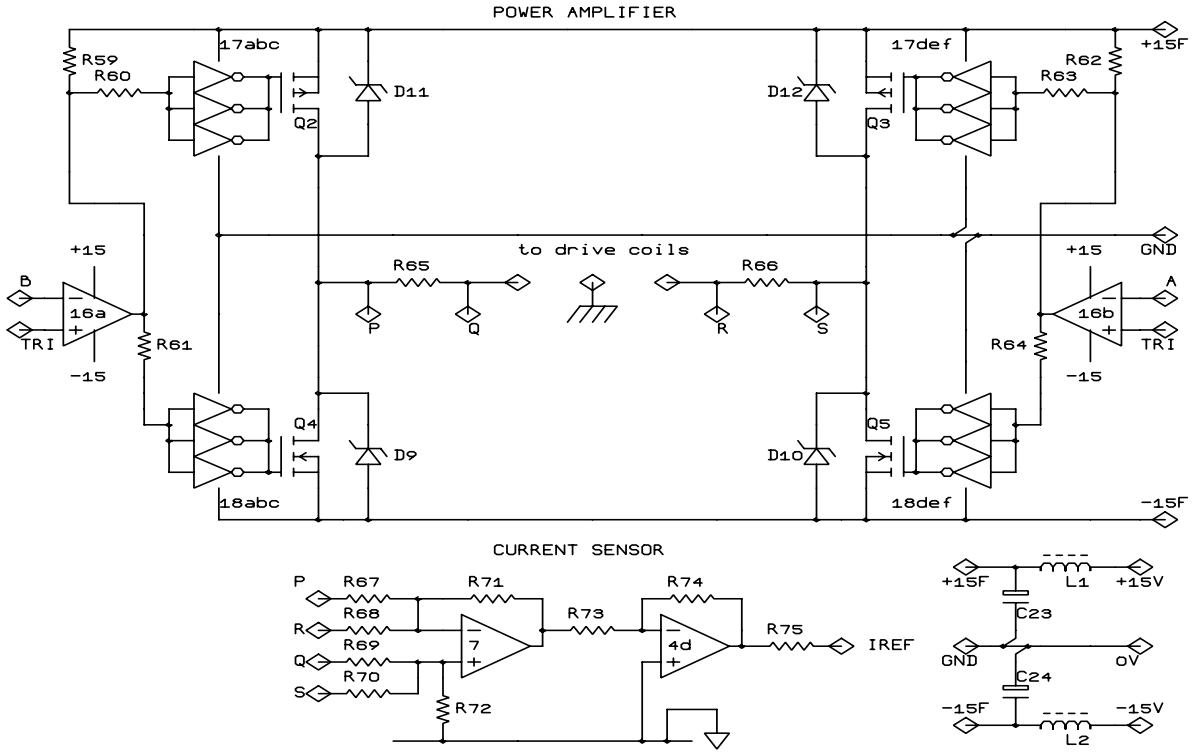

Figure 4.6: PA and current sensor. IC16 compares the two antiphased voltage demands with the high-frequency triangle waveform to control the switching of the four HEXFETs Q2-Q5. The power supply to this stage is filtered by L1–2 and C23–24.

R1a-d	47k §	R48	110k	C1,2	10n §
R2	12k	R49-51	100k	C3	47n §
R3	1M	R52,53	1k	C4,5	1µ
R4	100k §	R54,55	39k	C6	1n
R5	47k §	R56	27Ω *	C7-9	1µ
R6	100k §	R57	2k	C10	47n §
R7	160k	R58	24k	C11	10n §
R8-12	1M §	R59	10k	C12	1µ
R13	10M	R60,61	30k	C13	150n
R14	1M	R62	10k	C14	2.2n
R15	470Ω	R63,64	30k	C15	150n
R16	100k	R65,66	0.1Ω §	C16	33n
R17	10M	R67-72	100k §	C17	10n
R18	1M	R73	10k	C18	100n
R19	470Ω	R74	1M	C19	150p
R20	100k	R75	3k §	C20	1µ
R21	47k		C21	100n	
R22a-b	47k §		C22	47n	
R23	47k *		C23,24	10000µ	
R24	82k *		C25,26	1µ	
R25	3.3M				
R26	1k				
R27	470k	IC 1-5	OP420	D1,2	1N4148
R28	200k	IC 6	OP421	D3	5V1
R29	47k	IC 7	OP15	D4	1N4148
R30	130k	IC 8	CD4046B	D5	BAT 85
R31	3.3k	IC 9	CD4040B	D7	5V1
R32	100k	IC 10	CD4017B	D8	1N4148
R33	82k	IC 11	CD4071B	D9-12	40V 3A Schottky
R34	6.2k	IC 12	CD4013B	D13	BAT 85
R35	27k	IC 13,14	CD4066B	D14	1N4148
R36-38	100k	IC 15,16	LM393N		
R39	11k	IC 17,18	CD40106B	VR1	10k ten turn wirewound §
R40,41	100k		VR2	10k ten turn preset §	
R42,43	47k	Q1	2N3820		
R44	91k *	Q2,3	IRF9530	L1,2	50mH 1A dc
R45,46	200k	Q4,5	IRF530		
R47	56k				

§= High precision. * = Adjust on test.

Table 4.I: Parts list for the drive circuit

and 4.6. The component values are presented as table 4.I.

4.2.2 Sectional description

4.2.2.1 Input buffer

The input buffer (figure 4.3), copied from the ISAMS circuit, is a summing amplifier with a high-frequency roll-off to reduce the level of the PPO carrier frequency breakthrough and other high-frequency noise. A matched resistor gain network is used to reduce the temperature coefficient. The sum of the two PPO output voltages is proportional to the

piston separation.

Because of the long time constant formed by R1c, R1d and C3, the gain of the input buffer is weakly dependent on frequency. Thus the piston amplitude per monitored volt has a (calculated) gradient of 0.18% per Hz.

4.2.2.2 Current sensor

The current through the drive coils is sensed with a 0.1 Ω resistor in series with each side of the drive coils (figure 4.6). There is a very large common-mode signal at the input to the differential amplifier, with a 30V pk-pk common-mode swing at the power amplifier carrier frequency (see below) across the sensing resistors, compared with a minimum current of about 10mA for an empty modulator, which develops 1 mV across one of the resistors. Thus the maximum common-mode to differential-mode input signal ratio is about 100dB. Common-mode breakthrough in the differential amplifier output causes a noisy non-sinusoidal output waveform and a resulting phase error. The differential amplifier uses an OP15 operational amplifier IC, because of its high common-mode rejection-ratio (CMRR) and very low input bias and offset currents. It takes up to 5 mA of supply current, and therefore can dissipate up to 150 mW (almost half the static power consumption of the whole circuit). However, the power saving when running the PMU at pressures above about 20 mb significantly outweighs the loss in the IC.

The drive current reference is amplified to make the amplitude approximately the same as that of the PPO sum signal at the normal operating point.

4.2.2.3 Reference phase generator

In order to achieve precisely the desired phase shift, the output of the current sensor is amplified by a circuit identical to that used to buffer the PPOs (figure 4.3), so that the relative phase of the PPO and drive signals remains unchanged. The filter also attenuates any residual carrier frequency noise from the switching power amplifier.

Both the PPO and current sensor signals are taken to Schmitt triggers (figure 4.4) whose reference (mean switching) level is adjusted automatically (by integration) to provide unity mark/space ratio at the output. The 1:1 mark/space ratio is necessary to optimize the phase comparator. This technique provides zero DC gain, so that the phase of the square-wave output is the same as that of the alternating part of the input signal, without over-accentuating any high-frequency components such as the carrier frequencies. The hysteresis level in the Schmitt trigger should be small enough to prevent a significant phase-lag between the input and output, but large enough to prevent output bounce at edges. In the absence of a signal the circuit behaves as a low-frequency oscillator, because of the long time constant required to prevent noticeable change in the comparator levels over a PMU cycle.

4.2.2.4 Phase locked loop (PLL)

Phase locked loops consist of a phase comparator, filter and voltage-controlled oscillator to provide a phase or frequency servo of varying precision. Typical applications are the demodulation of weak radio signals, and the synthesis of high frequencies. The PLL in the PMU drive circuit is importantly unusual in its function. Most PLL circuits slave the

frequency of the output to the *frequency* of the input. In the PMU drive application, the frequency of the voltage-controlled-oscillator (VCO) is adjusted to cause the difference in phase between the modulator driving force and piston separation (displacement) to assume a preset value. The circuit itself has no information as to what that frequency may be; it is solely determined by the modulator mechanism. Thus the PLL must slave the *phase* around the loop, with the result that the frequencies of the two inputs of its phase comparator are always the same, since the PPO output is the result of the motion of the piston at the frequency of the forced oscillation caused by the drive current. For that reason the PLLs used in PMU drive circuits have no natural frequency.

The first version of the drive circuit used the edge-sensitive phase comparator design from the ISAMS circuit, in which the durations of the two (positive and negative) output pulses are adjusted until the phase of the two inputs is equal. Because the two phases are equal, the phase lag between the PPO output and drive current must be added externally. In the ISAMS circuit this was achieved elegantly by using an integrating triangle-generator.

The circuit is similar to Phase Comparator 2 in the 4046, but resolves an ambiguity in the behaviour of that comparator when used for PMUs. This edge-sensitive type of comparator gives a precise control of the output phase, but behaves very badly in the presence of noise in either input; the phase lock is lost, and the VCO runs freely, so that the PMU frequency is uncontrolled. The induced noise was reduced by screening the PPO output and drive coil connection wires in the lead to the PMU separately, and by earthing the PMU body independently.

I devised an alternative phase comparator, shown in figure 4.4, to solve these problems. The PLL uses the voltage-controlled oscillator (VCO) in the 4046, with resistor and capacitor values chosen to set an appropriate frequency range which bracketed the expected PMU operating frequencies. (The VCO operates at 320 times the drive frequency). The phase comparator is based on the multiplying phase comparator 1 in the 4046, which is an exclusive-or (XOR) gate. The mean output of the gate, when the mark-to-space ratio of both inputs is unity, is proportional to the phase difference between them; if the two inputs are coincident, the output is always zero, whereas if they are in antiphase the output is always logic '1'; the mean output voltage varies continuously between these extremes. In its normal use as a phase comparator, the output of the XOR gate would be filtered through an RC filter to provide the VCO input voltage; in this application is integrated against a settable voltage near 2.5V, so that the VCO input voltage stabilizes when the mark/space ratio of the XOR gate output is near unity. This occurs when there is a near-90° phase difference between its inputs when they are at the same frequency. The phase difference between the two inputs at which the output settles is determined by the precise level of the reference voltage to the integrator, and may be adjusted to minimize the power consumption of the PMU drive coils under operational conditions. The integrator has a long time constant to reduce the output ripple, and thus the VCO frequency modulation, to an acceptable level, and to provide a dominant phase lag in the loop, to ensure stability under expected operating conditions.

I tested the PLL with a PMU simulator circuit, which consisted of two RC low-pass filters. The circuit exhibited a 90° phase shift between input and output at a settable frequency in the expected range. The phase comparator locks more quickly with the modulator than with the simulator, because the resonance is more sharply defined, as the phase lag/frequency slope is greater.

4.2.2.5 Amplitude control

The positive and negative peak values of the PPO sum output are detected with a precision diode/capacitor circuit (figure 4.3). A leak resistor across the capacitor gives a 1 second decay time constant. This removes most of the ripple, whilst allowing rapid control of the amplitude. The difference between the demanded and actual amplitude levels is integrated to give the drive level, which controls the amplitude of the sinusoidal part of the drive waveform. If the PPO output exceeds the demanded amplitude, the output of the integrator can become increasingly negative, causing a 180° shift in drive phase, positive feedback in the amplitude control loop, and thus unstable operation. A diode and operational amplifier circuit follows the integrator to prevent the drive level falling below 0V in this way.

The frequency and amplitude control loops are coupled by means of the transfer function of the pressure modulator. Near resonance a small change in frequency entails a large change in amplitude. To overcome this problem the amplitude control loop has a smaller time constant (1s) than the PLL (3.3s). In the normal pressure régime this prevents the pistons hitting the end stops or each other as the frequency settles. However, for pressures below about 2 mb, the Q (reciprocal of the fractional energy lost per cycle) of the PMU is large, and the response very steep, so that the amplitude control loop does not act quickly enough. As the filling pressures for the PMIRR PMUs are much greater than 2 mb, this is not a problem in practice.

4.2.2.6 Synthesised sine-wave generator

The square wave output of the VCO is used to generate a sinusoidal voltage waveform at a subharmonic frequency. This eventually causes a sinusoidal current in the drive coils, preventing inefficiency arising from the excitation of unwanted harmonics in the drive coils. The resulting sine wave is synchronous with the power amplifier (PA) carrier frequency, which simplifies the design of the PA.

Originally I used the VCO output to drive an address-generator for an EPROM, which with an eight-bit multiplying DAC produced a sine wave at the drive frequency. Because of problems procuring parts to implement this method, the present sine-wave generator (figure 4.5) uses a circuit copied from the ISAMS cooler drive circuit. A Johnson counter is used to switch four analogue switches, thereby varying the gain of the op-amp inverting amplifier, which produces an inverted full-wave stepped sine wave. A flip-flop inverts alternate halves of this by switching the gain of the next stage, using a p-channel JFET. The result is a stepped sine-wave which with the correct weights contains no harmonic (apart from the fundamental) below the twentieth. This, and higher frequencies, are removed by the third order low-pass filter which follows, resulting in a smooth sinewave, shifted in phase from the original stepped version, which forms the AC component of the drive. The phase shift is cancelled by the phase control loop.

4.2.2.7 DC servo

As a result of the form of the flow rate in the gap past the modulator piston (the expression is given in chapter 5), the pistons have a tendency to pump gas away from the cell, so that there is a net force pushing them together. To counteract this piston creep, the PPO sum output is integrated (figure 4.3) with a time constant much longer than the drive period. The result is added to the drive voltage (figure 4.5), and amplified by the final stage, to

produce a DC component in the driving force, which tends to restore the mean piston separation to the static value, thereby maximizing the piston travel.

4.2.2.8 Triangle generator

The high-frequency VCO output is also converted into a triangle wave, the lowest points in which are at zero volts. The circuit, in figure 4.4, was adapted from the ISAMS cooler drive triangle-generator. Analogue switches are used to isolate the digital and analogue sections, thereby removing noise spikes from the output. An integrating servo keeps the amplitude of the triangle constant with frequency, providing a flat overall frequency response in the amplifying section of the drive circuit. The modifications were suggested by J.R.Locke [private communication] at JPL.

4.2.3 Power amplifier

The power amplifier uses a pulse-width modulation technique (PWM). The analogue input is transformed into a series of rectangular pulses, of constant peak voltage, at a high carrier frequency. The width of each pulse, and thus the average output voltage over the period of the carrier, is proportional to the instantaneous input voltage. In practice the conversion from voltage to pulse width is achieved by comparing the input with a triangle wave at the carrier frequency. When the input voltage is greater than the triangle wave, the load is connected to the power supply through high-speed, low-resistance transistors; otherwise it is short-circuited.

This type of amplifier is equivalently referred to as 'switched-mode', because the power to the load is switched on or off, rather than being applied continuously, or as 'Class D', (by extension from classes A, B and C) because no bias is applied to the switching devices, the gain and efficiency are large, and the conduction times per cycle may be extremely small.

4.2.3.1 Theory of PWM amplifiers

The Fourier transform of an analogue signal $V_0 + V_1 \exp i\omega_1 t$ may be expressed as

$$\int_{-\infty}^{+\infty} (V_0 + V_1 \exp i\omega_1 t) \exp i\omega t dt \tag{4.1}$$

The Fourier transform of a pulse-width-modulated version of the same signal, in which the pulse widths are proportional to the analogue signal level, for frequencies less than half the carrier frequency, and neglecting switching transient effects, is

$$\int_{-\infty}^{+\infty} V_S \exp i\omega t (V_0 + V_1 \exp i\omega_1 t)/V_T dt \tag{4.2}$$

in which V_S is the supply voltage to the PA stage, and V_T is the amplitude of the triangle wave against which the signal is compared.

It is evident by rearrangement that equation 4.2 is merely equation 4.1 multiplied by V_S/V_T. Thus PWM amplifiers introduce no distortion at frequencies below half the carrier frequency, assuming that the resistance of the switches is small compared with that of the load, and the switching time is small compared with the carrier period. Both of these criteria are met by the use of HEXFETs as switches.

4.2.3.2 Practical circuit

The physical arrangement (figure 4.6) is a bridge circuit, with 2 p-channel and 2 n-channel FETs whose gates are driven by 3 paralleled 40106 Schmitt trigger inverter gates. This allows gate-voltage rise and fall times of about 1µs. The p-channel FETs conduct when the gate is lower than +15V, whilst the n-channel FETs require the gate to be higher than -15V. The FETs are protected against inductive reverse voltages by Schottky diodes.

The drive voltage and its inverse are compared with the triangle signal. This determines the conduction period for the power FETs on the two sides of the PA bridge. The n-channel FET on either side conducts only when its analogue signal is positive and greater than the output of the triangle-generator; the p-channel FET on the same side is off when the n-channel FET is on, and there is a short interval when neither conducts, to prevent a short-circuit path between the supply rails. In the PMIRR flight model drive circuit, the power supplies are at 0 and 28 V so that the connecting circuit between the comparators and the FETs is more complex.

4.2.3.3 Inductive load considerations

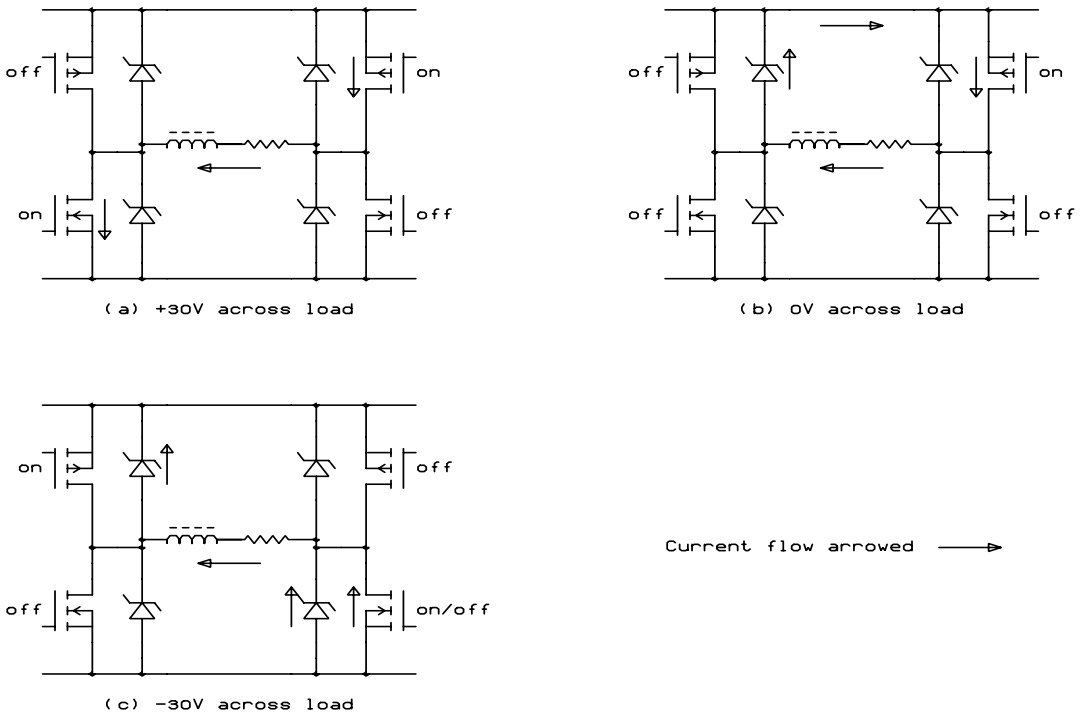

Figure 4.7: Current paths through a PWM power amplifier of the 'duty-cycle/steer' type (see text). In (a), current flows from the positive rail through the load to the negative rail, as expected. In (b), the voltage across the load is zero, but current continues to flow through the short circuit, because of the inductance of the load. In (c) the steering direction has changed; the current continues to flow in the same direction, contrary to expectation, until it has been dissipated by the resistance of the new circuit. This mechanism causes commutation spikes leading to distortion, and is avoided by always having one FET conducting on each side.

In an earlier version of the PA, the p-channel FET was kept conducting for a whole half-cycle; this is common practice in circuits of this type, as it reduces the amount of drive power required for the FETs, and for a resistive or nearly-resistive load makes no difference to the behaviour of the amplifier. The full-wave rectified drive signal was compared with the triangle to give a 'duty-cycle' signal, steered to the appropriate side of the bridge by a 'sign' signal. This works well for resistive loads, but for inductive loads the fundamental

assumption about the current paths involved in the method cease to be true and this leads to commutation spikes.

When the left-hand n-channel FET and right-hand p-channel FET were conducting, towards the end of a half-cycle, the current path was as shown in figure 4.7(a). The voltage across the load was therefore 30V. When the n-channel FET was turned off, the current continued to flow, because of the inductance of the load. It passed through the (intrinsic and external Schottky) diodes across the p-channel FET on the left-hand side, so that the voltage across the load was zero (figure 4.7(b)). Thus the mean voltage across the drive coils (over a period of one triangle cycle) was proportional to the duty cycle, and thus to the drive voltage, was required.

When the sign of the drive voltage changed (figure 4.7(c)), the direction of current flow through the load remains the same, because of the inductance of the coils. The right-hand p-channel FET had been turned off, but current still flowed (because of the back EMF) from the right-hand side, through the diodes across the right-hand n-channel FET, or through the FET itself when it was pulsed on. So the right-hand side of the load remained at -15V regardless of the duty cycle of the n-channel FET, until the stored inductive energy had been dissipated as heat in the load resistance, diodes, and power supply.

The single-stage LC filter used in the original design changed the form of the constant -15V level for this time into a rounded spike, or, with different filtering capacitors, into ringing. When the filter was removed, and the true cause of the spikes became apparent, the circuit was modified so that current did not 'free-wheel' through the Schottky diodes, but was actively pulled through a FET to one supply rail or the other. Thus, when the voltage had changed sign but the current had not, the (left-hand) p-channel FET was turned on for most of the time, so that the mean voltage across the load is near 0V, rather than -15V as occurred previously. The filter was discarded in the final version, as there is no component in the output of the amplifier between the signal frequency and half the carrier frequency. The drive coils form an LR circuit, the current through which is a low-pass-filtered version of the voltage across it; the knee frequency (3dB) is about 25 Hz, similar to the PMU frequency, so that the carrier-frequency component of the drive voltage produces negligible drive current, and dissipates negligible power.

4.3 Drive circuit tests and power measurements

To verify that the new drive circuit was functioning properly, and to provide engineers at JPL with data on the expected frequencies and power consumptions of the two modulators in the PMIRR instrument, I conducted a series of tests on the system performance, using the brassboard PMU described in chapter 3. The frequency and power consumption were measured for different amplitudes and pressures; the pressure range encompassed the flight specification for both water and CO_2, although the PMU had insufficient free travel to allow the flight compression ratios to be used for measurement. The maximum attainable amplitude with the PMU as it then was was 2.7 mm (pk-pk) differential; this could not be attained at high pressures. For the power measurements I used amplitudes of 1, 1.5, 2, 2.5, and 2.7 mm pk-pk.

4.3.1 Experimental details

The brassboard modulator was attached to a pumping/filling rig with a bottle of CO_2 and an H_2O ampule. A mechanical gauge, with a resolution of approximately 1 mb, and a

baratron with a 100 mb head, were used for measurement of the pressure. The PMU drive unit was connected to an oscilloscope, a frequency counter with a resolution of 0.1 Hz, a true-rms digital voltmeter (DVM) for measuring the PPO outputs, and two digital current meters for measuring the current in the power supply rails.

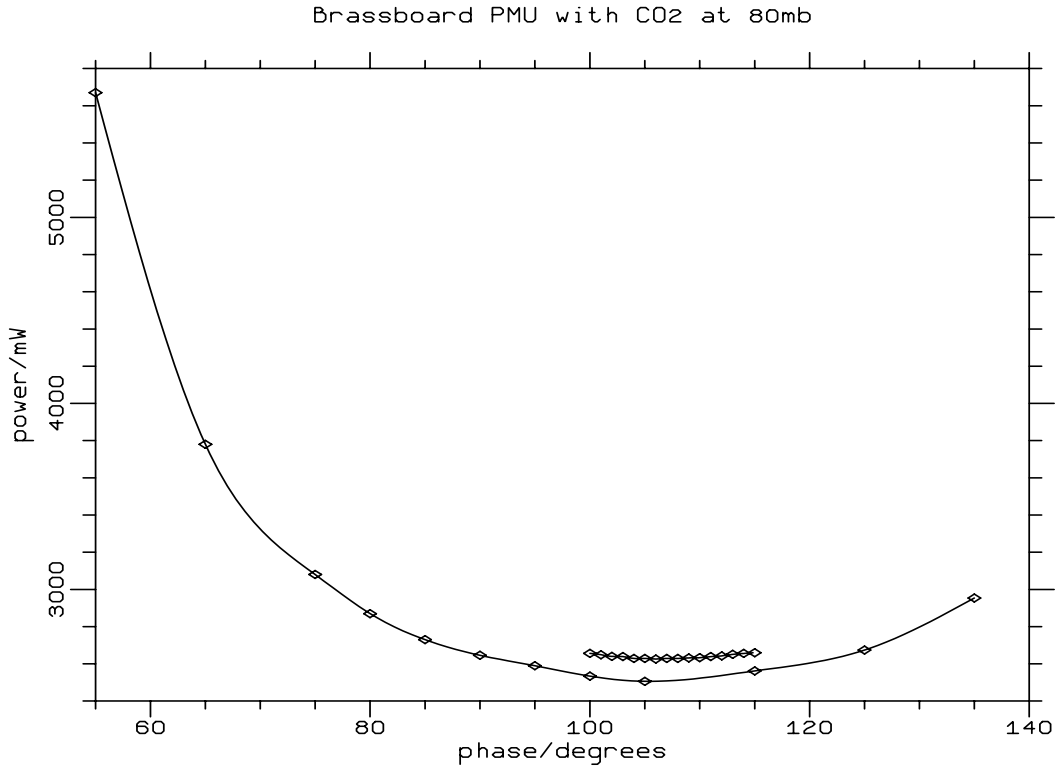

Figure 4.8: Power consumption *versus* phase, measured for CO_2 in the brassboard PMU. The offset between the two curves occurs because the coarse- and fine-mesh measurements were taken at different times.

I first measured the power consumption and frequency against phase setting (figures 4.8 and 4.9) with 86.6 mb of CO_2 in the PMU and used the minimum to set the phase for the further measurements. The nature of the power curve confirms the assumption that the phase must be carefully controlled for efficient operation. From the phase setting for minimum power the magnification factor Q may be estimated, using the expression from appendix B:

$$Q = \frac{\exp(2\pi\beta)}{\exp(2\pi\beta) - 1} \tag{4.3}$$

where β is the absolute value of the cosine of the phase difference between the PPO output and the drive current. The power minimum occurs at a phase of $105(\pm\ 2)°$; the Q is therefore 1.25 ± 0.06. The difference between the 90° and 105° power consumptions is not very great, so that the drive circuit could possibly have operated from a constant 90° phase shift; however, the phase error of 37° or 58° obtained by ignoring the phase shift in the drive coils (described above), so that the abscissa in figure 4.8 is 53° or 32°, obviously increases the power consumption well beyond acceptable limits.

Subsequently I measured the power consumption and frequency as a function of pressure and amplitude up to and beyond the flight pressures, and as high as the PMU would allow in amplitude (2.7mm pk–pk). The mean power was measured by adding the currents

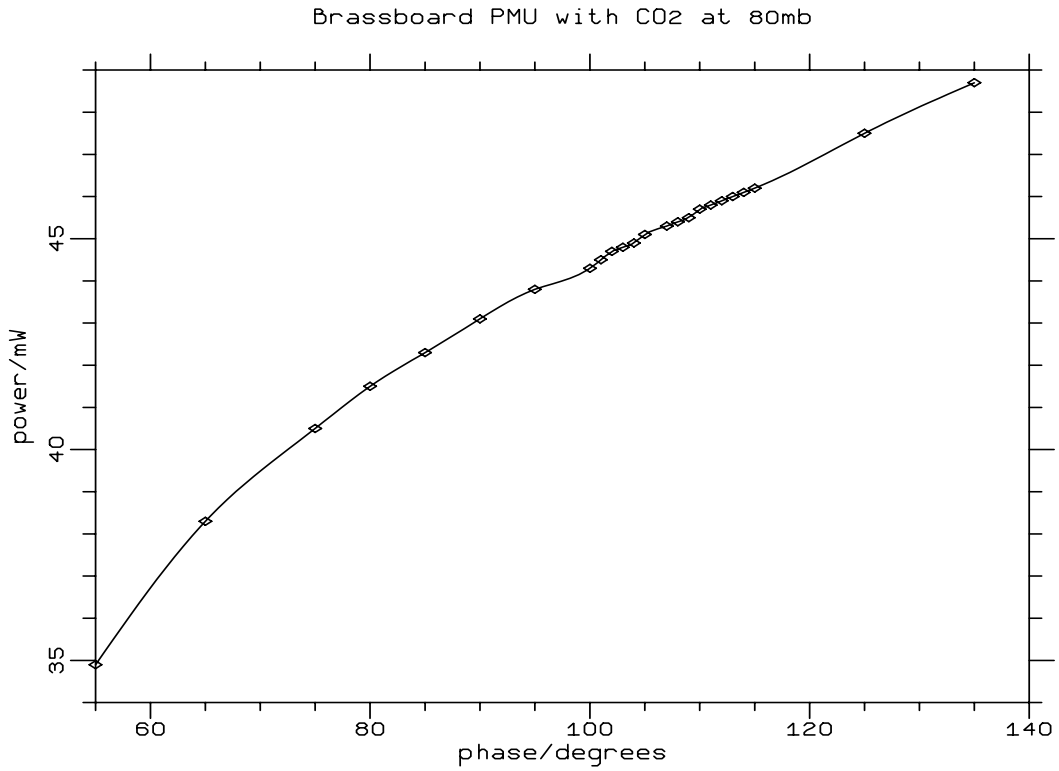

Figure 4.9: Frequency of the brassboard modulator, containing CO_2, *versus* phase.

delivered in the two supply lines and multiplying by the supply voltage of 14 V. The power figures given therefore include the drive circuit housekeeping and PPOs as well as the drive coils. Measurements made with the drive coils disconnected indicated that the static power consumption of the drive circuit and PPOs together was 600 mW. The resistance of the current meters was nulled by setting the voltage delivered to the drive circuit to $\pm$14 V for each amplitude or pressure increment, to provide conditions similar to the flight model drive circuit, which would have to operate from a 28V power rail.

4.3.2 CO_2 measurements

For CO_2 measurements, the a 4cm-long cell was attached to the modulator, and filling pressures of 0, 20, 40, 60, 80 and 100 mb were used. The power consumption is shown in figure 4.10. The graph shows that with 2.5 mm amplitude and 80 mb mean pressure, the modulator, PPOs and drive consumed 3.16 Watts. As expected, the power consumption increases monotonically for increases in both mean pressure and demanded amplitude. The frequency characteristics of the brassboard modulator when containing CO_2 have been described in chapter 3.

4.3.3 H_2O measurements

For H_2O measurements, a 10cm-long cell was attached to the modulator, and mean pressures of 0, 5, 10, 15 and 20 mb were used. The results are shown graphically in figure 4.11. The power delivered to the whole system was roughly constant at 600–650 mW, although for large piston amplitudes it does show an increase with pressure. Generally at these pressures the power required for the modulator itself is negligible in

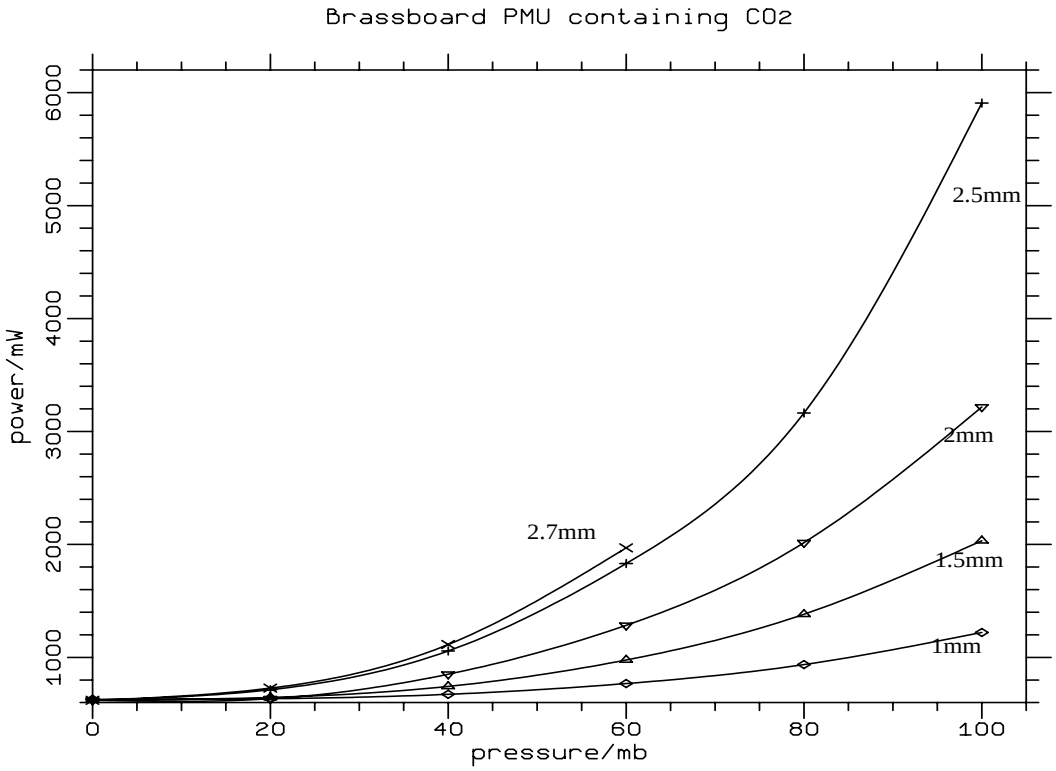

Figure 4.10: Power consumption of the brassboard modulator containing CO_2, *versus* pressure and amplitude. The flight pressure and amplitude are 80 mb and 2.9 mm respectively. The brassboard PMU could not be run at 2.7 mm amplitude at pressures above 60 mb.

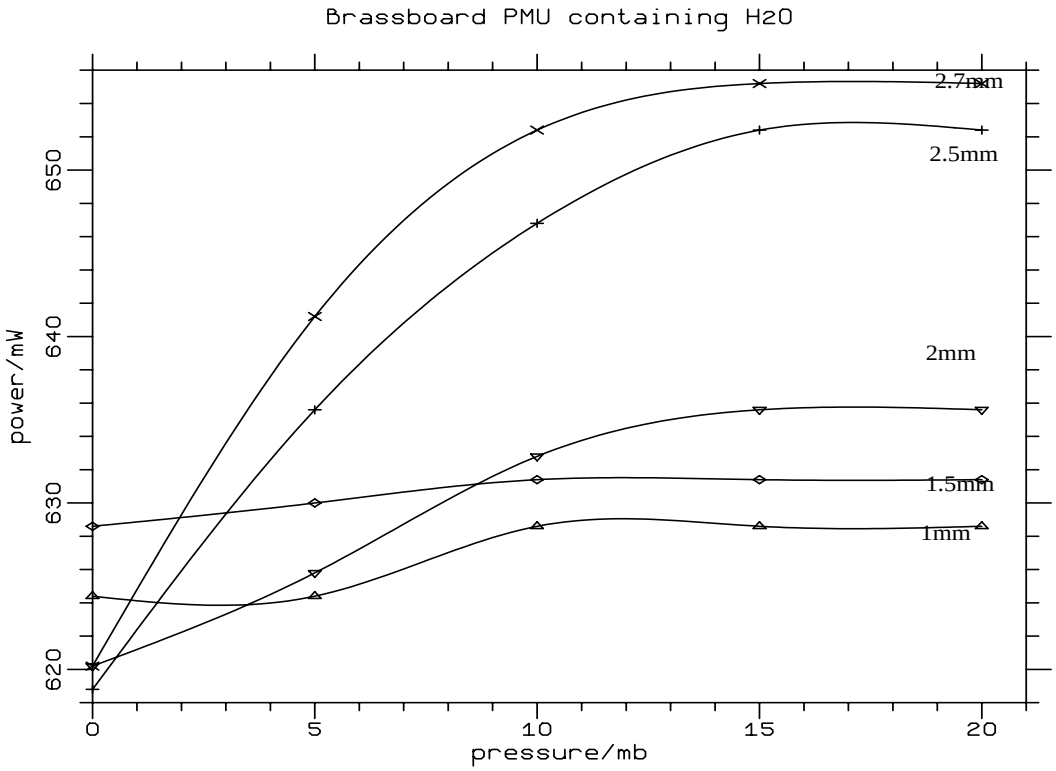

Figure 4.11: Power consumption of the brassboard modulator containing H_2O *versus* pressure, for five amplitude settings. At low amplitudes and pressures the static dissipation of the drive circuit and PPOs dominates.

comparison with the static dissipation of the drive circuit and PPOs. This explains the form of the curves for low pressures. The frequency characteristics have been described in chapter 3.

4.3.4 Conclusion

The results of the tests showed that the PMU's power consumption could, with the new high-efficiency drive circuit, meet the PMIRR allocation. The drive circuit satisfied all its requirements, running the PMUs stably, except at low pressures, when the current feedback is insufficient.

The circuit was modified by J.R.Locke at JPL for use on the PMIRR flight model. The power amplifier operated from +28V and 0V; this made the PA driver more complex but reduced the common-mode input into the current sensor, since the resistors could be in the source leads of the n-channel FETs. The sinewave drive voltage was produced by filtering a low-frequency squarewave, rather than by using an intermediate stepped-sinewave. Components were chosen from the JPL approved parts lists.

Future drive circuits should perhaps include more diagnostics; for example, when the modulator pistons hit each other, the differential mode amplitude decreases as a result of the phase difference between the two sides of the PMU. The drive circuit attempts to correct for this by increasing the drive level, thus exacerbating the problem. Correction of this requires monitoring of the common-mode motion of the PMU pistons.

Chapter 5:
Steady-state Pressure Modulator Model

This chapter describes a numerical model of the PMIRR pressure modulators. This model was developed to allow calculation of the pressure-modulated transmission profile, necessary for the interpretation of the spectroscopic measurements described in chapter 6, and to determine the spectral response of the PMIRR PMR channels to atmospheric radiation, as discussed in chapter 2.

The model is controlled by a modulator description file, which contains nominal values for and expected uncertainties in the physical dimensions, spring constants, piston masses *etc.*. The nominal values are used for a first run, which calculates the nominal cell pressure cycle, cell temperature cycles, and frequency. Each parameter is then perturbed in turn by its expected error, and the difference in the resulting cell pressure and temperature cycles and frequency are combined to give the root sum square error in each. Verification of the model was achieved by comparison with experimental measurements of the operating frequency, the cell pressure cycle and the optical transmission spectrum. The first two of these are compared in this chapter; the spectroscopic measurements will be described in chapter 6, and analysis will follow in chapter 7.

The physics used in the model is described in this chapter and in Appendix B. Since previous work [*e.g.*, Davis 1987] has raised many questions about the behaviour of water vapour in pressure modulators, the processes involved in H_2O cycling are discussed in detail.

5.1 Overall approach

The pressure modulator model was largely based upon the ISAMS model developed by Venters [1991], adapted to account for the high pressure of gas in the PMIRR CO_2 PMU and the different drive circuit described in chapter 4. Important changes have been made to the physics, with the inclusion of adsorption, saturation effects, and some consideration of the temperature modes in the cell. Instead of using a constant timestep and integrating the PMU's state from rest until it stabilizes, as in previous work (*e.g.*, Schofield [1980], Venters [1991]), I have considered the PMU behaviour as a two-point boundary-value problem, in which the boundary conditions are that each coordinate should have the same value at the end of the cycle as it had at the beginning. This achieves a fast convergence on the steady state, which is of most interest in spectroscopic prediction, and ignores transient behaviour.

5.1.1 Solution method

At the suggestion of C.J.Marks, the boundary-value problem is addressed using the 'shooting' method given in *Numerical Recipes in C* [Press *et al.* 1988]. The mechanical state of

the moving parts, the important drive circuit signals, and the thermodynamic state of the gas in each of three parts of the modulator are combined into a 'state vector'. This is set at the beginning of the cycle, and integrated using a fourth-order adaptive-timestep Runge-Kutta algorithm to derive the state at the end of the cycle. The difference between the initial and final states generates a 'score' vector, whose magnitude in the steady state should be zero, as the state at each end of the cycle should be the same as that at the beginning. The initial value of each variable is perturbed in turn, and the integration is performed after each iteration, providing local gradients of the score vector with respect to each state coordinate. Using these gradients, the initial state is updated by a Newton-Raphson method, until the behaviour becomes periodic. When the states at the beginning and end of the cycle are the same, so too are the time-derivatives, because they depend only on the state variables, and not on the cycle phase or time. Therefore there is no need to ensure separately that the derivatives are continuous.

As the modelled processes are non-linear, several iterations are necessary to reduce the discrepancy to a tolerable level. The model continues iteration until the modulus of the score vector falls below a preset threshold value. Typically five iterations are required for convergence; the frequency stabilizes first, and the temperatures last, since all the rate-dependent processes contribute to the frequency gradient of the score vector, while the sensitivity of the model to the initial temperatures is relatively small.

The component of the 'score' vector for each variable is in most cases the difference between the initial and final values of the variable, scaled by the first guess value, to give a non-dimensional matrix of error derivatives with approximately unit magnitude, the form preferred by the Newton-Raphson algorithm.

5.1.2 Model state variables

The time is represented by the cycle phase. The beginning of the cycle is defined as the point where the differential-mode piston velocity is zero and the piston separation is equal to the demanded operating amplitude.

The positions and velocities of the two pistons are stored as their equivalent sums and differences, as the drive circuit (chapter 4) is controlled entirely by the piston separation. For the first integration, the common-mode piston position and velocity are set to a small proportion of the required differential-mode values. When the modulator is balanced, the 'sum' position and velocity are both very small compared to the 'difference' values. The modulator cannot be represented as perfectly balanced, as the error derivative matrix would be singular in that case.

The drive circuit has been discussed in chapter 4. The model includes representations of the drive amplitude (as sensed by the capacitor in the peak-detecting circuit), drive level (amplitude of the drive force) and DC correction level (mean force). The piston driving-force is sinusoidal in form, calculated from the drive level, the drive phase offset set in the phase locked loop, and the cycle phase. The drive level is held constant throughout the cycle, in a slight simplification of the real drive circuit, in which the time constant for the drive level is about 100 cycles. The 'drive level' coordinate is adjusted until the predictions for the piston separation at each end of the cycle are the same. The DC correction force is similarly held constant throughout each cycle, but adjusted until the time-integral of the piston position is zero (as in the physical system).

The frequency is an eigenvalue of the system, remaining constant throughout the cycle.

Its error term is provided by the discrepancy between the values of the differential-mode velocity at each end of the cycle. The first guess frequency must be chosen with some care, to ensure that the frequency error function used in the Newtonian iteration scheme falls on a linear section near the correct root. The initial trial values of the drive level and DC correction level must be also be carefully chosen to prevent the pistons from passing through each other. There is no diagnostic for this event; the volume between them becomes negative and usually the Runge-Kutta integration fails. In the real modulator, the pistons would hit each other, preventing stable running. Provided that the frequency and drive and DC levels are within these broad ranges, there seems to be no dependence of the converged state on the first choice of other parameters.

To calculate the thermodynamic changes in the gas, the PMU is divided into three volumes. These are the optical cell, the 'fore' region, between the pistons, and the 'back' region, on the outside of the pistons, which is connected to the molecular sieve. The cell is connected to the fore region by the cell tube (which allows easy passage of gas). The fore region communicates with the back region via the piston gap, which has a very low conductance on account of the small difference (about 75µm) between the diameter of the piston and that of the bore. The back region is connected to the gas reservoir (molecular sieve) by a tube of intermediate conductance. General expressions for the conductances of the gas paths between each region are given below.

Three thermodynamic coordinates are used for the state of the free gas in each region. The volume is, of course, obtained from the geometry. Although any other two independent coordinates could be used, I use temperature and gas amount as they yield the simplest equations. The gas amounts are represented in the form Nk_B, where N is the number of molecules and k_B is Boltzmann's constant, to reduce the number of multiplications or divisions by k_B and to reduce the magnitude of the exponent in the numerical representation of these quantities. The amount of gas adsorbed onto the walls is also represented in this way.

The thermodynamic coordinates are all freely adjustable, with the constraint that at equilibrium there is no net gas transfer to or from the molecular sieve or heat transfer to or from the pressure modulator body. The gas amounts and temperatures in each region for the first integration are determined from the ideal gas equation, considering an isothermal change of state of the gas, initially at the sieve pressure and wall temperature, as the pistons move to their positions at the start of the cycle, and assuming that no gas flows past the piston. Thus the amount of gas on each side of the piston remains constant, although it is redistributed between the 'fore' region (between the pistons) and the cell. The temperature of the gas in each region is set to the wall temperature, and the adsorbed amount is calculated from the free gas amount, on the basis of the reasoning described below.

5.2 Model physics

The time-derivative of the model state is calculated from the instantaneous state and integrated as described above. All the physics is contained in the calculation of the state derivative. This includes mechanical forces acting on the pistons, thermodynamic changes in the gas, and transfer of gas between each of the three regions. The following sections describe these in turn.

5.2.1 Mechanical forces

The forces on each piston arising from the springs, the viscous drag in the piston gap, the AC and DC components of the driving force, and the pressure difference are used to provide the accelerations, which are then resolved into common- and differential-mode components. Newton's second law of motion is used to derive the time-derivative of the velocities from the forces acting on the pistons.

The force provided by the springs includes terms proportional to the extension and to its cube, in order to take some account of the springs' expected non-linearity, *i.e.*,

$$F = k_1 x + k_3 x^3 \tag{5.1}$$

This is required to reproduce the observed small dependence of frequency on piston amplitude.

5.2.2 Gas flows

Consideration of the Reynolds numbers ($\approx$ 500 max) and the mean free path (a few microns) in the regions of gas flow indicate that the flow is viscous in all cases. For the cell and sieve tubes the flow rate is derived by considering the viscous shear force gradient at the boundaries of elemental annuli; the net force on the annulus as a result of the rate of change of velocity gradient across it is balanced by that produced by the pressure difference from one end of the tube to the other. This gives a paraboloidal velocity *versus* radius characteristic and an expression for the volume flow rate:

$$Q = \frac{\pi \Delta p R^4}{8 l \eta} \tag{5.2}$$

where Q is the volume flow rate, p is the pressure, R is the radius of the tube, l its length, and η the viscosity. There is [Newman and Searle, 1948] a correction to l of $1.64R$ from non-uniformity near the tube mouth and a correction to Δp of $-Q^2\rho/(\pi^2 R^4)$ from the kinetic energy imparted to the fluid. I have neglected this second correction as it is small, but retained the first.

The gap around the piston, where the thickness is much smaller than the radius, is treated as a straight section. Here gas flows both because of the pressure difference and because the moving piston leaves a 'wedge' of gas behind it, since the boundary condition at the wall is that the gas velocity is zero. In this case,

$$Q = \pi a t \left(u - \frac{\Delta p}{l \eta} \frac{t^2}{6} \right) \tag{5.3}$$

where a is the radius of the piston, t the thickness of the gap, and u the piston velocity.

The pressure of the injected gas quickly reaches equilibrium with that already in the cell because the pressure wave moves at the velocity of sound; the modulator frequency is much lower than that of a gas resonator of similar dimensions. Thus spatial pressure variations can be ignored, as they are dissipated in a very short time. Since the gas is compressible, and the gas is assumed to remain in the pipe long enough to relax to the wall temperature, in view of the short thermal time constant, the volume flow rates are multiplied by the mean pressure and divided by the wall temperature to give the rate of gas injection,

$$\frac{d(N k_B)}{dt} = \frac{Qp}{T} \tag{5.4}$$

5.2.3 Temperature changes

The temperature change in each region arises from adiabatic compression by a change in physical volume, adiabatic compression by the addition of more gas, heat conduction to or from the walls, and mixing of the added gas which is assumed to be at the wall temperature. The first three give the work and heat terms in the first law of thermodynamics, and the mixing changes the internal energy. The temperature change is deduced from the remaining change in internal energy.

The heat conduction is derived from calculations using the cylindrical symmetry of the cell. Each region is considered as being made up of one or more cylinders, in each of which the temperature distribution is that of the lowest radial and axial modes, i.e. a zero-order Bessel function in radius, and a half-sine distribution along the cylinder (see Appendix C). The resulting ratio of the heat conducted to the walls to the difference between the mean temperature of the gas and the wall temperature is

$$\frac{H}{\Delta T} = 2\pi^2 \sigma p_0^2 L \left(1 + \left(\frac{\pi}{p_0 s}\right)^2\right) \tag{5.5}$$

where σ is the thermal conductivity of the gas, p_0 is about 2.4 (see Appendix C), L is the total length of all the cylinders, and s, the shape, is the ratio of the length to the radius in each cylinder.

The approximation of the regions by cylinders is reasonable given the symmetry, although the CO_2 cell is conical in section with about a 1:10 slope, and the fore region is very complicated in shape. The back region is divided into 14 segments to allow for the springs, stops and magnets on both sides, and the fore region into 2 for the coupling spring. These regions are oblate and most heat flows along the axis of the PMU; the cells are prolate ($s > 5$) and most heat is transferred via the curved walls.

Combining the four terms above, the time derivative of temperature is

$$Nk_B \frac{dT}{dt} = \frac{(H/T)(T_w - T) - p\frac{dV}{dt}}{\zeta} + \frac{d(Nk_B)}{dt}(\gamma T_w - T) \tag{5.6}$$

where (H/T) is the ratio of conducted heat to temperature difference, T_w is the wall temperature, ζ is the ratio of the specific heat to k_B for one molecule, and γ is the ratio of specific heat capacities, $c_P/c_V = (\zeta + 1)/\zeta$.

The body of the modulator is assumed to be an infinitely-conducting heat sink, and therefore remains at constant temperature.

5.2.4 Pressure: equation of state

For the majority of gases, under the pressure and temperature régimes of pressure modulators, the ideal gas equation may be used to calculate the pressure for a given gas amount, volume and pressure. However, it is not obvious that this is a reasonable assumption for water vapour, particularly for the PMIRR H_2O modulator, because of the high operating pressure. In particular, if the saturated region is entered during the cycle, condensation will occur, reducing the number of free molecules and thus the pressure. However, condensation is not an instantaneous process, because of the finite impingement rate of

vapour molecules on the enclosing walls and each other, so supersaturation is also expected. The reduction of pressure in and near the saturation region, as a result of the form of the equation of state, or because of saturation, would also reduce the gradient dP/dV, and thereby the rate of change of PMU frequency with mean pressure. In extreme circumstances it could also cause a drop in frequency at high filling pressures and low temperatures, in agreement with experiment (chapter 3).

The equation of state for water has been the subject of a great deal of research. For the vapour phase, the equation of Keyes [1949],

$$log_{10}\frac{RT}{pv} = log_{10}\frac{\omega}{v} + \frac{\psi\omega}{v^2} \tag{5.7}$$

gives good agreement with experiment, especially the molecular volume at which saturation occurs. In equation 5.7, when p is in atm, v in cm^3/g, and T (in K) is below the critical temperature, the other terms take the following values [Eisenberg and Kauzmann, 1969]:

$$\tau = 1/T \tag{5.8}$$

$$\delta = 2.0624\exp(-0.87498/v) \tag{5.9}$$

$$\psi = \psi_0\left(1 + \frac{\psi_1}{v} + \frac{\psi_2}{v^2}\right) \tag{5.10}$$

$$\psi_0 = 1260.17\tau\exp(17.09 \times 10^4\tau^2) \tag{5.11}$$

$$\psi_1 = 305.6\psi_0\tau\exp(34.19 \times 10^4\tau^2) \tag{5.12}$$

$$\psi_2 = 0 \tag{5.13}$$

When the gas is uncompressed, *i.e.*, at large values of v and T, this equation is identical with that of Dieterici,

$$p = \frac{RT}{V-b}\exp(-a/RTV) \tag{5.14}$$

from which the well-known Van der Waals equation

$$\left(p + \frac{a}{V^2}\right)(V-b) = RT \tag{5.15}$$

can be formed using the first two terms of the expanded exponential. In equations 5.14 and 5.15, b represents the volume of the molecules themselves, and a is a factor dependent on the strength of the cohesive forces in the substance, which are assumed to be proportional to the square of the density [Roberts and Miller 1956]. The values of a and b may be derived from the critical data for the substance.

Figure 5.1 shows a comparison between 300K isotherms calculated from the ideal gas, Van der Waals, Dieterici and Keyes equations of state for water around the pressure and temperature of the PMIRR flight model H_2O PMU. The differences between the pressure predictions of all of these equations are extremely small. Since the Keyes equation is in excellent agreement with the measured state of water under these conditions, and since the differences between the pressures predicted by Keyes's equation and the ideal gas equation are extremely small compared with other uncertainties in the model (adsorption, for example), and the ideal gas equation is much quicker to compute, the ideal gas equation is used for water vapour as well as for CO_2 and N_2O.

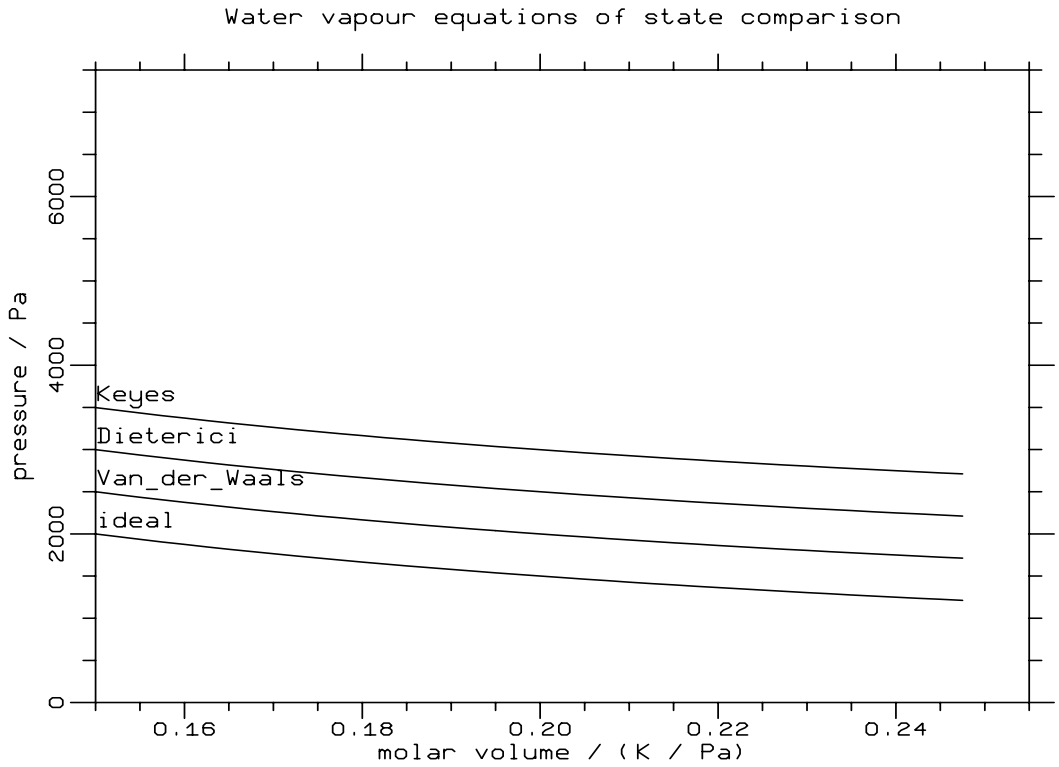

Figure 5.1: 300K isotherms for water vapour near the operating point for the PMIRR H_2O PMU. Each is offset by 500 Pa for clarity; the ideal gas curve gives the true pressure. The saturated vapour pressure at this temperature is about 3500 Pa. At the minimum volume shown on this graph, the ideal gas equation predicts a pressure of 2000 Pa, the Keyes equation 1997.99Pa , and the Van der Waals and Dieterici equations 1999.99Pa using constants derived from the critical data.

5.2.5 Adsorption

Adsorption, the tendency of gas molecules to remain on the surface of materials in the solid phase by physical or chemical bonding, is an important factor in the thermodynamics of the pressure modulator. According to a simple theory of chemical adsorption for the majority of adsorbates, molecules of the gas are bound to sites in the lattice of the solid surface, forming a monomolecular layer, and preventing further adsorption at those sites, effectively reducing the available surface area and causing a saturation effect. Physical adsorption, however, can lead to the presence of multiple layers on the adsorbent, and the semi-empirical Brunauer-Emmett-Teller (BET) equation [Brunauer *et al.* 1938] is generally used for adsorption isotherms at equilibrium:

$$\frac{p/p_0}{a(1-p/p_0)} = \frac{1}{a_m c} + \frac{c-1}{a_m c}\frac{p}{p_0} \tag{5.16}$$

In equation 5.16, p is the pressure, p_0 is a saturation pressure, a is the adsorbed amount, a_m is the amount adsorbed into a monolayer, and c is a constant, related to the latent heats of adsorption Q_A and vaporization (Q_V):

$$c = \exp\left(\frac{Q_A - Q_V}{k_B T}\right) \tag{5.17}$$

The form of the adsorption isotherms for a particular adsorbate-adsorbent combination is determined by the coefficient c. When c is very large, the adsorbed amount reaches saturation; this is appropriate when a monolayer is expected. For most cases, c is in the

range of 3–4 to several hundred, and the pressure saturates at p_0. In a very few cases, when the latent heat of vaporization is smaller than that of adsorption, c is much less than 1. The values of these latent heats have unfortunately not been measured for many of the gas/surface combinations in the PMIRR PMUs.

For CO_2 and N_2O, adsorption is a minor effect at the operating temperatures of the PMIRR PMUs, as the affinity of the gas (adsorbate) molecules for each other is much smaller than that for the walls (adsorbent), and thus c is expected to be large. However, for H_2O, hydrogen bonds can form between a hydrogen atom of one water molecule in the liquid phase and one of the lone pairs in the oxygen atom of another, greatly increasing the adsorbate-adsorbate affinity. This has the result that more water can be adsorbed onto the first monolayer, and still more onto subsequent layers, so that a large number of layers can be accumulated. Thus water vapour molecules impinge upon a surface which is essentially liquid, and c is expected to be fairly small.

Venters [1991] has suggested that the dynamics of the adsorption process are essentially the same as those of condensation onto the vessel walls. (Surface tension effects [Newman and Searle 1948] imply that condensation will occur preferentially onto flatter surfaces, and therefore the interaction with the walls is more significant than aggregation into droplets in the bulk of the vapour.) His development of the theory is questionable, as it uses the equilibrium between liquid / adsorbate and *saturated* rather than unsaturated vapour, but, with this important distinction in mind, a first-order differential equation may be deduced as follows. From kinetic theory, the number of gas molecules impinging on area A in unit time is

$$\frac{dN_i}{dt} = -\alpha A \frac{\bar{c}}{4} \frac{N_g}{V} \tag{5.18}$$

where N_g is the number of gas molecules in volume V, $\bar{c}$ is the mean speed of the molecules, and α is the 'sticking factor', *i.e.*, the probability that an impinging molecule will remain attached to the surface.

The number leaving the liquid in unit time is

$$\frac{dN_l}{dt} = \frac{A}{6r_s^3} \bar{c} \exp\left(-\frac{\lambda}{k_B T}\right) \tag{5.19}$$

where r_s is the mean separation of the adsorbed molecules on the adsorbing surface, λ is the latent heat of adsorption per molecule, and T is the temperature. By hypothesis, the mean speed in the two phases is the same if they are in thermal equilibrium. Equations 5.18 and 5.19 may be combined, to give the net loss from the gas:

$$\frac{dN_g}{dt} = \frac{A}{6r_s^3} \bar{c} \exp\left(-\frac{\lambda}{k_B T}\right) - \alpha A \frac{\bar{c}}{4} \frac{N_g}{V} \tag{5.20}$$

where N_g is the number of molecules in the gas phase. Using an integrating factor $\exp(\alpha A\bar{c}/4V)$, and an integrating constant C, equation 5.20 may be solved:

$$N_g = \frac{2}{3r_s^3 \alpha} \exp(-\lambda/k_B T) + C \exp\left(-\frac{\alpha A \bar{c}}{4V} t\right) \tag{5.21}$$

The nature of this equation, of exponential decay to a constant value, suggests that the saturated vapour pressure is associated with the first term. Venters [1991] has shown

from an Arrhenius plot that the temperature variation of the saturated vapour pressure for water conforms well to this form. The sticking factor α may thus be calculated:

$$\alpha = \frac{2k_BT\exp(-\lambda/k_BT)}{3pr_s^3} \tag{5.22}$$

Comparison with measured saturated vapour pressures implies that around 300K, α should have the value of about 6×10^{-4} for water.

The saturated vapour pressure term in equation 5.21 may also be substituted into the Clausius-Clapeyron equation,

$$\frac{dp}{dT} = \frac{\lambda}{T(v_l - v_g)} \tag{5.23}$$

in which v_l and v_g are the volumes occupied by one molecule in the liquid and gas phases respectively, to yield a slightly different expression for α,

$$\alpha = \frac{2k_BT\exp(-\lambda/k_BT)}{3pr_s^3}\left(1 + \frac{k_BT}{\lambda}\right) \tag{5.24}$$

which gives values which are generally 5% larger than those derived from equation 5.22, putting a limit of that order onto the self-consistency of this adsorption model.

In equation 5.19, the de-adsorption rate is a function of the temperature and the adsorbate density. Since it is possible at any pressure below the saturated vapour pressure, and at any temperature, to reach an equilibrium between the adsorbed and free gas, the density must be a function of the amount adsorbed. The nature of this function can be deduced from the BET equation 5.16. From equation 5.21, the pressure at constant temperature is proportional to $1/r_s^3$. Thus

$$\frac{p}{p_0} = \frac{r_0^3}{r_s^3} \tag{5.25}$$

where r_0 is the minimum possible separation of the molecules, derived from the density of the substance in the liquid phase. a_m, the maximum number of molecules which can be adsorbed into a monolayer, is A/r_0^2, and a is equal to the number of molecules actually adsorbed, N_A. Rearranging equation 5.16 results in a quadratic expression for p/p_0, which may be solved:

$$\frac{p}{p_0} = \frac{r_0^3}{r_s^3} = \frac{(A/N_Ar_0^2 - 1)c + 2 \pm \sqrt{\left((A/N_Ar_0^2 - 1)c + 2\right) - 4(1-c)}}{2(1-c)} \tag{5.26}$$

From the boundary conditions for very small and very large adsorbed amounts, which are that the pressure p should be zero and p_0 respectively, the negative sign is required in the numerator of equation 5.26.

Some adaptations of this theory are required in order to apply it to pressure modulators. The adsorbed gas and free gas are at different temperatures, so the mean molecular speeds in the two terms on the right hand side of equation 5.20 are different. This causes the adsorbed amount to vary throughout the modulator cycle. The adsorbing areas are considerably greater than those predicted by geometry alone, because of the roughness of the machined surfaces. An area multiplier or 'roughness factor' of up to 100 [Elsey 1976] is required to give the effective area. In the model, only one roughness term is used, and

applied to all the regions. The expressions 5.18 and 5.19 are used for the adsorption and de-adsorption rates, respectively, and 5.26 is used to calculate the adsorbate density. The sticking factor, different for each gas, is discussed below.

The optical effect of the adsorbate may be estimated by calculating the expected number of adsorbed layers in the water vapour cell. The area occupied by each adsorbed molecule is approximately d^2, where d is the molecular diameter (about 0.3 nm for H_2O). The number of layers is therefore approximately $N_A d^2/A$, where N_A is the number of adsorbed molecules in a region and A is the adsorbing surface area. For the operating conditions of the PMIRR H_2O PMU, the model predicts that 7.2×10^{17} molecules are adsorbed in the cell. This corresponds to an average of 1.9 molecular layers, assuming a roughness factor of 6.62 (derived from frequency fits). Optical absorption by this amount of water is negligible in comparison with that in the free gas. For CO_2, a maximum of one layer is expected, the effect again being negligible.

5.2.6 Operating parameters

Most of the physical parameters required to run the model are derived from the design drawings and measurements of the PMU, cell, and drive circuit, and from published tables of physical properties. The remaining parameters, (the spring strengths, surface roughness, and, in the case of CO_2, the latent heat and sticking factor) were estimated from least-squares fits to the measured frequency characteristics of the modulators, using the Levenberg–Marquardt non-linear method described in *Numerical Recipes in C* [Press *et al.* 1988]. For H_2O, the measured frequencies at 8 pressures and 2 amplitude settings were used for the flight model PMU; for the brassboard modulator at 295K, 4 pressures and three amplitudes, and for the same modulator at 298K, 5 pressures and three amplitudes. For CO_2, 6 pressures and three amplitudes were used for the flight model modulator, and 3 pressures and 3 amplitudes for the brassboard.

The three spring constants have equivalent effects on the differential mode behaviour, so only the coupling spring strength was fitted. The spring non-linearity term, *i.e.*, the extension which produces equal force components from the linear and cubic terms in equation 5.1 ($\sqrt{k_1/k_3}$), was similarly fitted, in order to match the observed variation of frequency with amplitude as far as possible. It was assumed to be the same for all the springs in each PMU.

For water, the latent heat λ is the latent heat of vaporization. The values used were interpolated from the data published in *Kaye and Laby* [1973], and a sticking factor was calculated from the saturated vapour pressure, using equation 5.22. This is not appropriate for CO_2 and N_2O, since these molecules tend to bind to the walls more readily than to each other, so that the latent heat of adsorption is different from that of vaporization. For these gases, the first term in equation 5.21 is not the saturated vapour pressure, but a saturated pressure for adsorption, *i.e.*, the pressure which causes all available sites to be occupied. The measured frequency *versus* pressure characteristics for CO_2 show a decrease in gradient at high pressures, and I estimated the saturation pressure to be about 200mb from these curves. This estimate was probably too low, as the resulting modelled frequency/pressure characteristic is too curved (see below), but it enabled a sticking factor to be calculated.

The coefficient c in the BET equation 5.16 exerts little control over the frequency or pressure cycle, so it was estimated as 3 for water, for which the adsorption process is essentially

one of condensation, and 1000 for CO_2 and N_2O, for which monolayers are expected. Surface roughness factors of about 5–10 were obtained from fits to water vapour frequency measurements, and factors of about 100 from fits to the CO_2 data. This large number results from the complete ignorance of the latent heat and sticking factors for CO_2 in view of the many different material interfaces in the modulator. The covariances between the roughness, latent heat, sticking factor and temperature are large, as all these terms trade off against each other in the adsorption rate equation 5.20.

The viscosity is treated as a linear function of temperature, the gradient and intercept of which are obtained by a least-squares fit to the data published in the CRC handbook [1984], Kaye and Laby [1973] and Newman and Searle [1948]. In the region of 300K a linear relation is quite adequate for all three gases, and for water vapour provides a better fit than the usual expression (Sutherland's law)

$$\eta = \eta_0 \frac{T^{3/2}}{T + T_S} \tag{5.27}$$

where T_S is Sutherland's constant for the gas. The specific heat capacities are as published in the CRC handbook [1984], and as their variation with temperature is small they are treated as constant. The thermal conductivity differs substantially from the kinetic theory value (proportional to ηC_V) for complex molecules such as CO_2 and H_2O [Newman and Searle 1948], and so these are included separately. The constants used are given in table 5.I. All the values for CO_2 and N_2O are similar as their molecular weights are the same.

Parameter	unit	H_2O	CO_2	N_2O
Thermal conductivity	W m^{-1} K^{-1}	0.01781	0.01660	0.01734
Specific heat ratio (γ)		1.334	1.300	1.324
c_v	J kg^{-1} K^{-1}	1397.0	648.8	662.8
c_p	J kg^{-1} K^{-1}	1864.0	843.4	877.5
Fitted viscosity at 0 K	10^{-6} Pa s	-1.821	1.198	0.7712
Fitted viscosity change per K	10^{-6} Pa s K^{-1}	0.03856	0.04650	0.04689
Latent heat of vaporization/adsorption	10^6 J kg^{-1}	2.43741	0.908	0.908
BET c coefficient		3	1000	1000

Table 5.I: Gas constants used in the model

5.3 Model results and comparison with measurements

The model produces an output file, which contains the cell pressure and temperature at 256 points, equally spaced around the cycle, and their variances and covariances. This is used to calculate the optical transmission, as described in chapter 7. The frequency and its total uncertainty are also written to the output file, as are the contributions to the uncertainty in frequency, piston position, cell pressure and cell temperature. Graphs of the cyclic variations of the model variables can also be plotted (some are shown below).

The predictions of the operating frequency and cell pressure cycle made by the model were compared with experimental results for validation. Frequency was also used to fit the unknown model parameters, described above. In principle, it should also be possible to compare the modelled and measured power consumptions. However, for the pressures

used in most of the work undertaken, the static dissipation of the drive circuit is similar in magnitude to or larger than the dynamic dissipation of the modulator, and this prevents a useful comparison.

To measure the pressure cycles, a small piezoelectric strain-gauge pressure sensor was included in the brassboard PMU cells. The sensors used had a large bandwidth (> 1kHz) and a small combined hysteresis and non-linearity (<1%). The varying component of the cell pressure was measured for H_2O and CO_2 at different filling pressures and piston amplitudes. The gauges were calibrated against a baratron over the range of pressures expected, and confirmed to be linear to better than 2%, the estimated limit of accuracy of the measurements. To compare these measurements with the model predictions, the PPO output was recorded simultaneously, and then used to align the model output with the frequency and phase of the measurements. Because there was a large DC offset in the pressure sensor output, only the AC component was measured, so that the modelled cycles were displaced by the mean pressure to allow direct comparison. In principle it would be possible to use the pressure cycle measurements in the fits as well, but this is not yet computationally feasible.

Quantitative estimates of the goodness of fit between the modelled and measured frequencies and pressure cycles are made, on the assumption that the errors are random and normally distributed. This is only partly true in practice. Most of the uncertainties in the measurements are random, but those arising from uncertainties in the dimensions of the modulator, for example, are actually systematic. However, the dominant uncertainty in the modelled frequencies and pressure cycles arises from the uncertainty in the wall temperature, which does vary from measurement to measurement and may therefore be considered random. Since the other uncertainties are much smaller, it is not unreasonable to combine all of them in this way.

The formal probability that the measured and calculated values agree on the basis of the assumptions stated above is [Press *et al.* 1988]

$$P = \frac{1}{\Gamma(\nu/2)} \int_{\chi^2/2}^{\infty} e^{-t} t^{\nu/2-1} dt \tag{5.28}$$

in which Γ is the gamma function, ν is the number of degrees of freedom, and

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(y_i - y(x_i; a_1 \ldots a_M)\right)^2}{\sigma_i^2 + \sigma^2(x_i; a_1 \ldots a_M)} \tag{5.29}$$

in which x_i, y_i are the data points, $y(x_i; a_1 \ldots a_M)$ is the calculated value of y at x_i, with model parameters $a_1 \ldots a_M$. σ_i is the standard deviation in the measurement, and $\sigma^2(x_i; a_1 \ldots a_M)$ is the variance of the calculated value of y.

In all the comparisons made in this thesis, including those for the transmission measurements in chapter 7, the number of degrees of freedom is the number of data points less the number of fitted parameters. This allows comparison of data with measurements even when the data has been used for fitting model parameters, provided that enough measurements have been made.

The comparison fails to be meaningful when the measured and calculated x values do not agree. For example, there was a phase offset between the measured and calculated pressure cycles (see figures below). The large gradient of pressure with cycle phase causes

χ^2 to be unreasonably large. The relative amplitudes and phase offsets between the mean measured and modelled pressure cycles have therefore been estimated as an alternative comparison.

5.3.1 Results for water vapour

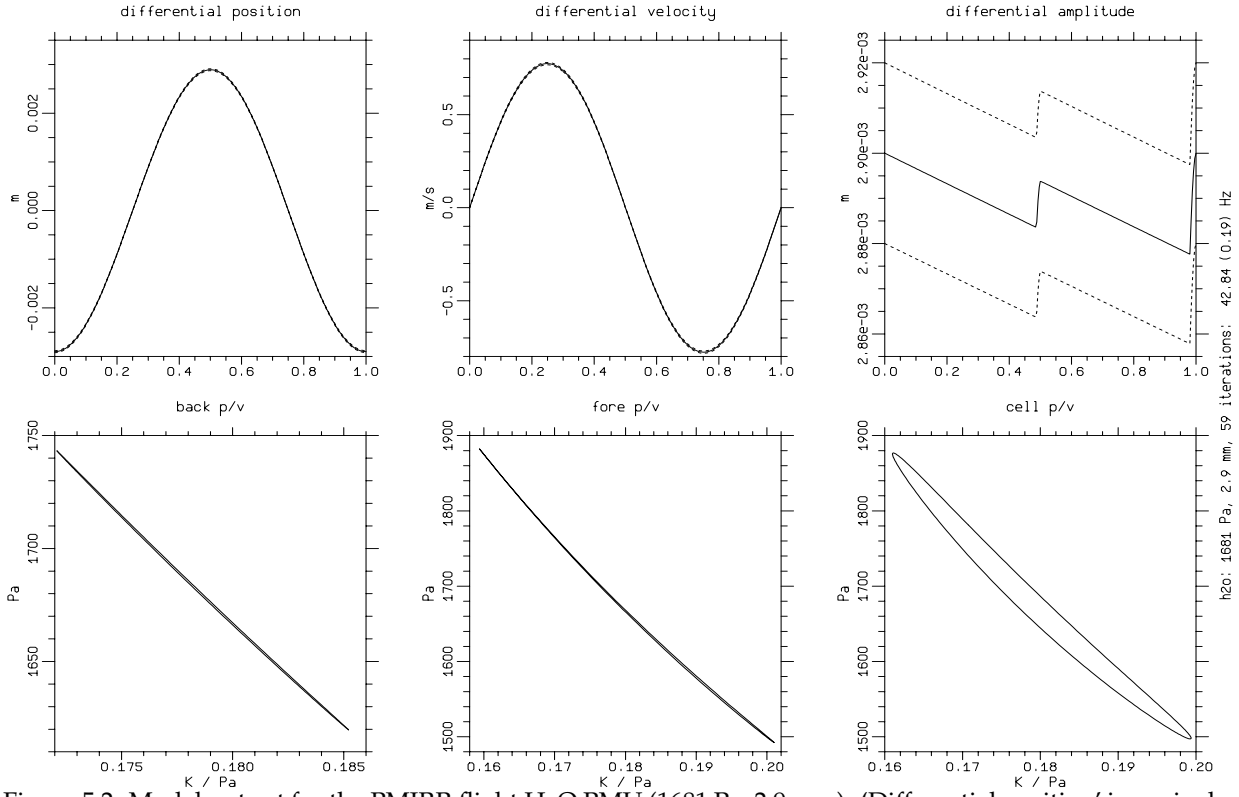

Figure 5.2: Model output for the PMIRR flight H_2O PMU (1681 Pa, 2.9 mm). 'Differential position' is equivalent to the piston separation; 'Differential velocity' is its time-derivative. 'Differential amplitude' is the distance equivalent to the voltage across the amplitude-sensing capacitor in the drive circuit. No uncertainties are given in the indicator diagrams, for clarity.

The output of the model for the PMIRR flight H_2O PMU operating conditions is shown in figures 5.2, 5.3 and 5.4. In these figures, the abscissa in each graph is the elapsed time in cycles, except in the indicator diagrams, when it is the molecular volume divided by Boltzmann's constant. The dotted lines show the ranges expected for each coordinate, when the expected uncertainties are added or subtracted. These uncertainties are the root-sum-squares of the individual deviations calculated by perturbing each model parameter by its expected error. Most of these parameters and errors are specified in the modulator description file; mean pressure, piston amplitude and body temperature errors are given when each model run is commenced.

Because of the small physical distances in the fore region, the temperature cycle in the volume between the pistons is essentially isothermal, especially when the pistons are close together near the start and finish of the cycle. The isothermal nature of the cycles in all three modulator regions is shown by the 90° phase shift between pressure and temperature (figure 5.3). The indicator diagrams (figure 5.2) show that most of the work done on the gas is in the cell.

The measured and calculated frequencies for different pressures and amplitudes are shown in figures 5.5, 5.6 and 5.7. It is evident that the variation of frequency with

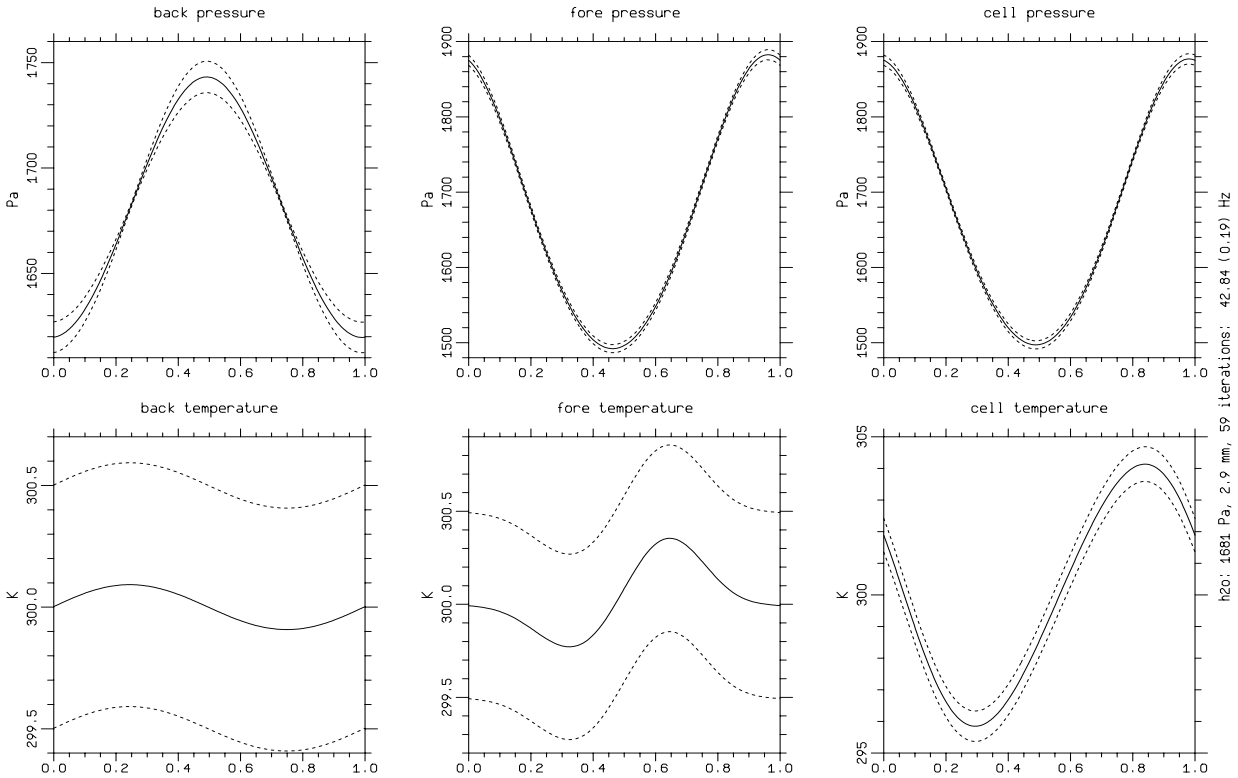

Figure 5.3: Model output for the PMIRR flight H_2O PMU (1681 Pa, 2.9 mm). The large phase offsets between temperature and pressure in each region indicate a largely isothermal cycle. The form of the temperature cycle in the fore region occurs because of the continuously-changing thermal relaxation time constants as the piston separation changes.

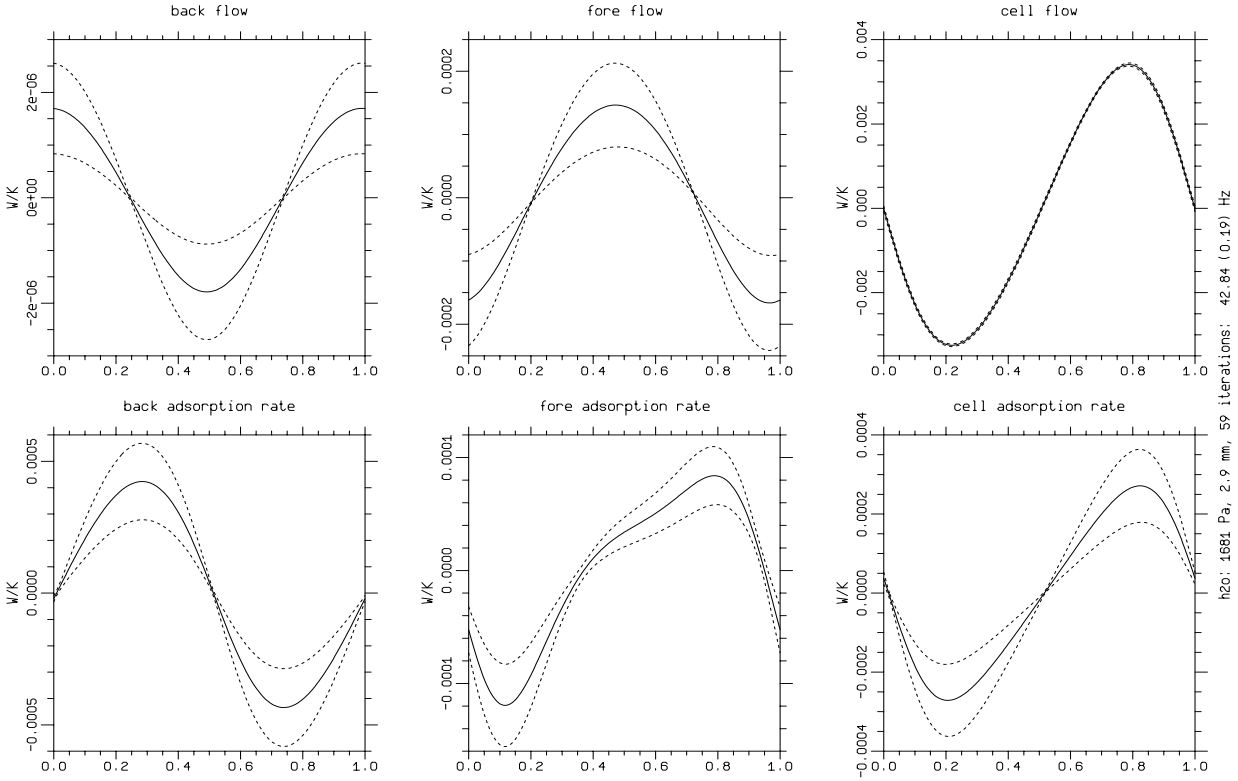

Figure 5.4: Model output for the PMIRR flight H_2O PMU (1681 Pa, 2.9 mm). 'Flow' is the amount of gas (in units of J/K or number of molecules multiplied by Boltzmann's constant) entering each region from that 'behind' it per second, *i.e.*, 'back flow' comes from the sieve; 'fore flow' occurs past the piston, and 'cell flow' represents gas moving along the cell tube. The adsorption rates are positive when gas is being deposited onto the walls.

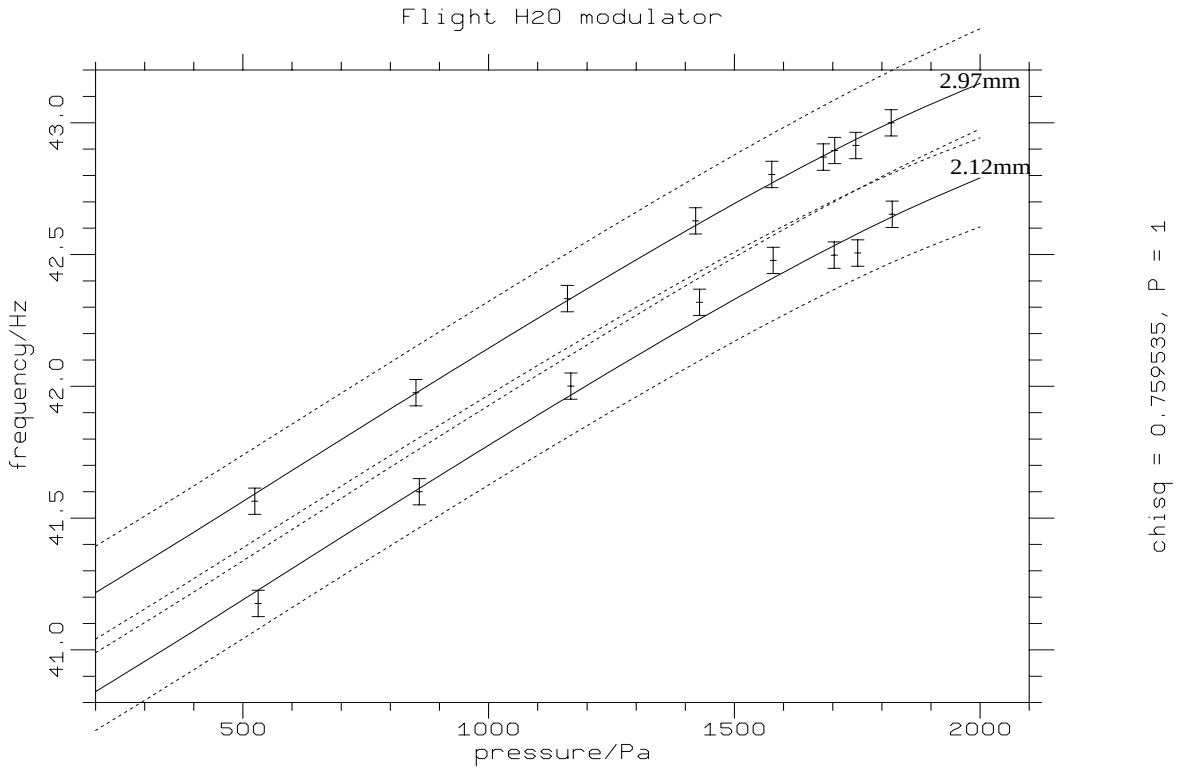

Figure 5.5: Measured and calculated frequencies for the PMIRR flight H_2O PMU against filling pressure, for piston amplitudes of 2.12 and 2.97 mm. The measured frequencies shown here form a subset of the frequencies measured during filling, which have been presented in chapter 3, and are those used to optimize the modulator model spring and roughness parameters.

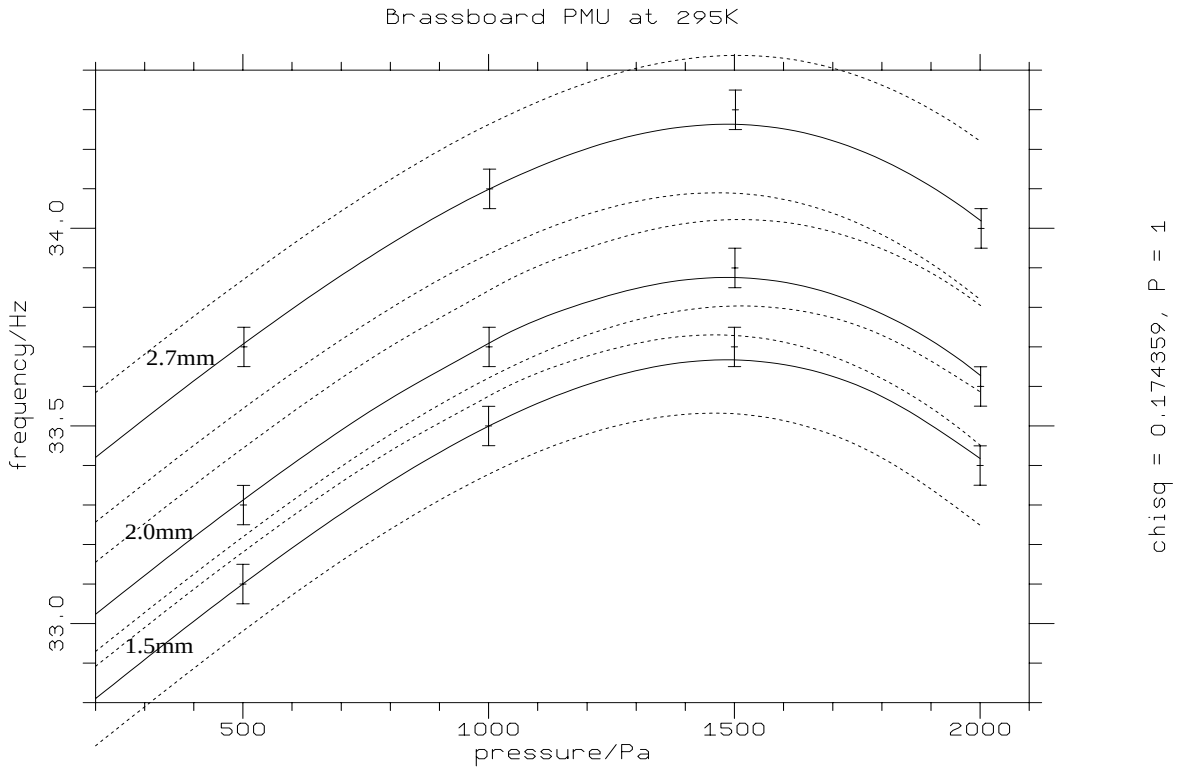

Figure 5.6: Measured and calculated frequencies of the brassboard PMU (at 295K) against H_2O filling pressure, for three piston amplitudes (1.5, 2.0 and 2.7mm).

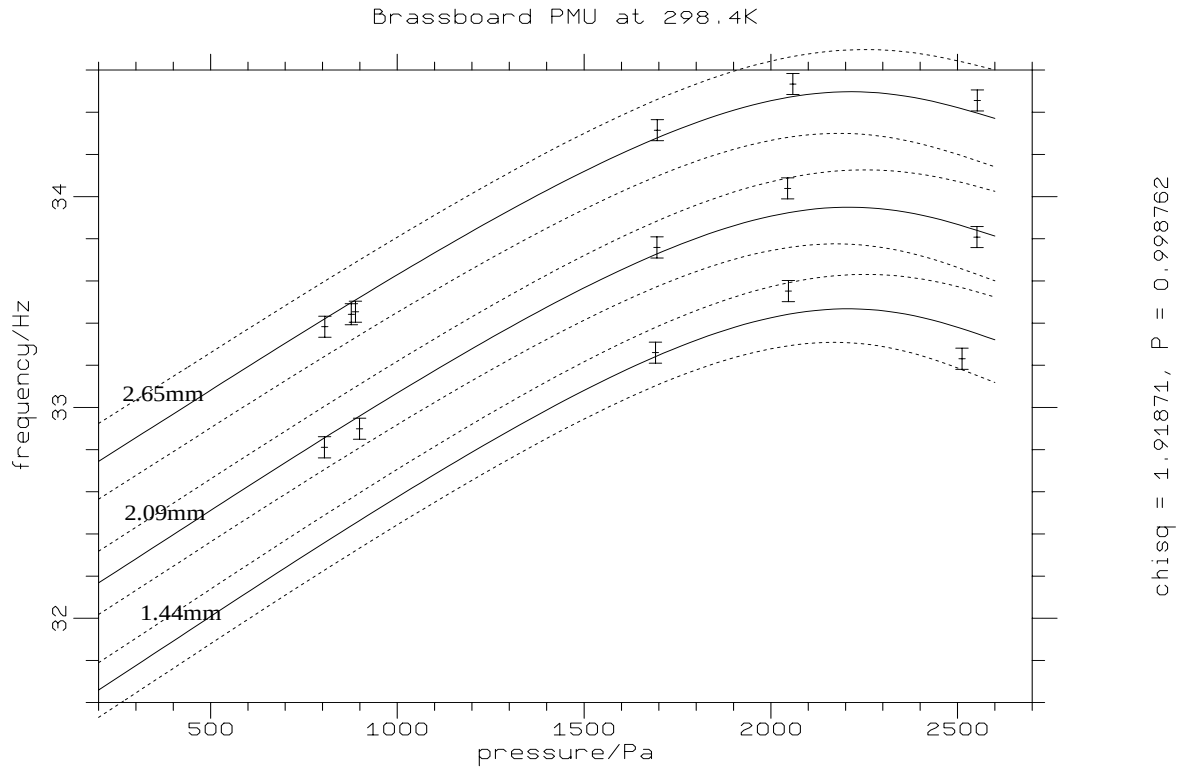

Figure 5.7: Measured and calculated frequencies of the brassboard PMU (at 298.4K) against H_2O filling pressure, for three piston amplitudes (1.44, 2.09 and 2.65mm). The measured frequencies shown here form a subset of the frequencies measured during the laser diode experiments described in chapter 6.

amplitude has been well represented in the model by the fitted cubic spring term. Without the term, the frequency is independent of amplitude. The fact that the variation of frequency with amplitude is largely independent of pressure indicates that the cause of the effect is mechanical rather than thermodynamic.

The frequency fits have an extremely high probability of agreement: 100% for the flight modulator (figure 5.5), 99.9% for the brassboard modulator at 298.4K (figure 5.7), and 100% for the brassboard modulator at 295K (figure 5.6). It is, however, unfortunately true that these high fit probabilities are caused by the large uncertainties in the calculated frequencies, particularly at high pressures.

Davis [1988] postulated that the reduction of PMU frequency at high pressures of water vapour (described in chapter 3) occurred as a result of adsorption. The results presented here seem to confirm this hypothesis, although the magnitude of the reduction is also controlled by the modulator geometry and piston masses. However, at least for the case of the PMIRR PMUs, the frequency reduction is caused more by the saturation effects in the adsorption representation than by the lower-pressure adsorption itself.

Figures 5.8–5.17 show the measured and modelled cell pressure cycles, the root sum square uncertainties (dotted), and the residuals divided by the uncertainties. The measured pressure cycles are larger than the calculated ones, and, in absolute terms, the discrepancy increases with filling pressure, from 9 to 24%, although the uncertainty in the calculated pressure cycle increases faster. The measured cycles lag the predicted cycle.

Table 5.II shows the frequency change and cell pressure cycle variance contributions from the perturbation of each parameter, from the model output for the flight modulator, at

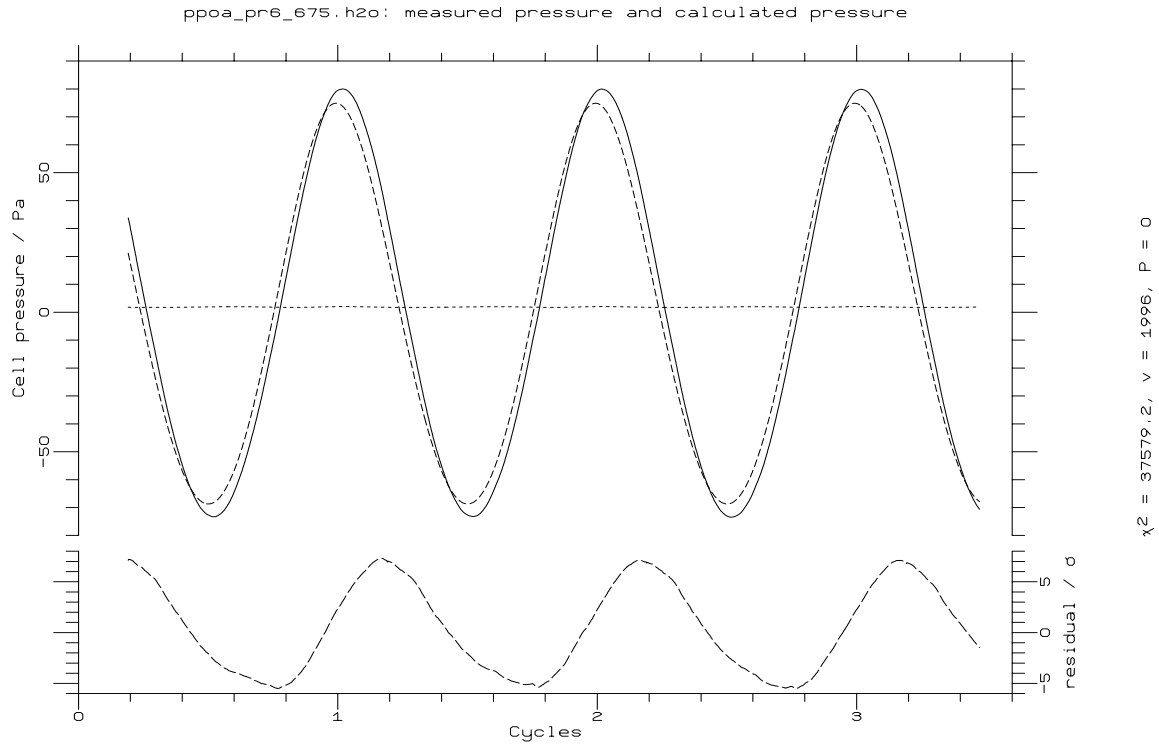

Figure 5.8: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 8.7mb mean pressure, 2.08mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 9% in amplitude, and lags it by 8.9° in phase.

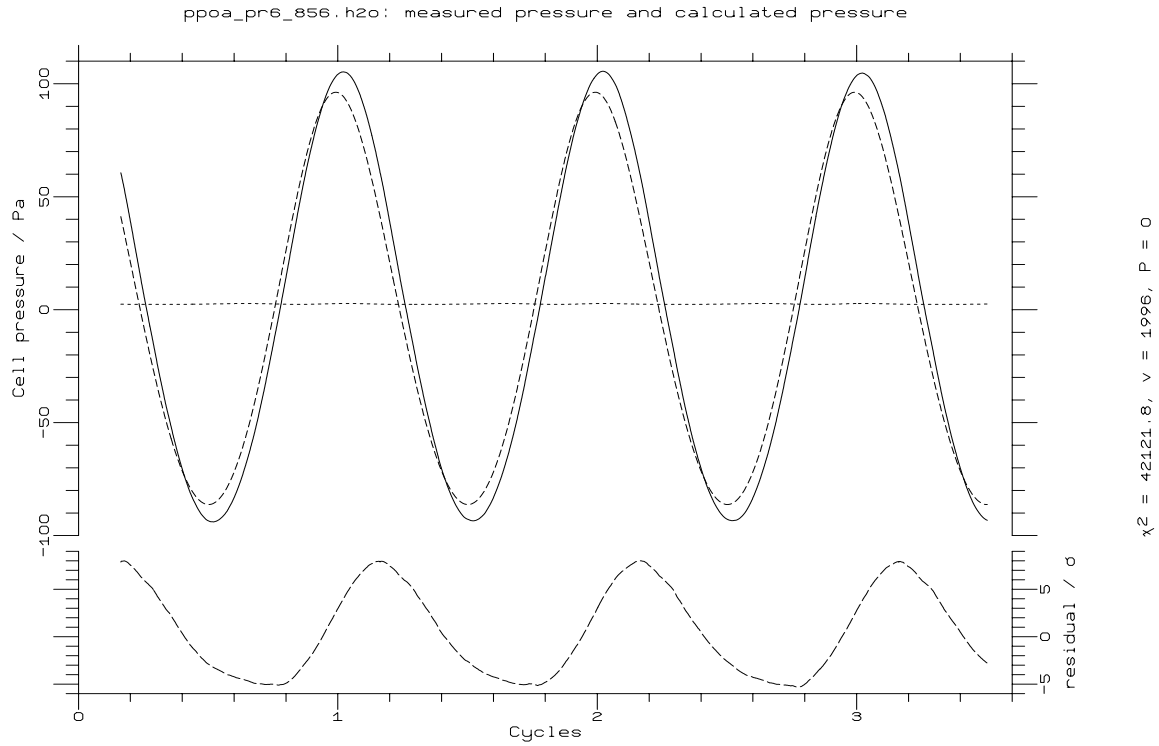

Figure 5.9: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 8.7mb mean pressure, 2.64mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 8% in amplitude, and lags it by 9.3° in phase.

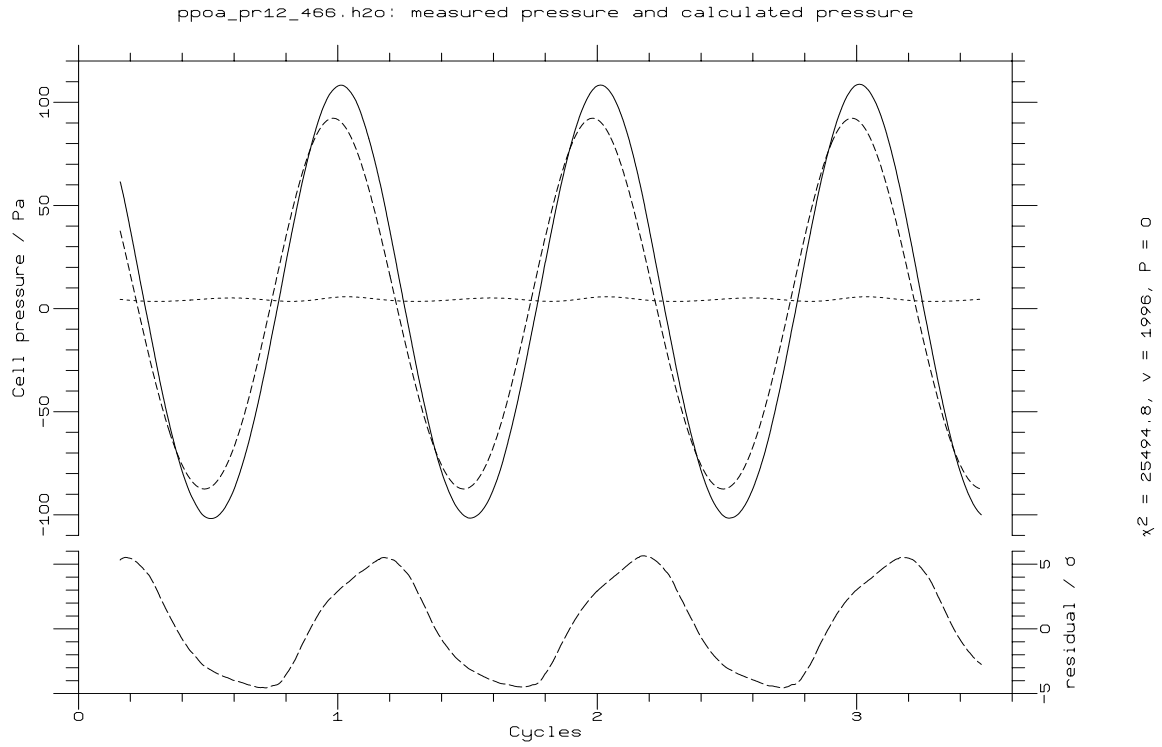

Figure 5.10: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 16.9mb mean pressure, 1.44mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 15% in amplitude, and lags it by 10.6° in phase.

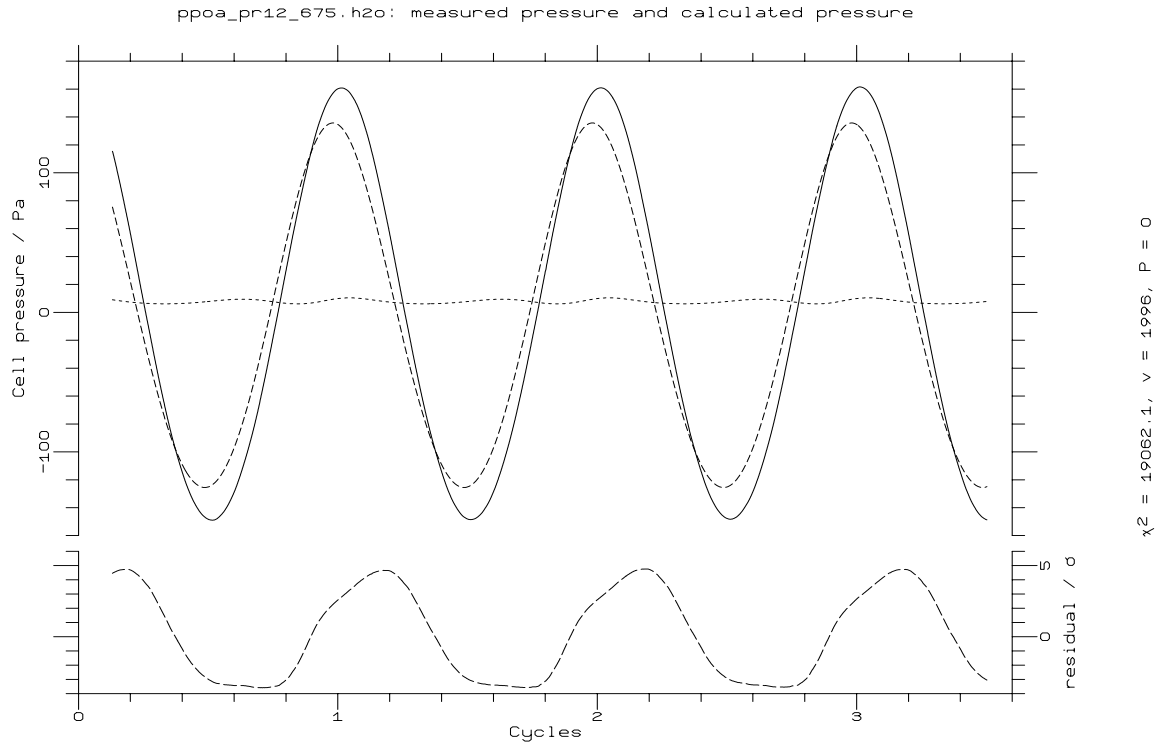

Figure 5.11: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 16.9mb mean pressure, 2.09mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 15% in amplitude, and lags it by 11.5° in phase.

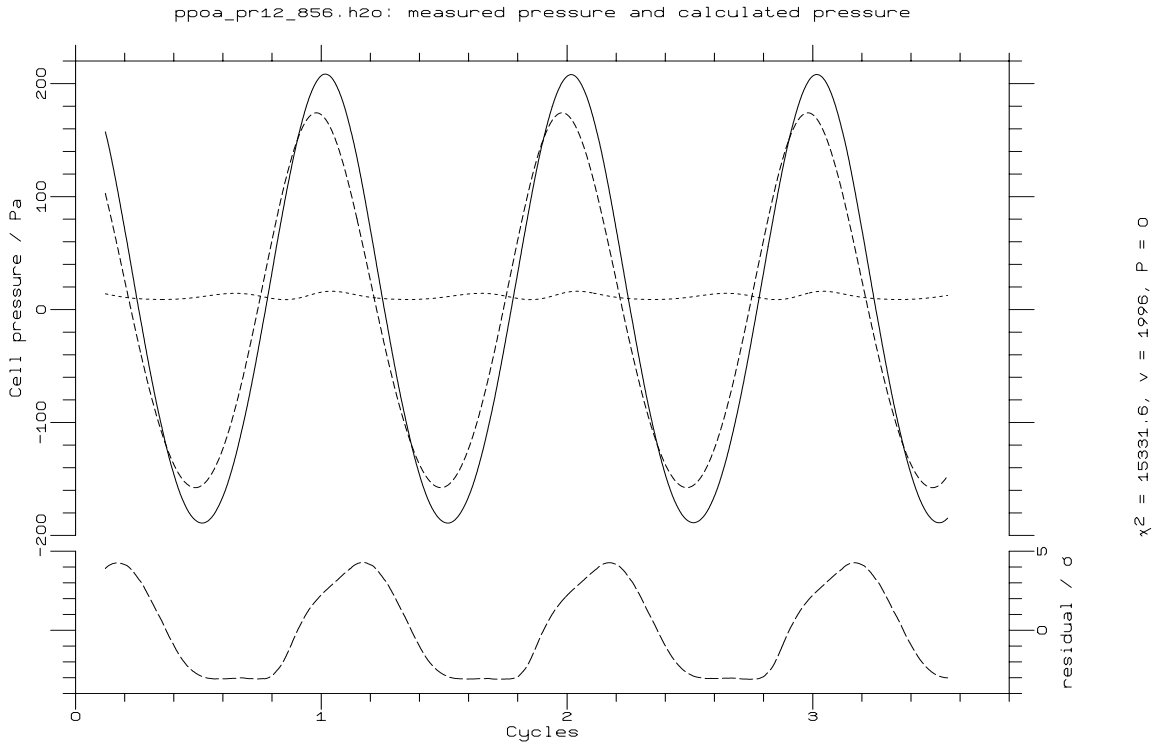

Figure 5.12: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 16.9mb mean pressure, 2.65mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 17% in amplitude, and lags it by 12.5° in phase.

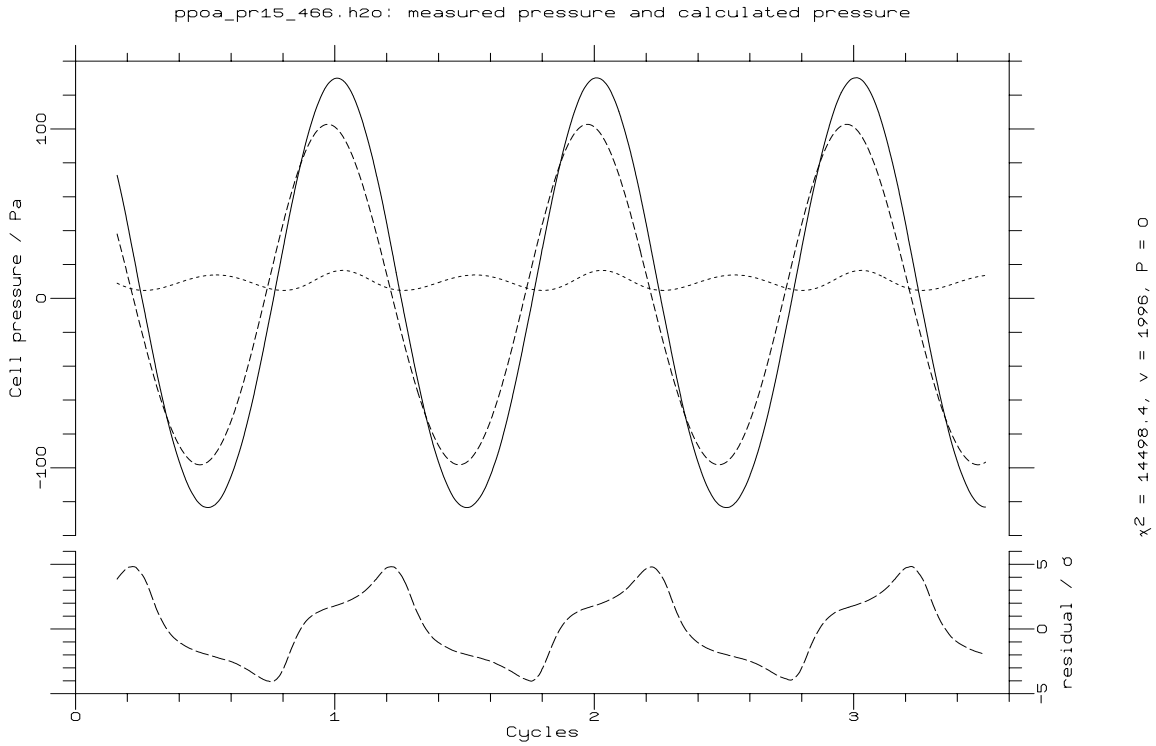

Figure 5.13: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 20.6mb mean pressure, 1.44mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 20% in amplitude, and lags it by 12.3° in phase.

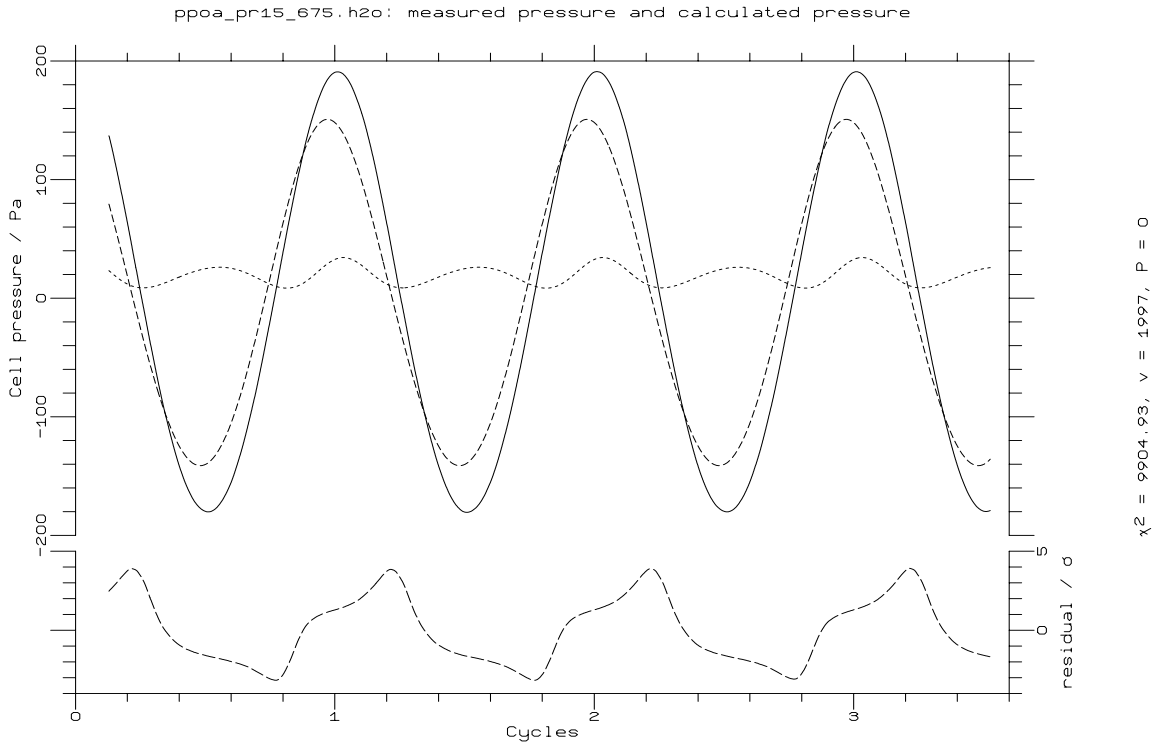

Figure 5.14: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 20.6mb mean pressure, 2.09mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 21% in amplitude, and lags it by 13.4° in phase.

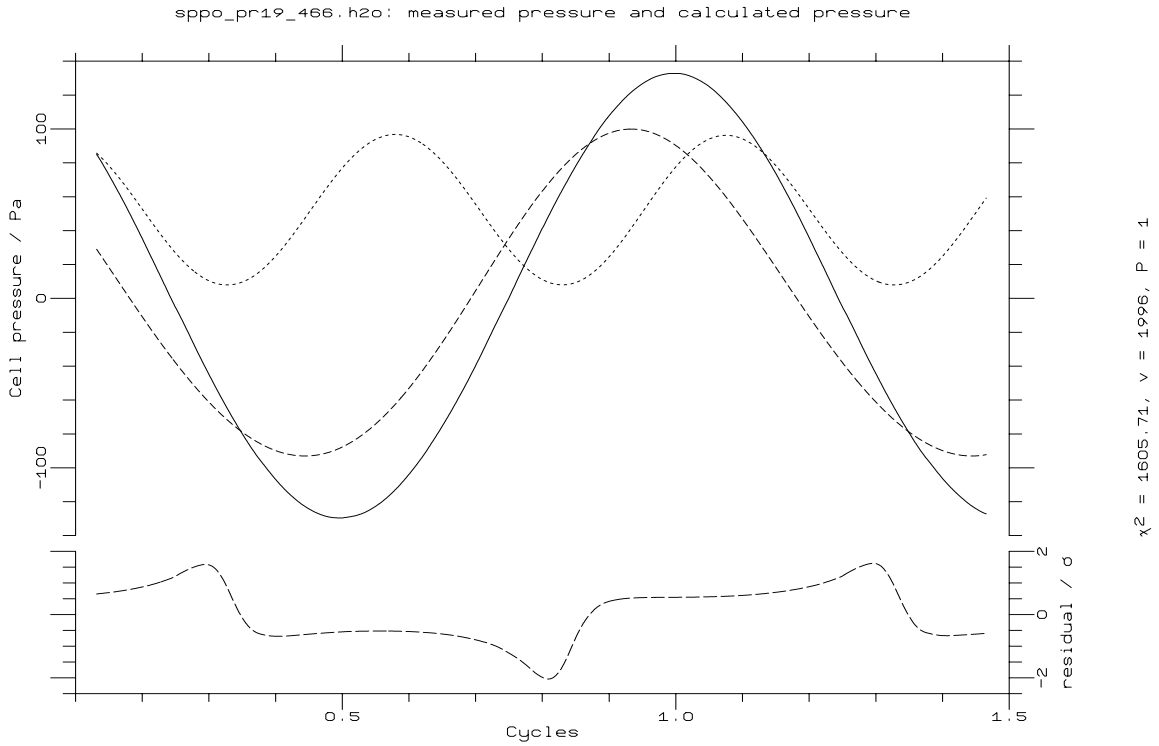

Figure 5.15: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 25.7mb mean pressure, 1.45mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 25% in amplitude, and lags it by 9.4° in phase, when the phase shift introduced by the drive circuit is removed.

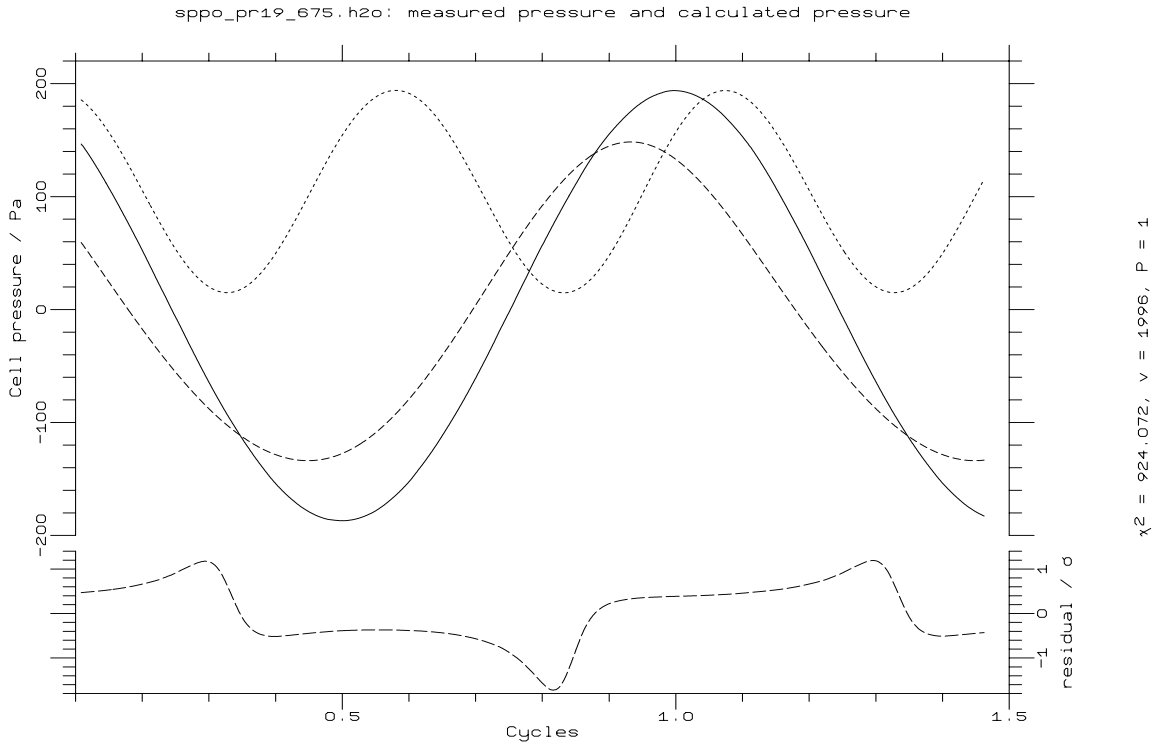

Figure 5.16: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 25.7mb mean pressure, 2.10mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 22% in amplitude, and lags it by 9.7° in phase, when the phase shift introduced by the drive circuit is removed.

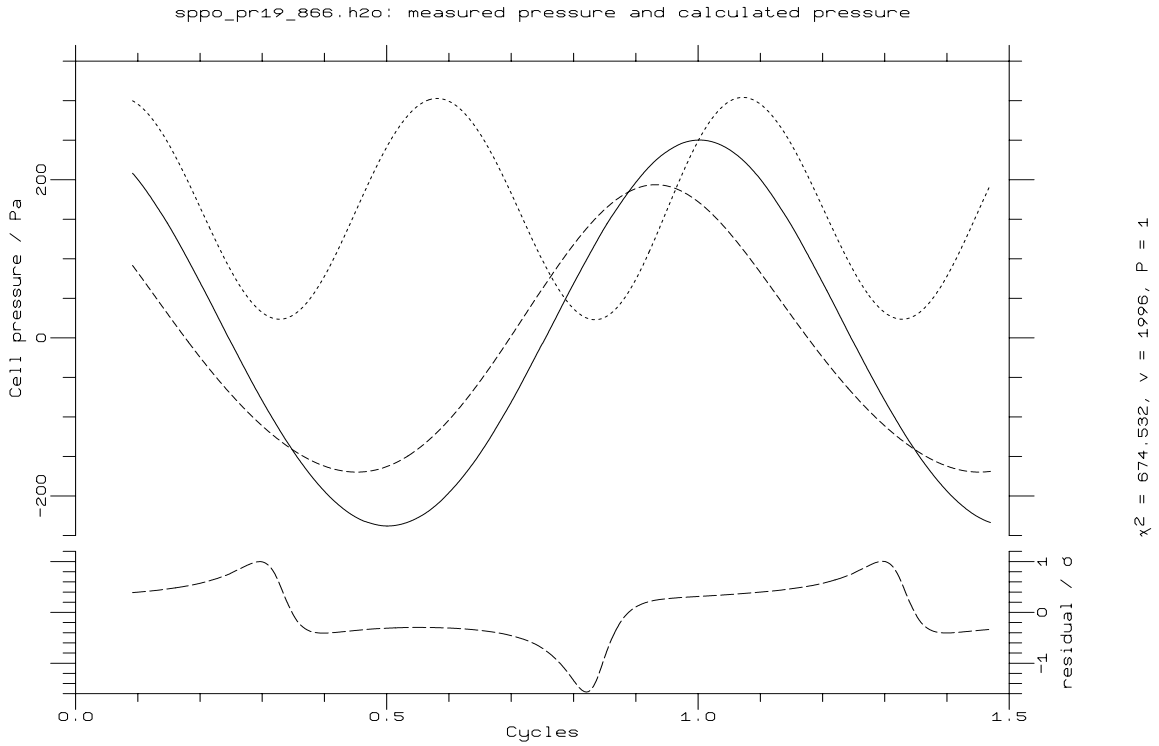

Figure 5.17: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for H_2O at 25.6mb mean pressure, 2.67mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by 24% in amplitude, and lags it by 11° in phase, when the phase shift introduced by the drive circuit is removed.

Varied	by	Δf/Hz	pressure cycle variance/Pa^2
Demanded amplitude	2e-05	0.00968914	0.933
Sieve pressure	1	0	0.000189
Wall temperature	0.5	0.01477354	1.06
Drive phase	0.035	0.01390372	0.00346
Suspension spring strength	10	0.05695979	0.0124
Coupling spring strength	3.95	0.09465776	0.0194
Spring cubic term	0.00067	-0.09883939	0.0291
Rest separation	2e-05	-0.001157159	0.108
Back length	0.01	-0.05204819	0.132
Piston mass	0.0001	-0.0797455	0.0169
Piston radius	2.5e-05	0.008822892	0.304
Piston gap	1e-05	-0.006042142	2.21
Gap length	5e-05	5.051741e-05	0.0329
Sieve tube radius	5e-05	1.247303e-05	0.000221
Sieve tube length	0.1	-4.483359e-06	7.93e-05
Cell radius	5e-05	-0.01185443	1.17
Cell length	5e-05	-0.00052879	0.0511
Cell tube radius	-5e-05	0.005818996	0.46
Cell tube length	0.0001	8.998772e-05	0.0179
Roughness	2.5	-0.0559186	3.91
BET c coefficient	1	0.0002925065	0.0207

Table 5.II: Frequency changes and pressure cycle variances for H_2O. 'e' represents a base 10 exponent. The numbers are taken directly from the cycle file produced by the model, so that the number of figures quoted is not an indicator of accuracy.

1681Pa filling pressure and 2.9mm piston amplitude. The sizes of the errors are estimated either from the measurement technique or, in the case of fitted parameters, from the covariance matrix produced by the fitting program. The largest contributions to the frequency uncertainty are from the spring constants, which provide an offset largely independent of pressure, and from the roughness, which alters the amount of gas transfer by adsorption. The cell pressure cycle variance is dominated by the roughness, but the piston gap is also significant, as it controls the conductance between the fore and cell regions, and therefore the 'leakiness' of the gas viscosity seal. The large dependence implies that the piston gap should be at least controlled, if not measured, to high precision, for repeatability of pressure cycle between different modulators.

The poor agreement between the modelled and measured pressure cycles is almost certainly due the detailed temperature structure in each volume (see Appendix B), and to overestimation of the amount of adsorption in the cell relative to the fore and back regions, since both are controlled together with one roughness factor. In connection with the temperature structure, it is significant that Irwin [1991] measured the emission signal from the brassboard modulator, and concluded that the cell cycle was nearly adiabatic. In that case the expected pressure cycle would be much larger too, by a factor of up to γ, which is 1.334 for water vapour (table 5.I). Overestimation of the adsorption would affect the pressure cycling as follows. The frequency behaviour is determined by the pressure cycle on either side of the pistons, and thus affected by the adsorption in the fore and back regions. The cell pressure cycle, however, is affected directly only by the adsorption in the cell itself. Enhancement of the adsorbing areas of all three regions together will therefore tend to

reduce both the frequency (which is achieved by the fit) and the cell pressure cycle. The phase shift could be caused by large adsorption or condensation in the cell tube, which is not considered as a separate thermodynamic volume, but divided between the cell and fore regions. Introducing such a reservoir into the connection between the fore region and the cell could effectively cause an imaginary component of the flow conductance, analogously to an RC electrical circuit.

5.3.2 Results for carbon dioxide

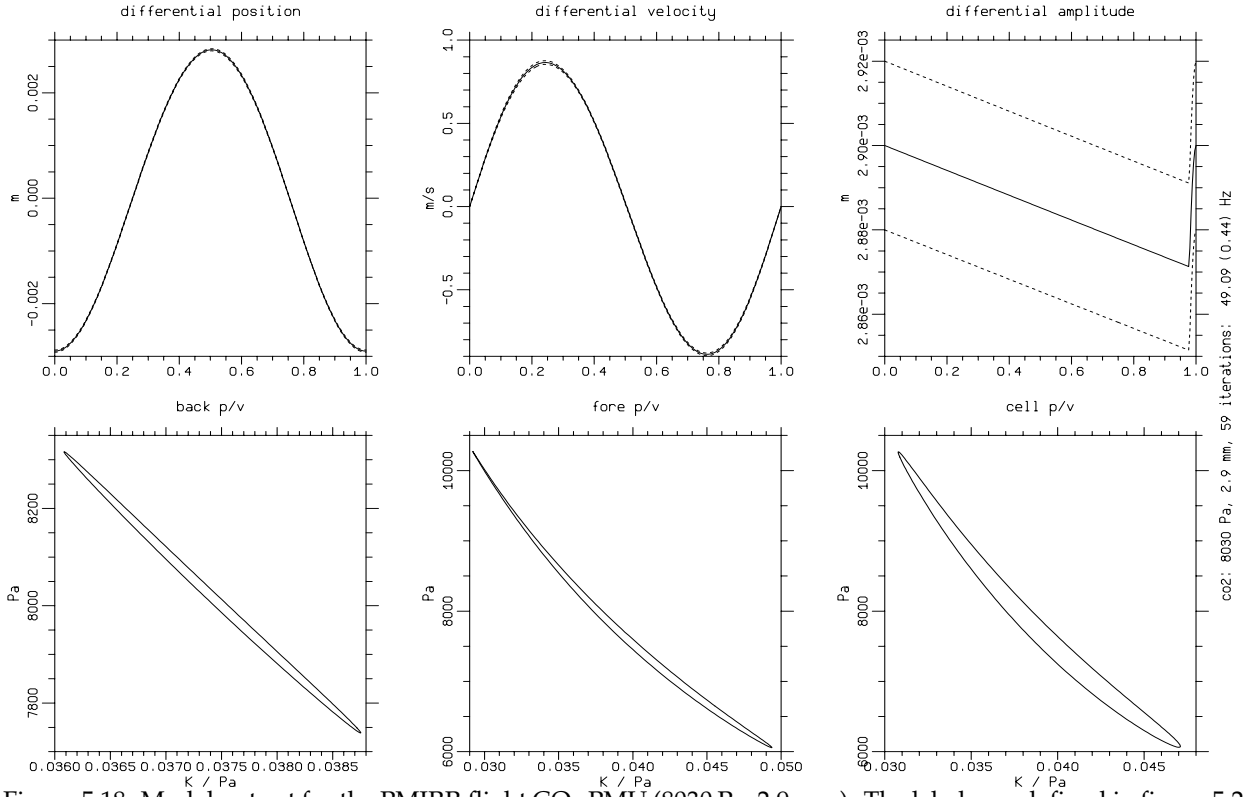

Figure 5.18: Model output for the PMIRR flight CO_2 PMU (8030 Pa, 2.9 mm). The labels are defined in figure 5.2. Note the larger asymmetry in the 'differential position' plot, which prevents the amplitude-sensing capacitor from recharging half-way through the cycle, as happens with H_2O (figure 5.2). A greater proportion of the work done on the gas occurs in the 'back' and 'fore' regions than is the case with H_2O (figure 5.2).

The output of the model for the PMIRR flight CO_2 PMU operating conditions are shown in figures 5.18, 5.19 and 5.20. The isothermal nature of the cycles in the fore and back regions is again evident from the 90° phase shift between pressure and temperature (figure 5.19). The CO_2 PMU exhibits almost adiabatic cycling in its cell, with very little phase shift between pressure and temperature. The phase shift between cell pressure and piston position in the CO_2 PMU is also greater than that for H_2O (figures 5.19 and 5.3) because of the increased flow rate through the cell tube (figures 5.20 and 5.4).

The predictions of the model regarding the frequency *versus* pressure characteristics of the modulators for CO_2 are accurate, with fitting probabilities of 99.2% and 94.7% for the flight model and brassboard PMU. The results are shown in figures 5.21, 5.22, 5.23 and 5.24; the uncertainty lines have been omitted from figures 5.21 and 5.23 for clarity.

The uncertainties in the calculated frequencies, particularly at high pressures, coupled with the large roughness factors (mentioned above) derived from the frequency fits, demonstrate the need for systematic measurements of the latent heats of vaporization and the sticking factors for CO_2 adsorbed onto the surfaces of the modulator.

Although the measured dependence of frequency on pressure decreases at high pressures, an effect which is not observed in the model unless adsorption is included, the amount of gas removed by adsorption has been overestimated. This is almost certainly due in part to the somewhat arbitrary values of the latent heat and sticking factors, but mostly because the temperature of the inside surface of the cell wall is greater than from the 300K external temperature, because of the poor thermal conductivity of the thick stainless steel cell used

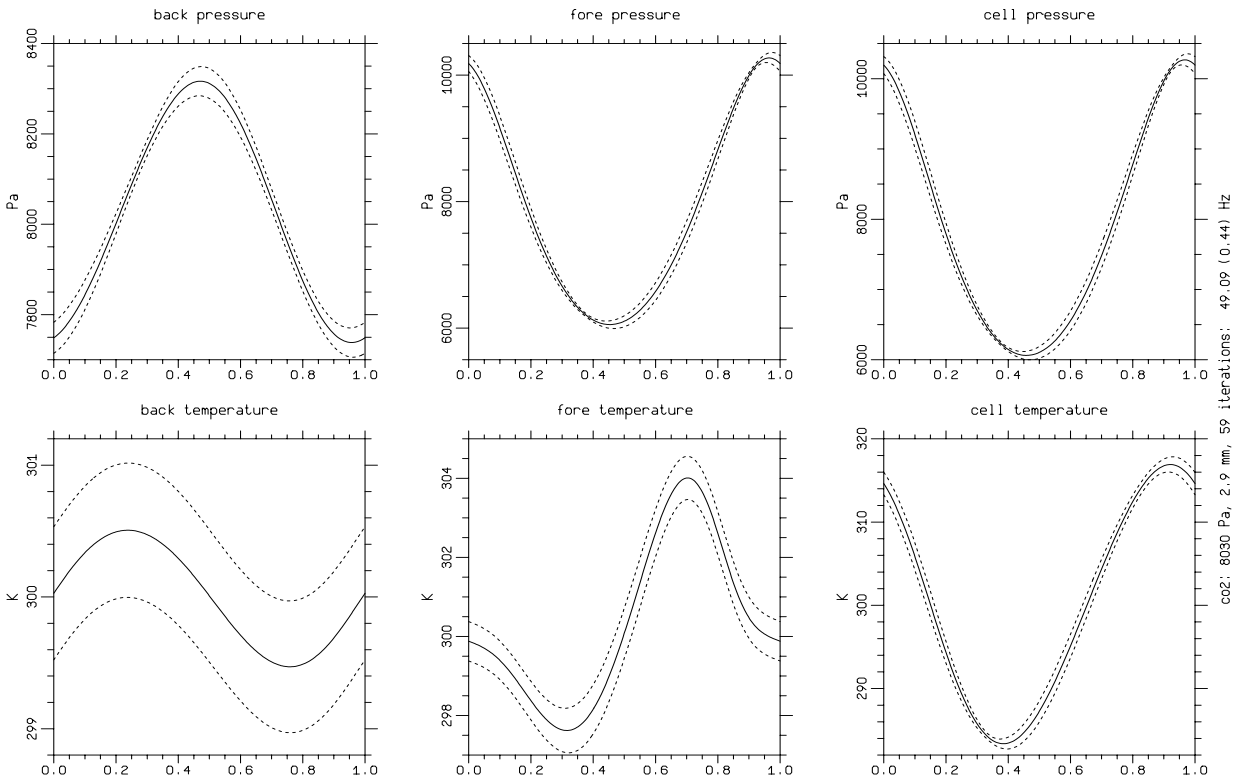

Figure 5.19: Model output for the PMIRR flight CO_2 PMU (8030 Pa, 2.9 mm). When this figure is compared with figure 5.3, it is evident that the cycle in the cell is more adiabatic for carbon dioxide, as the phase difference between temperature and pressure is smaller.

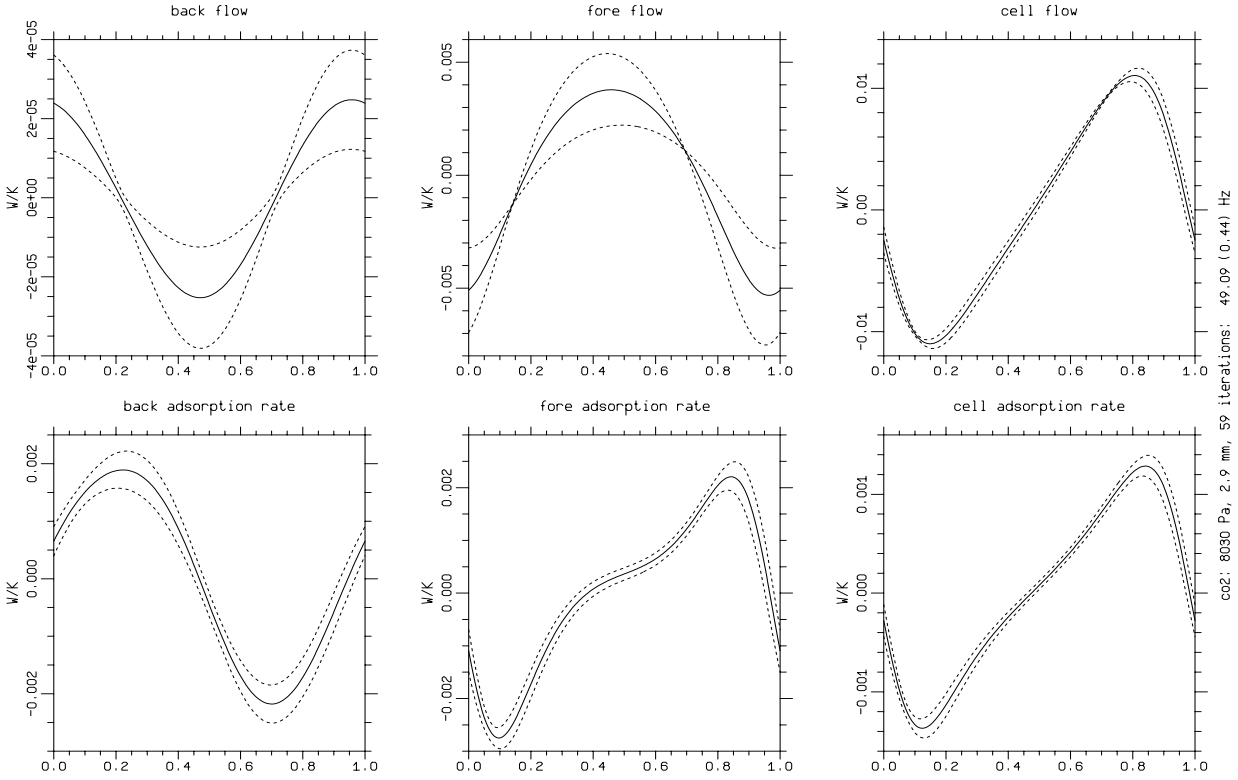

Figure 5.20: Model output for the PMIRR flight CO_2 PMU (8030 Pa, 2.9 mm). The terms used are defined in the caption for figure 5.4.

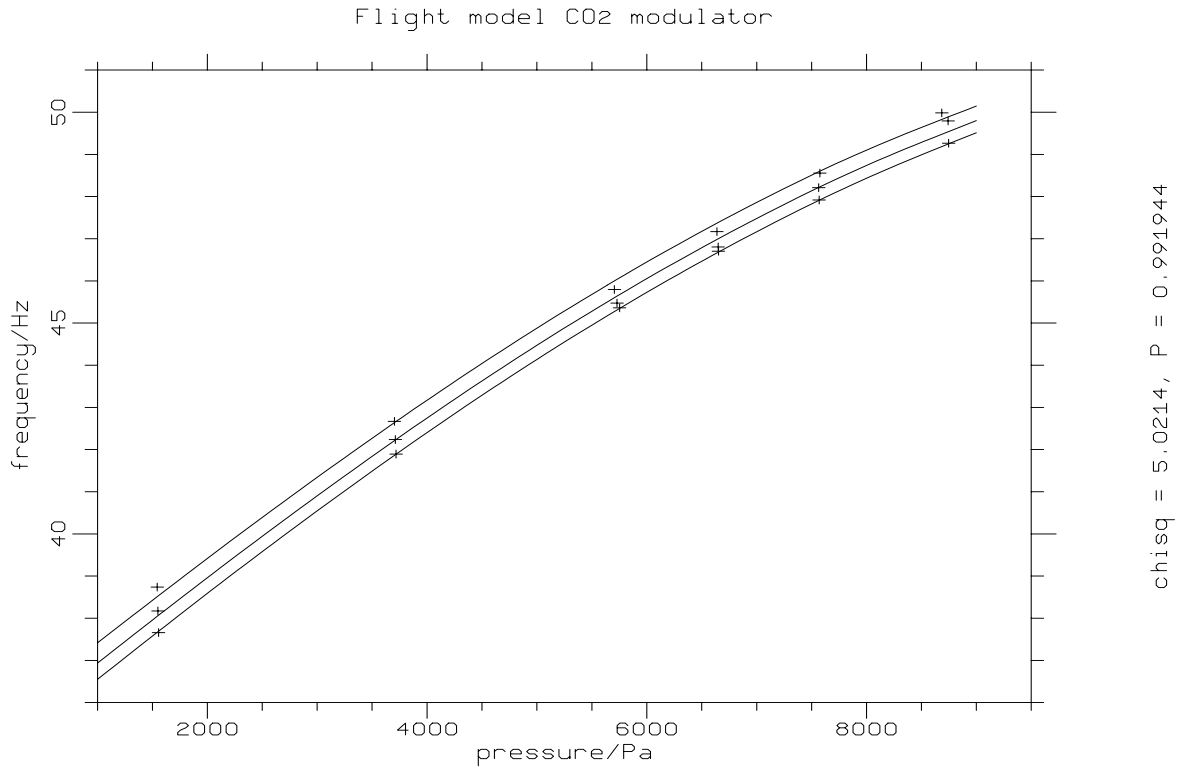

Figure 5.21: Measured and calculated frequencies for the PMIRR flight CO_2 PMU against filling pressure for piston amplitudes of 1.07, 2.15 and 3.0 mm. The measured frequencies shown here form a subset of the frequencies measured during filling, which have been presented in chapter 3, and are those used to optimize the modulator model spring and roughness parameters. The frequencies at all pressures increase with amplitude.

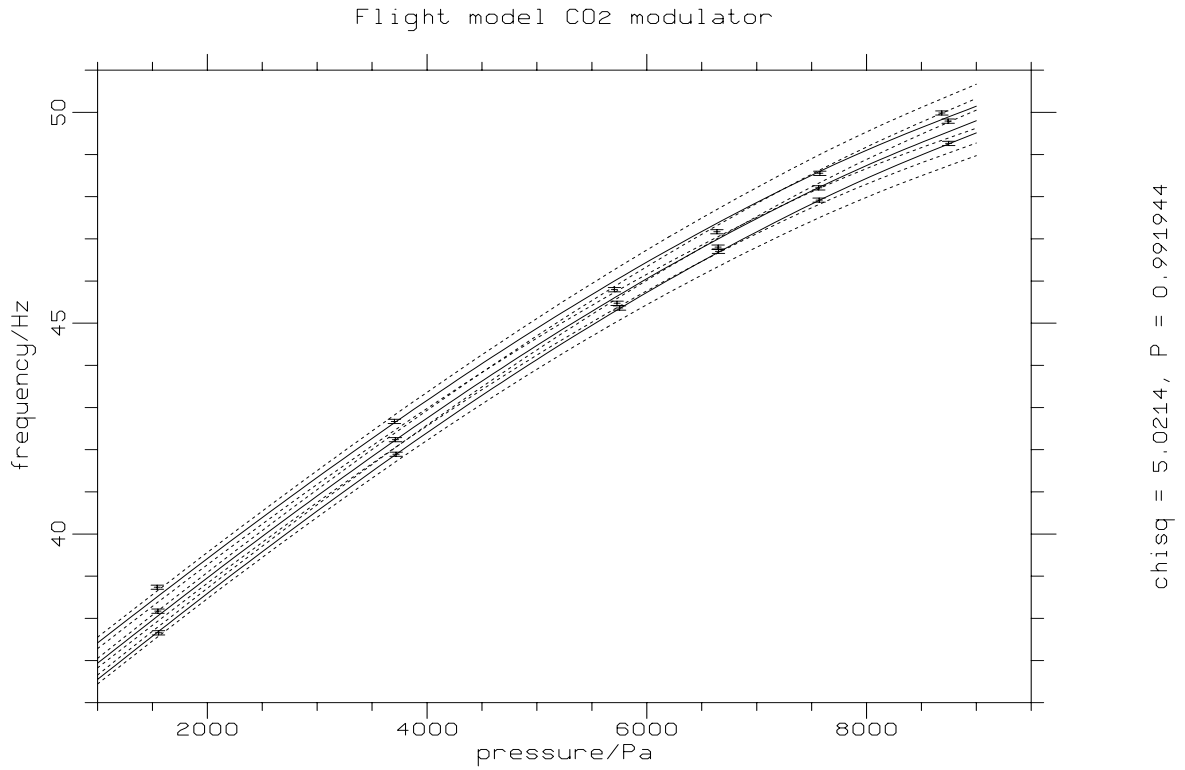

Figure 5.22: Measured and calculated frequencies for the PMIRR flight CO_2 PMU against filling pressure for piston amplitudes of 1.07, 2.15 and 3.0 mm. This figure also shows the ranges of the modelled frequencies and the uncertainty in the measurements.

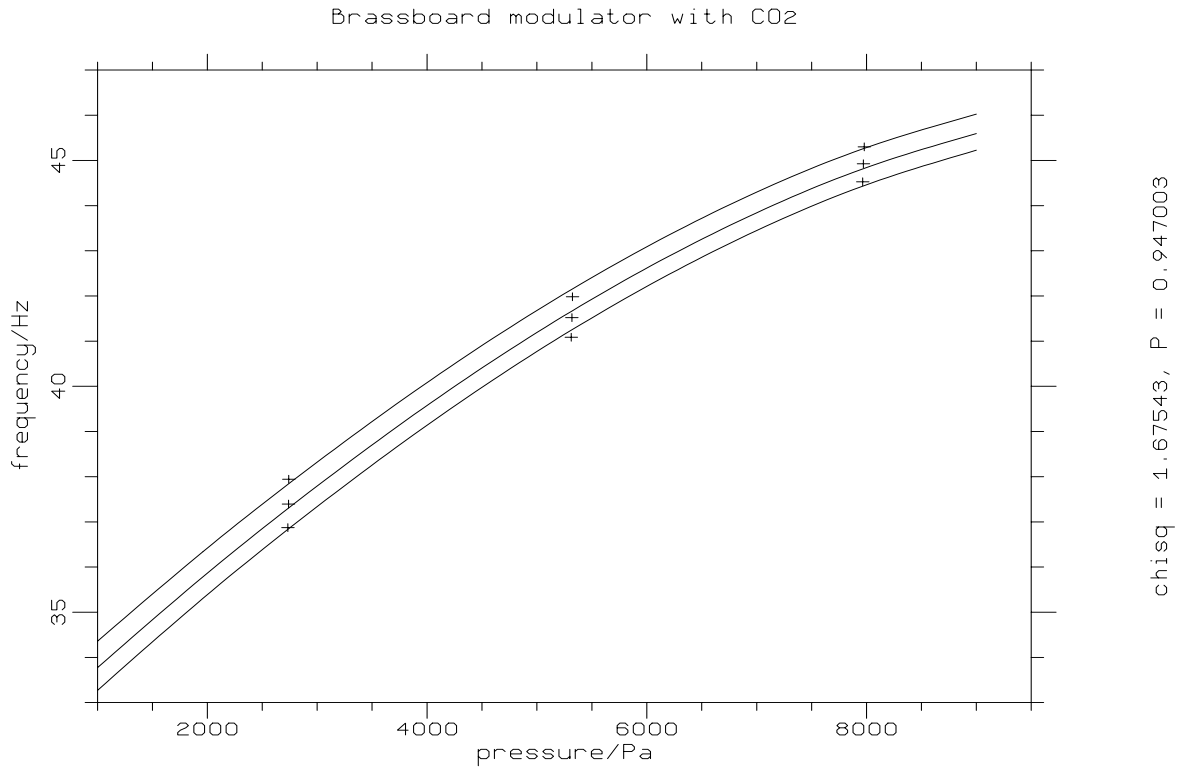

Figure 5.23: Measured and calculated frequencies of the brassboard PMU (at 300K body temperature) against CO_2 filling pressure, for piston amplitudes of 1.4, 2.1 and 2.7 mm. The frequencies at all pressures increase with amplitude.

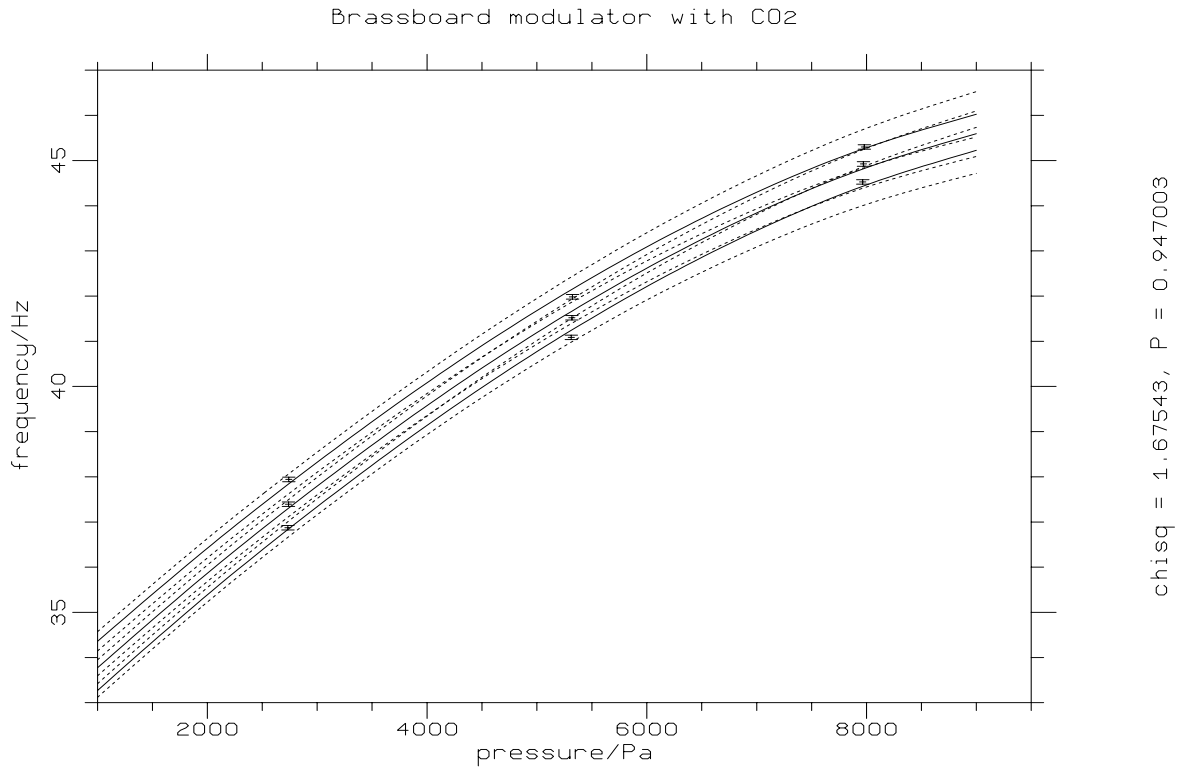

Figure 5.24: Measured and calculated frequencies of the brassboard PMU against CO_2 filling pressure, for piston amplitudes of 1.4, 2.1 and 2.7 mm. This figure also shows the ranges of the modelled frequencies and the uncertainty in the measurements.

for these experiments. In this connection it is interesting to note that the temperature of the flight model CO_2 modulator cell reached about 325K when the instrument was in a vacuum chamber, and the only mechanisms for removing heat from the cell were by radiation from its exterior surfaces and by conduction down the tube. The difference in the cell temperature attained in vacuum and laboratory environments implies that there is a large temperature gradient across the cell wall when in operation, and therefore that the de-adsorption rate from the cell wall is higher than that calculated when assuming that the inside of the cell is at the same temperature (300K) as its outside. Since the work done on the gas increases with mean pressure, this effect is more noticeable at high filling pressures.

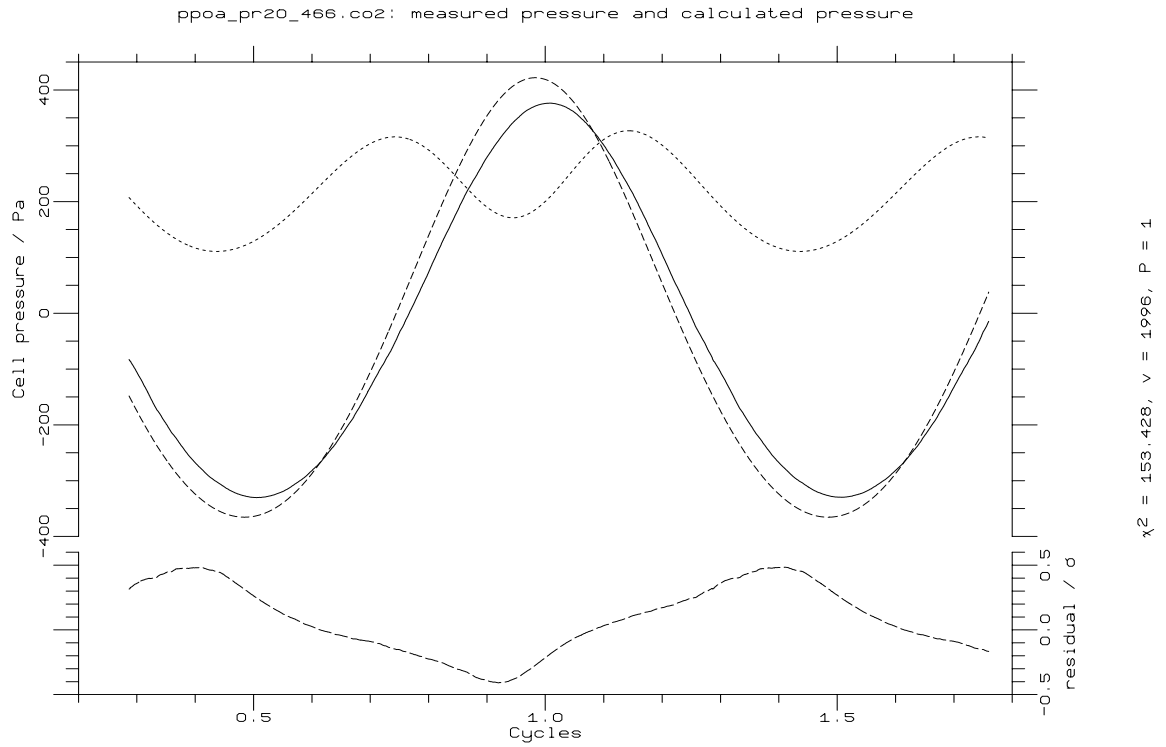

Figure 5.25: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 27.3mb mean pressure, 1.45mm piston amplitude. The calculated pressure cycle exceeds the measured cycle by 9% in amplitude, and leads it by 8.7° in phase.

The comparisons between the measured and modelled pressure cycles are shown in figures 5.25–5.33 for CO_2. The fit probability is high in all cases, but this arises largely from the large uncertainties in the modelled cycles (shown by the dotted curves). The variance contributions generated by the perturbation of each parameter are given in table 5.III, which is derived from the model output for the brassboard modulator, at 79.8 mb and 2.72 mm piston amplitude, near the flight model CO_2 operating conditions. The largest contribution to the frequency and pressure cycle uncertainties is from the piston gap tolerance of 10 μm, which significantly affects the gas flow rate between the back and fore volumes at high pressures. The wall temperature causes the second largest error in both, as it controls the gas desorption rate from the walls. The 'roughness', or area multiplier, is the third most important uncertainty; it controls the total amount of gas transfer by adsorption.

The pressure cycle amplitudes agree much better for CO_2 than for H_2O, and the calculated

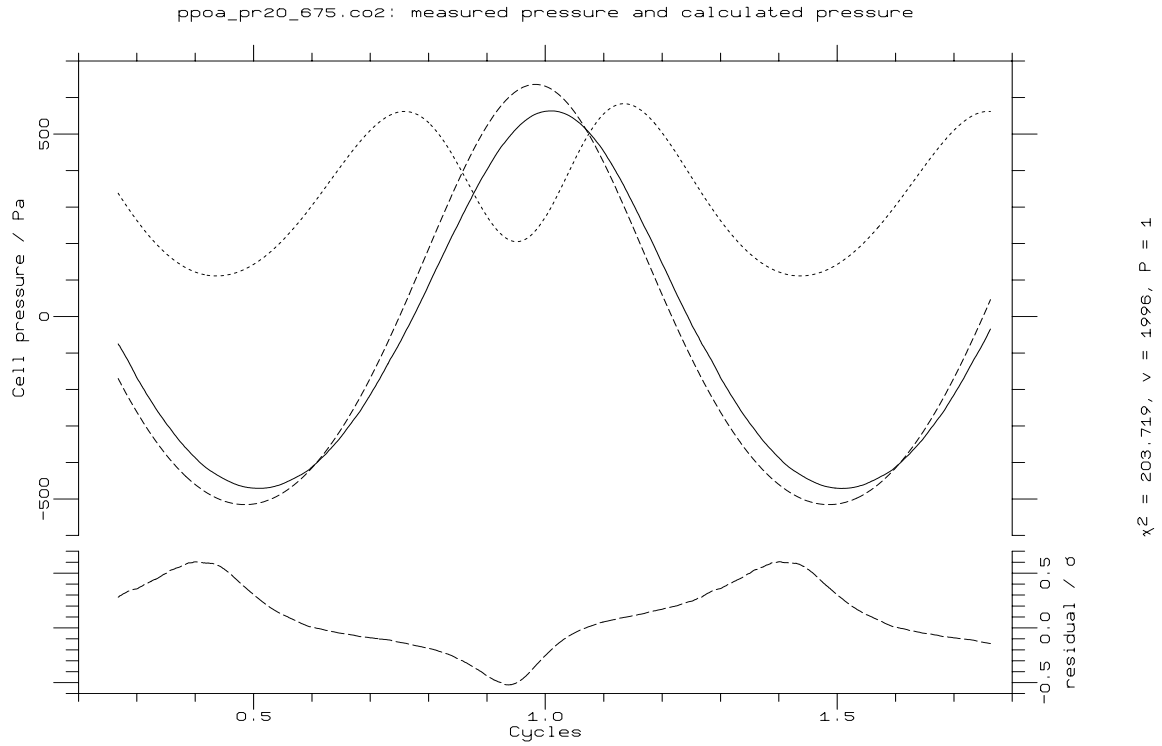

Figure 5.26: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 27.4mb mean pressure, 2.11mm piston amplitude. The calculated pressure cycle exceeds the measured cycle by 11% in amplitude, and leads it by 9.4° in phase.

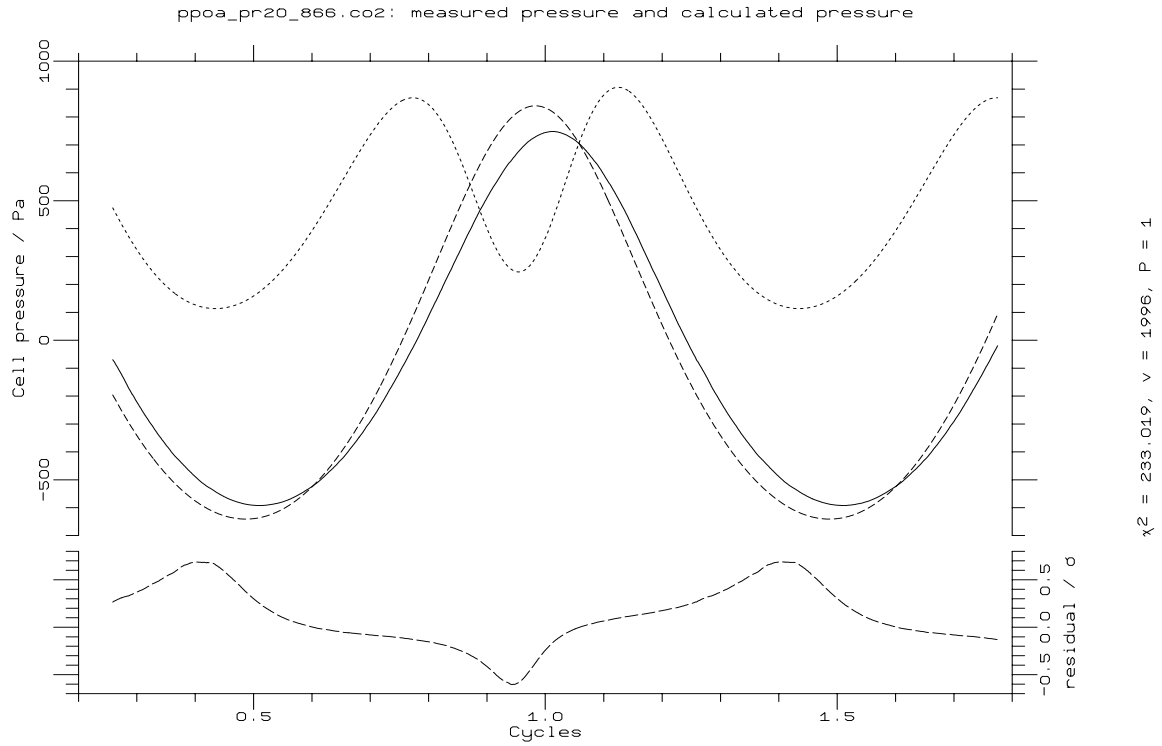

Figure 5.27: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 27.4mb mean pressure, 2.69mm piston amplitude. The calculated pressure cycle exceeds the measured cycle by 11% in amplitude, and leads it by 10.4° in phase.

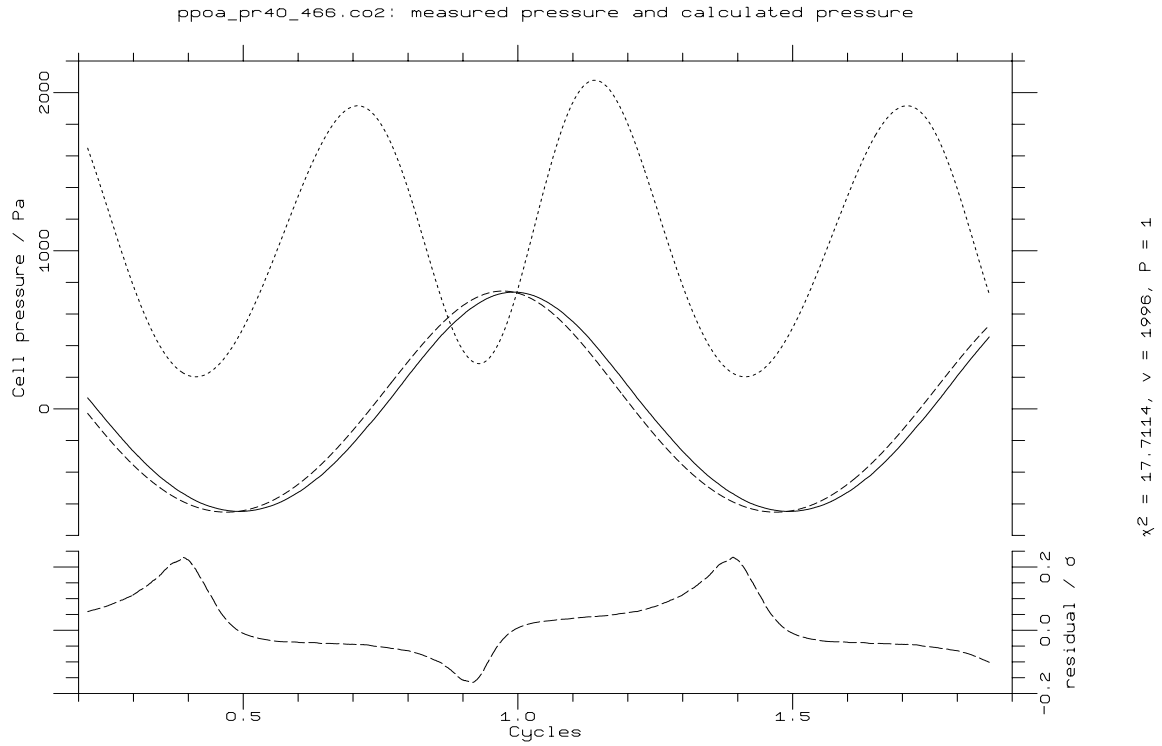

Figure 5.28: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 53.1mb mean pressure, 1.46mm piston amplitude. The calculated pressure cycle almost exactly matches the amplitude of the measured cycle, but leads it by 7.5° in phase.

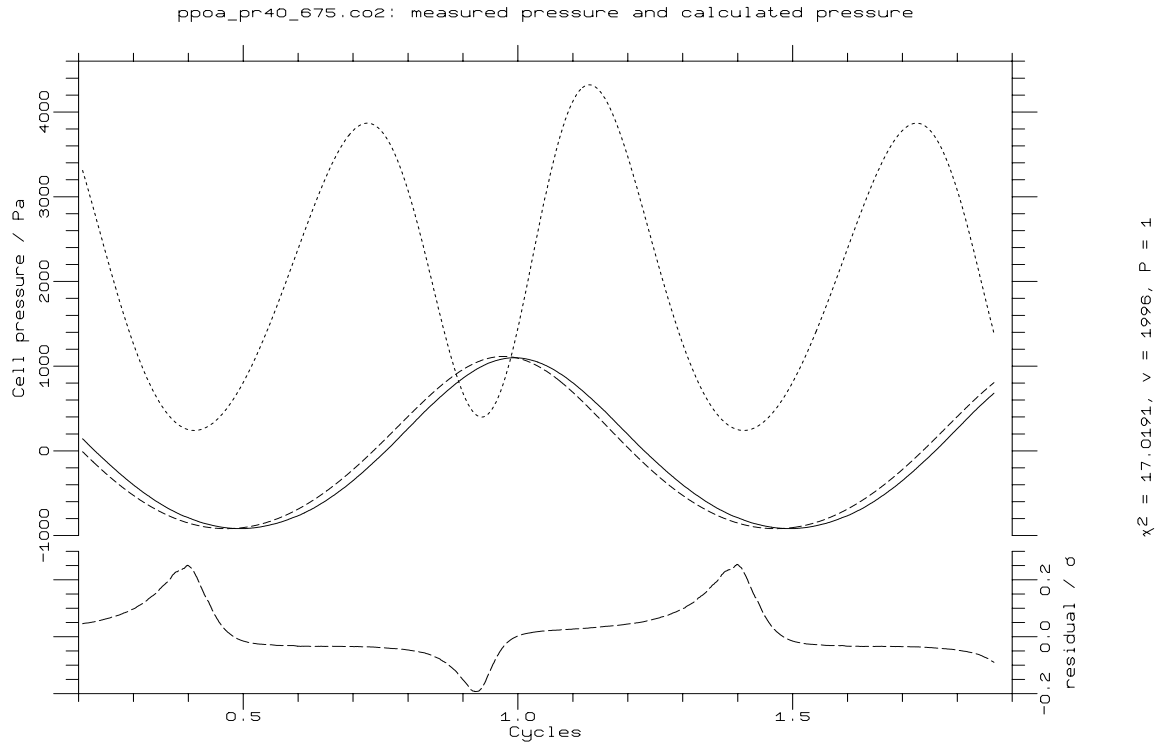

Figure 5.29: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 53.2mb mean pressure, 2.12mm piston amplitude. The calculated pressure cycle almost exactly matches the amplitude of the measured cycle, but leads it by 7.8° in phase.

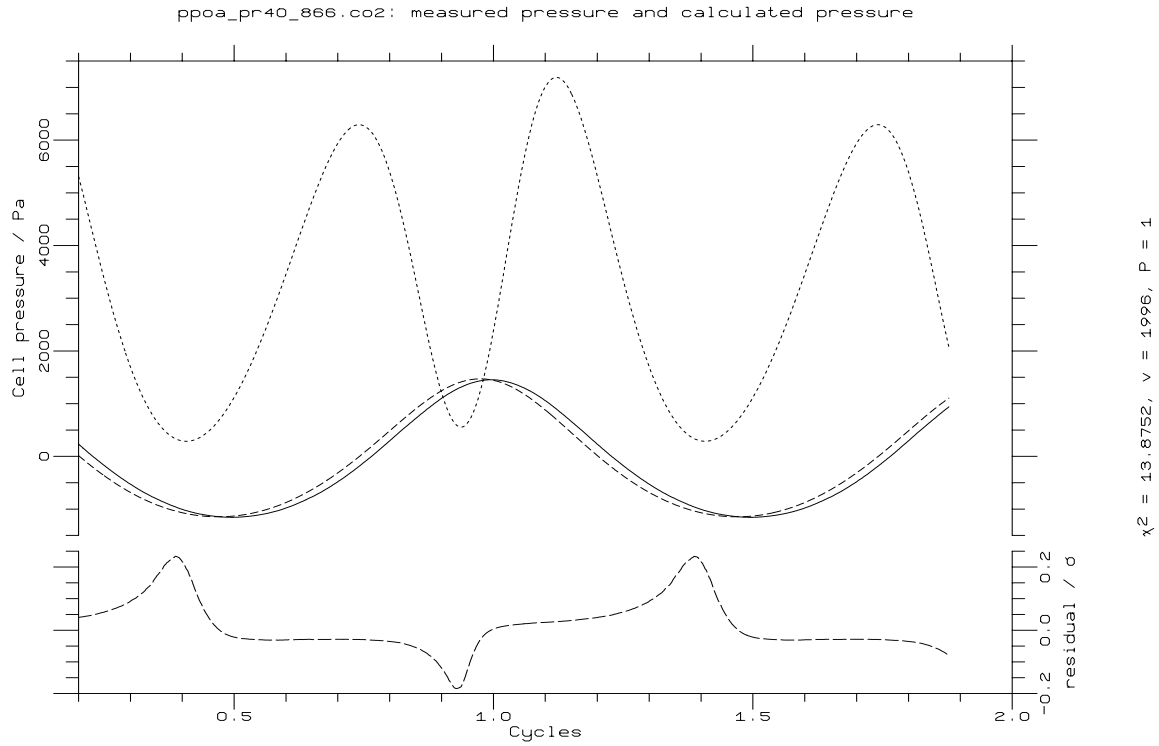

Figure 5.30: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 53.2mb mean pressure, 2.71mm piston amplitude. The calculated pressure cycle almost exactly matches the amplitude of the measured cycle, but leads it by 8.4° in phase.

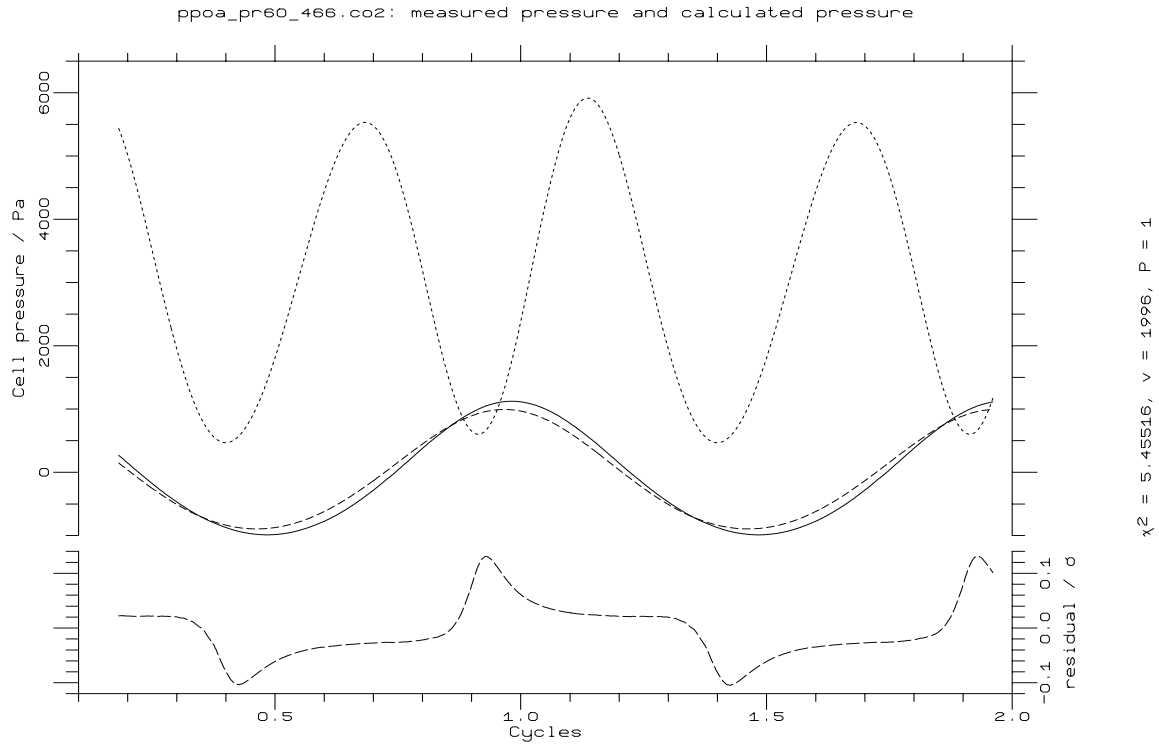

Figure 5.31: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 79.7mb mean pressure, 1.47mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by about 10% in amplitude, and lags it by 6° in phase.

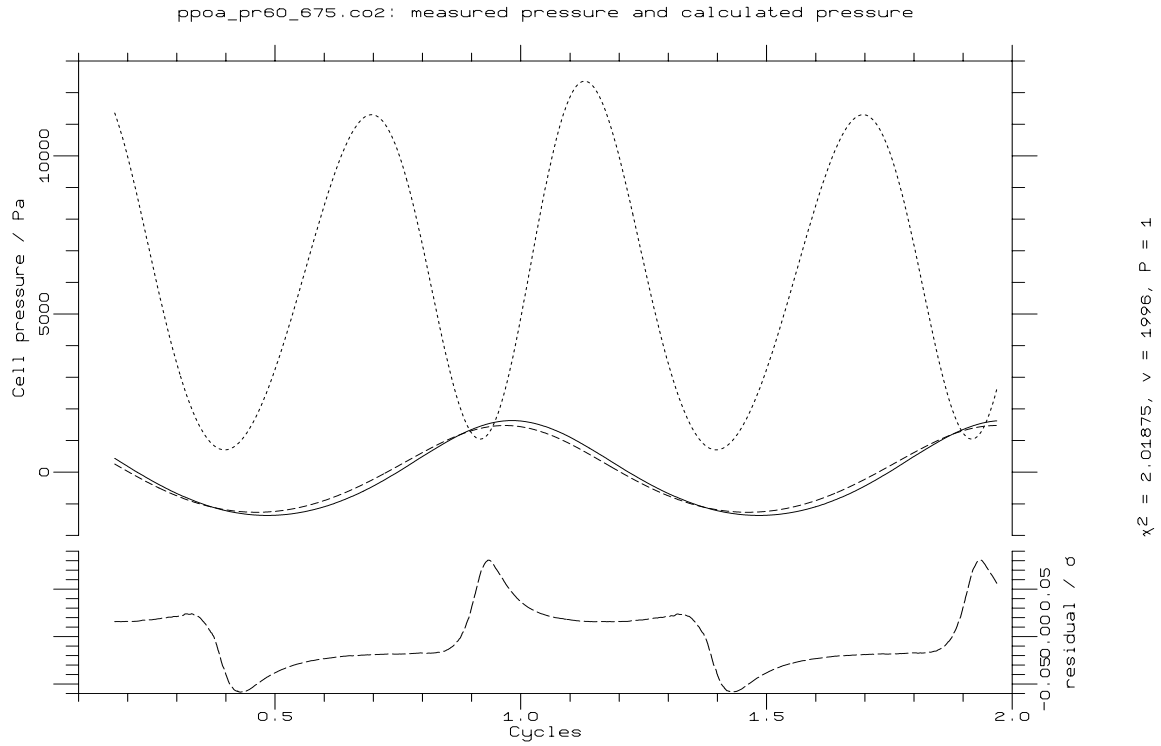

Figure 5.32: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 79.7mb mean pressure, 2.13mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by about 10% in amplitude, and lags it by 6.1° in phase.

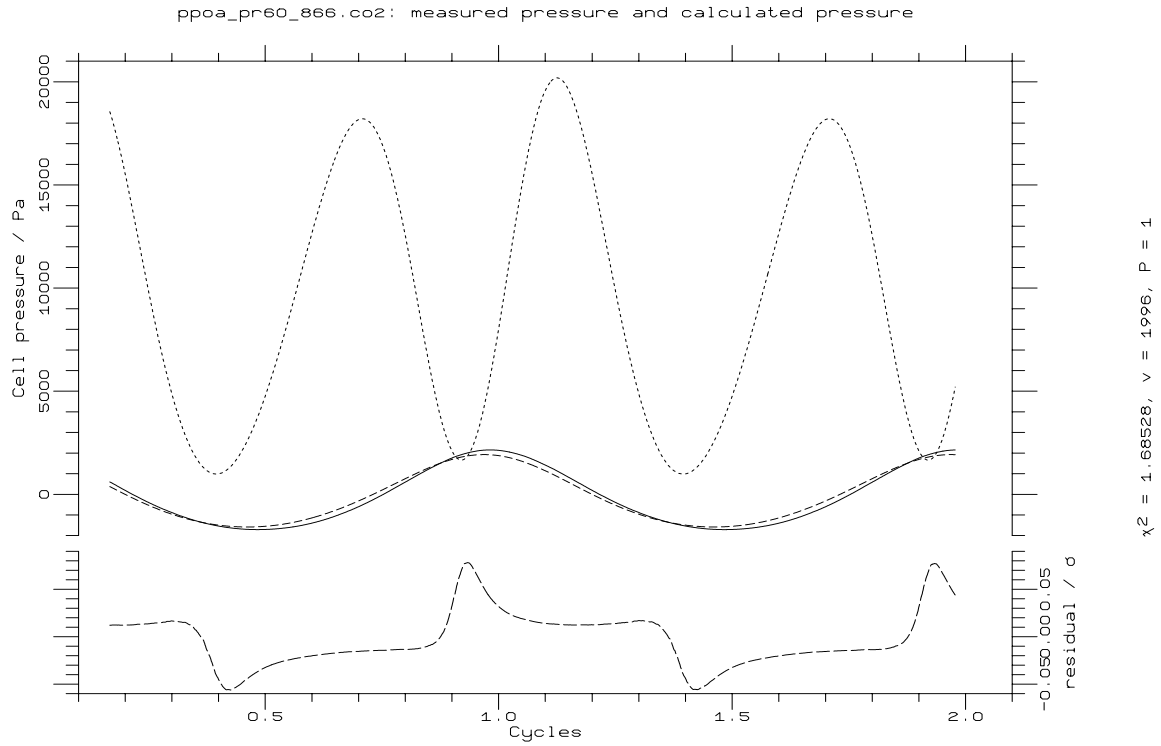

Figure 5.33: Measured (solid curve) and calculated (dashed curve) cell pressure cycles for CO_2 at 79.8mb mean pressure, 2.73mm piston amplitude. The measured pressure cycle exceeds the calculated cycle by about 10% in amplitude, and lags it by 5.9° in phase.

Varied	by	Δf/Hz	pressure cycle variance/Pa^2
Demanded amplitude	2e-05	0.009547084	10.5
Sieve pressure	10	0	0.0253
Wall temperature	0.5	0.1207766	16
Drive phase	0.035	0.1178518	0.693
Suspension spring strength	10	0.03865491	0.227
Coupling spring strength	2.95	0.04935189	0.285
Spring cubic term	0.000316	-0.05741208	0.334
Rest separation	2e-05	-0.01879932	2.66
Back length	0.01	-0.1594823	2.79
Piston mass	0.0001	-0.0713296	0.441
Piston radius	2.5e-05	0.05141245	2.33
Piston gap	1e-05	-0.3395986	125
Gap length	5e-05	0.003536566	1.88
Sieve tube radius	5e-05	-0.0001473563	0.0168
Sieve tube length	0.1	5.269693e-05	0.00602
Cell radius	5e-05	-0.05963365	8.52
Cell length	5e-05	-0.005704696	0.81
Cell tube radius	5e-05	-0.02355197	3.24
Cell tube length	0.0001	-0.00104507	0.165
Roughness	2.93	-0.04730565	6.36
BET c coefficient	10	0	0.0263

Table 5.III: Frequency changes and pressure cycle variances for CO_2. 'e' signifies a base 10 exponent. The numbers are taken directly from the cycle file produced by the model, so that the number of figures quoted is not an indicator of accuracy.

cycles lead the measured ones by a smaller amount. This is indicative of the smaller part that adsorption plays in the pressure modulation of CO_2, while still consistent with some reservoir action by the cell tube, which is longer for this modulator. At high pressures, the model overestimates the amount of adsorption, leading to a drop in the pressure cycle amplitude. This could be because of an imbalance between the roughness factors in the three modulator volumes, as was suggested for H_2O, but is also possibly because of the temperature difference across the cell walls, discussed above. The flow past the piston has a large uncertainty (figure 5.20), and this is also expected to contribute to the amplitude discrepancy.

5.4 Summary

The pressure modulator model described in this chapter is novel, in that it only addresses the steady-state PMU behaviour, and flexible, since in principle it can be run for the dimensions and operating conditions of most modulators without adaptation. It incorporates the first representation of adsorption, and gives good agreement with measured frequency characteristics for both carbon dioxide and water vapour.

Its predictions for the cell pressure cycles are not as satisfactory. For water vapour it seems as if the amount of adsorption is overestimated, since the calculated pressure cycle amplitude is significantly smaller than that measured, particularly at high pressures. For carbon dioxide, the magnitudes of the measured and calculated pressure cycles agree to

about 10%, but the uncertainty in the calculated cycle, arising from the uncertainties in each modulator parameter, is very large. For both gases the predicted cycle leads the measured cycle in phase by up to 12°.

Future improvements to the model could address the representation of the tube between the modulator and the cell, including the possibility of adsorption and condensation within it, which may cause the pressure cycle phase errors. The finite thermal conductance of the modulator and cell bodies should also be considered, especially at high pressures. The temperature structure within each volume, especially the cell, could be addressed in more detail, by considering more than one mode.

These theoretical improvements should be accompanied by better measurements of the modulator characteristics. In particular, it would be better if the spring constants were measured during assembly, and if frequency measurements were performed at several different temperatures. This is particularly important for water vapour, for which the de-adsorption rate varies strongly with wall temperature, and significantly alters the modulator behaviour. The various parameters required in the adsorption representation merit experimental investigation. The latent heats of adsorption and the sticking factors for the interaction between CO_2 (and other gases) and the various surface materials inside the modulator deserve closer definition, as do the BET c coefficients.

Finally, it may be possible in the future to use pressure cycle and other diagnostic measurements to assist in the fitting of unknown modulator parameters.

Chapter 6:

Tunable diode laser measurements of pressure modulated lineshapes

6.1 Objectives

Before the start of the present work, no direct measurement of pressure modulator transmission profile had ever been made, because spectrometers then available lacked the resolution required for such a measurement. Following a feasibility study by May *et al.* [1988] at JPL, using an Oxford-supplied SAMS-type modulator containing N_2O, I made high-resolution measurements of pressure modulator lineshapes using a tunable diode laser (TDL) spectrometer.

It was particularly hoped that the TDL might give information on the behaviour of water vapour in pressure modulators, thereby continuing to resolve the difficulties in interpretation of the Stratospheric and Mesospheric Sounder (SAMS) water vapour channel data [Mutlow 1984, Davis 1987, Munro 1991], which had led to a wide discussion about whether the technique could be used for highly polar ('sticky') molecules such as H_2O and NO_2 [Roscoe 1988]. This information is relevant to PMIRR, since one of its PMUs contains H_2O. Accordingly, greatest emphasis was placed on measurements of the pressure modulated transmission for water vapour. Measurements were also made using N_2O, and attempted for the '628' isotopic CO_2 band used in channel 1 of PMIRR.

These high-resolution measurements on the PMIRR brassboard and flight modulators formed part of the test and calibration of the PMIRR flight modulators. As part of the collaboration between JPL and Oxford on PMIRR, and by kind arrangement with Dr D J McCleese and Dr C R Webster, I was able to use the Atmospheric Laser Spectroscopy Group's Tunable Diode Laser Spectrometer at JPL for four weeks in February–March 1988 and another three weeks in December 1989.

This chapter includes a description of tunable diode lasers, the spectrometer system, the methods used to record pressure-modulator spectra, and the results. Conversion of the detector signal into transmission, and the comparisons between the measured spectra and those calculated from the modelled cell cycles, will be described in chapter 7.

6.2 Tunable Diode Laser Spectroscopy

6.2.1 General

To provide direct measurements of the pressure-modulator lineshape, a spectrometer of very high resolution (much smaller than the linewidth of a few GHz) must be used. Tunable diode laser spectrometers have adequate resolution for this. The laser linewidth

is of the order of a few tens of MHz, and the tuning range is typically about 30GHz (1 cm^{-1}); therefore a tunable diode laser spectrometer (TDLS) can be used for very high resolution spectroscopy over short spectral intervals. Most conventional spectrometers cannot achieve this resolution, although some Fourier Transform spectrometers approach it.

Normal spectrometers use a broadband source and tunable detector system: tunable diode laser spectroscopy uses a laser as a tunable monochromatic light source scanned across the spectral region of interest and viewed with a broadband detector system. In such a tunable-source spectrometer, ambient light reaching the detector will give rise to an unpredictable offset in its output. To circumvent this problem the laser beam is modulated either by chopping or by high-frequency scanning of spectral lines, which generates an AC signal as a result of the transmission variations across the spectral region.

6.2.2 Laser Diodes

Laser diodes are physically small (less than 1 mm^3) semiconductor devices which emit light when a current is passed through them. The active material is generally a lead salt, *e.g.*, lead tin telluride ($Pb_{1-x}Sn_xTe$), lead tin selenide ($Pb_{1-x}Sn_xSe$) and lead europium selenide telluride ($Pb_{1-x}Eu_xSe_yTe_{1-y}$). Different compositions and operating temperatures are used to provide a range of output wavelengths from 0.6 to 32 μm.

The diode is cut from the salt, and its ends are polished to provide the laser cavity. Because of the small size and the geometry (the total emitting area is about 10^{-4} cm^2) the light diverges typically in an $f/1$ cone, and a collimating lens or mirror is used to produce a parallel beam.

The operating temperature for mid-IR diode lasers is typically 50–80K. The temperature is controlled by placing the diode on a thermostatted cold finger, which is cooled by either a closed-cycle compressed-helium refrigerator or a liquid nitrogen dewar, depending on operating temperature. The current through the diode is typically in the range of 200–1200 mA.

6.2.3 Tuning

Over a small range (usually about 20–30 cm^{-1} discontinuously and about 3 cm^{-1} continuously) the frequency of a particular device can be varied by changing the bulk temperature or the bias current, the magnetic field or the pressure [Mooradian 1976]. Generally temperature and current are used simultaneously. The temperature of the laser alters the physical dimensions of the laser cavity and the state of the population inversions between the various states associated with the laser action, thereby changing the output frequency. Most diodes show a significant change in tuning and modal characteristics when warmed up to room temperature and re-cooled, largely from geometrical deformation.

The current tuning is a subtle effect. Whilst the current causes a small amount of heating due to non-radiative losses, with a consequent change in temperature, this is to a large extent compensated by the temperature controller, and reduced by the thermal mass of the diode. Furthermore, the diode output frequency would not be instantly repeatable for a given current, because of the time taken to reach a thermal steady state. In fact, the current causes small changes in both the bandgap energy and the electron density. The bandgap energy controls the frequency directly, while the electron density, by affecting the

refractive index, controls the effective length of the cavity, and thus adjusts the frequency of the gain peak. The cavity tuning rate is typically a third of the bandgap tuning rate. Both these process are very rapid, so the current tuning is instantaneous.

As the refractive index is also dependent on the temperature, it is possible to alter the tuning behaviour by simultaneously adjusting the temperature and the current. This effect was used to optimize the diode's characteristics over the spectral ranges of interest.

6.2.4 Mode output characteristics

The laser output consists of several longitudinal and transverse modes. Each mode has a *threshold current* below which no emission occurs, and near which the output power of the mode rises linearly with current. As the current is increased further, the power becomes more constant, and then falls abruptly as the mode dies. The peak power in a single mode is typically 0.1–1mW which is large compared with the throughput of conventional spectrometers of similar resolution.

Typically an individual mode can be tuned continuously over a range of 1–3 cm^{-1}; as different modes often overlap in frequency, particularly at high currents, a total tuning range of some hundreds of wavenumbers can sometimes be achieved. A monochromator is used to separate one mode from the rest, and the current is kept low to reduce the number of modes.

Because of the small size of the diode, parts of each mode, projected from the diode in different directions, appear as different parts of the collimated beam. This results in a non-uniform power/frequency relationship across the beam, so that two or more optical paths separated from the same mode often have different power envelope curves. This feature was observed in my measurements.

6.2.5 Laser linewidth

The spectral width of the laser output varies strongly with temperature and current. It arises both from quantum-mechanical and collisional line broadening processes within the diode [Mooradian 1976] and from noise in the current and temperature, which produce corresponding uncertainties in the output frequency. Typically the peak current noise is 100 μA, and the thermostat precision is a few mK.

When the linewidth is not negligible, the measured spectra must be deconvolved into instrumental and test-cell parts. The laser linewidth is predominantly due to random noise and can therefore be determined from the difference between the observed and calculated Gaussian halfwidths of lines in a long (*e.g.*, 15 cm) low-pressure gas cell, in which the gas lines are Doppler-broadened and the absorption is proportional to the gas amount. The linewidth measured for both sets of measurements at JPL was about 30 MHz; this is typical for an infrared diode laser cooled with a closed-cycle cooler. For the pressure modulator measurements this instrumental linewidth is negligible in comparison with the gas-cell linewidth (1.5 GHz) and for the purposes of these experiments the laser could be treated as monochromatic.

6.2.6 Power noise sources and characteristics

The optical beam is very small at the diode, and highly focussed by the collimating lens, so that there is a transfer of the displacer-induced vibration to the optical intensity through motion of the cold finger transferred from the closed-cycle cooler displacer. The piston motion also vibrates the optics and causes alignment changes. Random fluctuations in the diode itself, such as thermal and shot noise, are also transferred to the output. All these effects can be seen in the detector signal. The output noise spectrum of the diode also varies with the junction temperature and current.

The amplitude varied by approximately 3% over intervals of a second, although over a shorter time interval the variation was much smaller. Over the time taken to measure a spectrum (typically about 30 seconds for my experiments), the variation in amplitude limited the wideband SNR to about 30.

At frequencies below 1MHz the diode noise is dominated by a $1/f$ characteristic. When possible, the effects of this characteristic are reduced by various high-frequency modulation techniques. In my experiments, the modulation was provided by a chopper, at about 500Hz, and by the pressure modulator, operating at about 50 Hz.

6.2.7 Operating frequency identification

Since the laser diode output frequency cannot be set or measured directly, it is necessary to identify its operating region. The centre of the spectral region is loosely defined (to about 2 cm^{-1}) by the bandpass of the monochromator. The precise operating frequency is determined by measuring the transmission spectrum of a low-pressure gas sample and some étalon fringes or a reference gas spectrum measured under the same conditions. The peaks in the latter should have constant and known frequency separations, so that the HITRAN line list can then be searched until lines matching the measured relative intensity pattern and separations are found. Gases with irregular line spacings and intensities are most useful for this purpose. The frequency *versus* current relationship of the the diode may also be derived from this measurement. (The breadth of the lines in the pressure modulator makes them unsuitable for calculating the frequency *versus* current relationship in this way.)

6.2.8 Tunable Diode Laser Spectrometer

The spectrometer made available to me at JPL was a Spectra-Physics commercial tunable diode laser spectrometer with a modular construction. Its layout is shown in plates III and IV. All the optical components were mounted on a flat, stable optical table. The diode itself was mounted in an evacuated cold head, and connected to the closed-cycle cooler by a thermostatted cold finger. Light from the laser passed through a window in the cold head, was collimated by a germanium condensing lens, and then chopped by a disc chopper with a variable-speed drive. A plane mirror directed the beam into the entrance slit of the monochromator, selected one mode of the many emitted by the diode. Location holes in the baseplate allowed the insertion of 45° mirrors to replace the diode laser beam with that of a Helium-Neon laser, and to short out the grating, for alignment purposes.

On leaving the exit slit of the monochromator the beam was split into two channels. One passed through the pressure modulator cell to the main detector. The other beam was used as a frequency reference. In February 1988 a low-drift confocal étalon with a free spectral range of 286.9($\pm$0.3) MHz, was used. This was replaced by a low-pressure carbon disulphide cell for December 1989, because high accuracy (to 2 MHz) measurements of

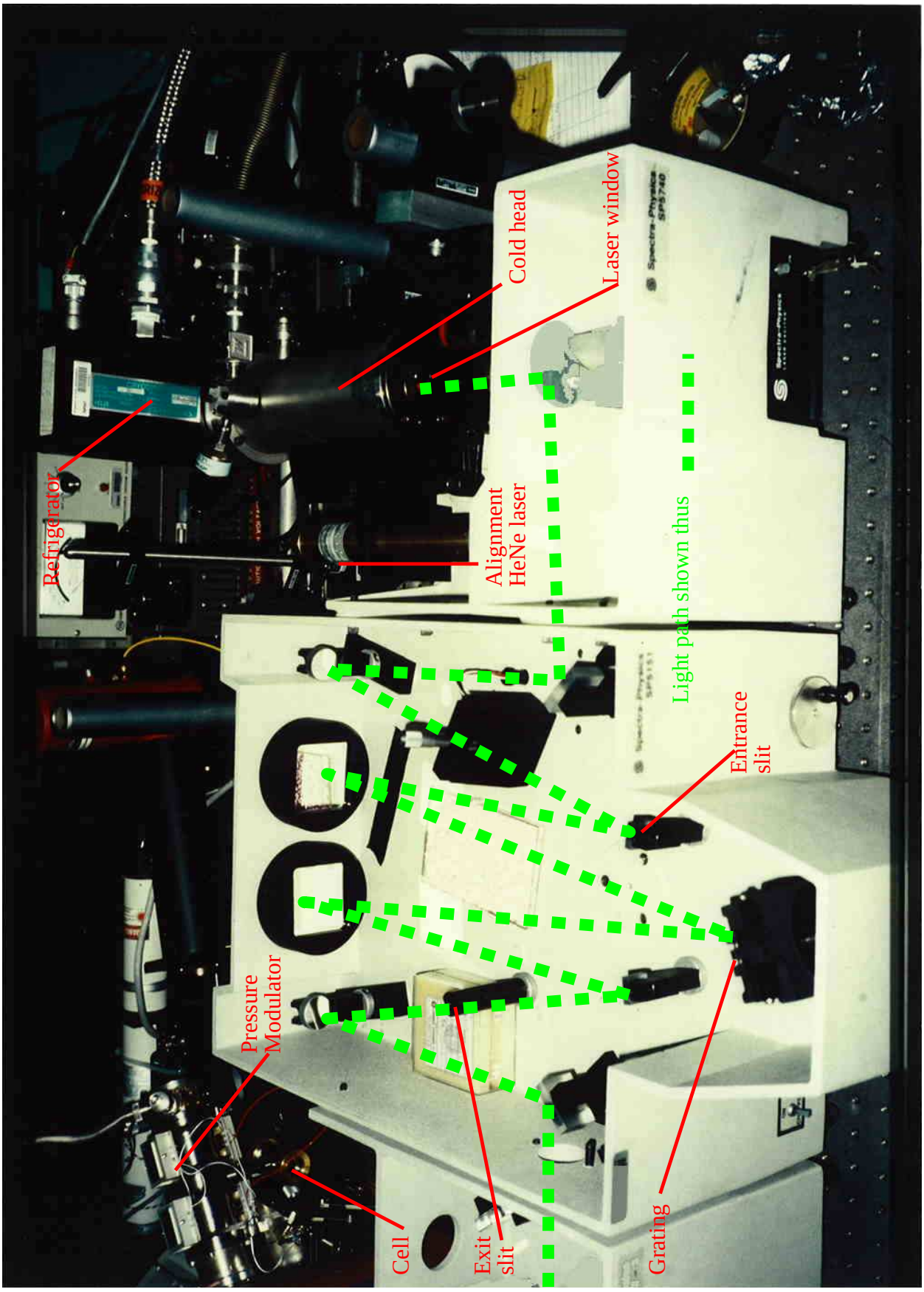

Plate II: Tunable Diode Laser Spectrometer (cold head and monochromator)

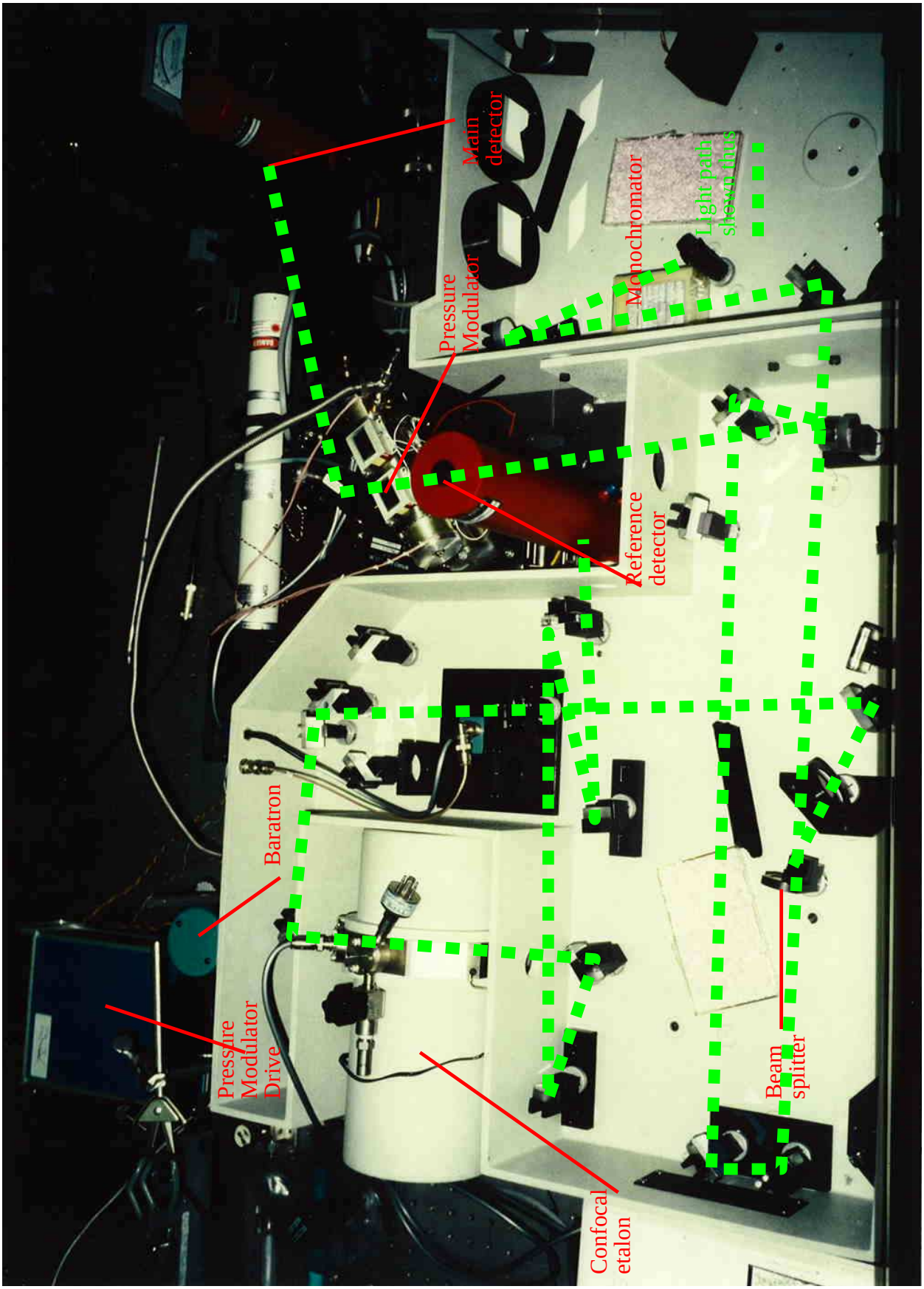

Plate III: Tunable Diode Laser Spectrometer (optics, étalon, PMU and detectors)

the line positions of CS_2 in the spectral bandpass of channel 4 of PMIRR had recently been published [Wells *et al.* 1988]. The 15 cm cell containing the CS_2 leaked slightly, resulting in the presence of a small amount of water vapour. This identified the low-pressure line centres of the two strongest water lines.

The detectors were liquid-nitrogen-cooled HgCdTe (MCT) photoconductive detectors. The detector preamplifiers were set up with passbands from 100 Hz to 10 kHz to reduce the noise bandwidth without affecting the signal. The preamplifiers produced an output voltage nearly linear with respect to the input intensity; when used with a chopper and AC-coupled amplifier the output was proportional to the incident intensity on the detector element.

6.3 Experimental details

6.3.1 Configurations

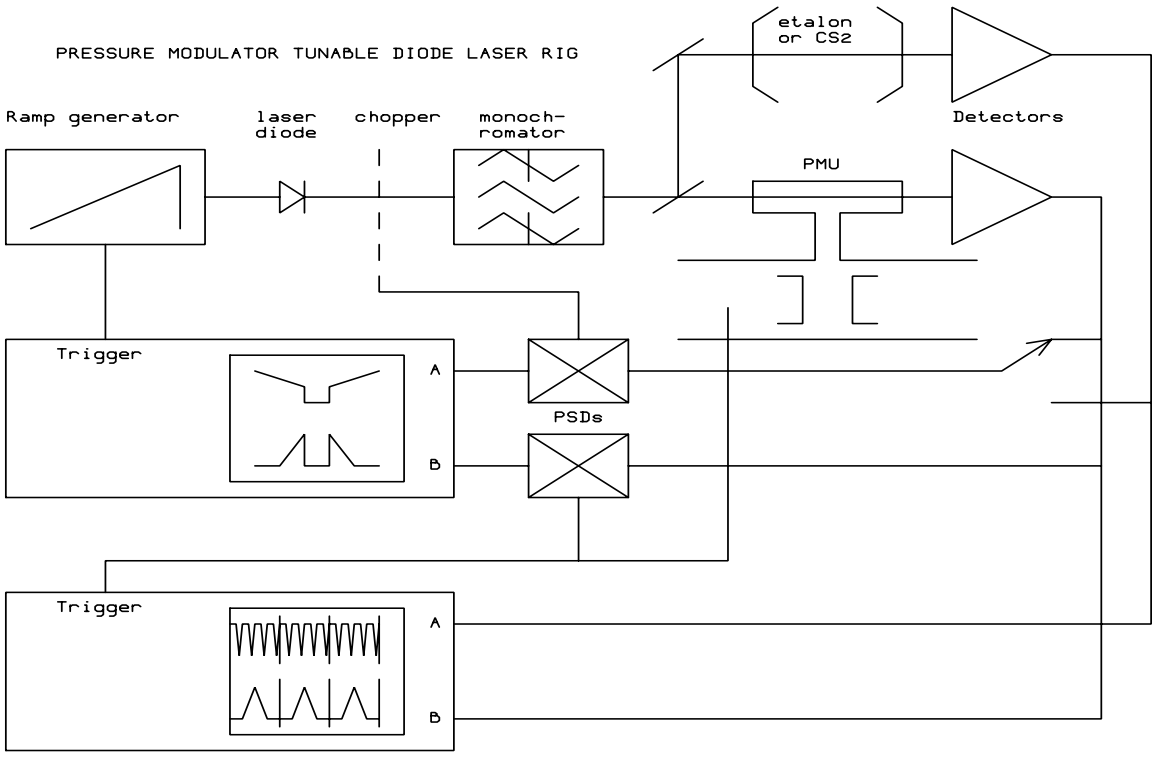

Figure 6.1: Experimental configurations used for the TDL experiments. The light path was from the diode, through the chopper and monochromator, to the beam-splitter. Part of the beam continued through the pressure modulator cell to one detector, and the rest through a reference cell or étalon to a second detector.

The two configurations of the equipment are shown in figure 6.1. For the normal (PSD-resolved) spectral measurements, the main detector signal was demodulated at the chopper frequency to give the mean transmission signal, and simultaneously demodulated at the modulator frequency for the modulated transmission signal. The modulated transmission, and either the mean transmission or the etalon signal, were recorded by the signal averager.

The sweep integration configuration was also used to give instantaneous diagnostics of

the pressure modulator cell. Without the chopper running, the laser was swept across the line 8 times per modulator cycle and each resulting spectrum recorded. For this type of measurement, the drive circuit produced two sets of pulses, one at the modulator frequency, adjustable in phase relative to the piston, and one eight times faster. The fast set of pulses started the diode current sweep and the slow ones triggered the signal averager, so that eight sweeps appear on one trace of the averager.

6.3.2 PMUs and cells

In February 1988, I made measurements of N_2O and H_2O using the brassboard PMU described in chapter 3, which was moderately well-balanced (8% imbalance). The maximum attainable amplitude was only about 2.3–2.4 mm peak-to-peak for each piston. In December 1989 I performed similar measurements with the PMIRR flight PMUs, which had then been filled and crimped (described in chapter 3), and again with the brassboard PMU to obtain measurements at different pressures. The brassboard modulator had been rebuilt and was well-balanced (4% imbalance, increasing the maximum attainable amplitude to 2.7 mm.

The flight PMUs both had cells with coated Ge windows and Platinum Resistance Thermometers bonded to the side. The cells were permanently attached to the PMU and sealed with a gold 'O'-ring. The brassboard PMU had two removable stainless steel cells, made to the same internal dimensions as the PMIRR cells, and sealed to the PMU using a Viton 'O'-ring. The length of the gas paths between the windows were 37.895 ($\pm$0.008) mm for the CO_2 cell, also used for N_2O measurements, and 101 ($\pm$0.07) mm for the H_2O cell. In order to have a free choice of wavelength region the brassboard cells had removable CaF_2 windows, 1 inch in diameter and 3 mm thick, transparent at 7 μm and in the visible region (for spectrometer alignment using a He/Ne laser). For the second set of experiments I added a pressure-sensor and a semiconductor temperature sensor to the cells.

6.3.3 Full scale measurement

The output power of the laser varies with frequency, as discussed above. Knowledge of the variation is necessary to convert the measured detector signal into transmission. It is preferable to measure this curve by recording the throughput of the gas cell when empty; this method compensates for the cell window opacities. However, for the flight modulators this could not be done, as the cells were sealed. Moving the cells out of the optical path was not an adequate substitute, because the light path would have been different. For this reason I fitted the power curve (usually called the 'baseline') from the measured wideband spectrum, as described in chapter 7. The same method was adopted for the brassboard modulator.

The form of the pressure-modulated spectrum baseline is the same as the wideband spectrum, as it is derived from the same source, but its absolute scale is uncertain, because the gains of the signal chains were frequency-dependent. I attempted to achieve an absolute measurement of the pressure-modulated transmission by replacing the modulator with a slow-running chopper (at about 47 Hz), to represent a modulator chopping between black (at ambient temperature) and transparent. By running this chopper at about the same frequency as the modulator, the gain *versus* frequency characteristics of the detectors, preamplifiers and PSDs remained the same as for the pressure modulator measurements. In practice, however, the chopper frequency was at the lower limit of its drive circuit's lock range, and consequently poorly controlled. Although a curve was obtained similar

in form to the baseline expected from observation of the wideband transmission signal, the measurement was of no practical use because of the resulting phase noise from the chopper reference signal, which confused the PSD. As a result the measured curve was not used except as a preliminary check on the approximate relative magnitude of the pressure-modulated signal.

6.4 Nitrous Oxide

The first measurements were made using N_2O, because of its strong lines, well-understood spectroscopy and physical stability. In particular, it was hoped to use N_2O for intercomparison with H_2O by investigating lines of similar ground state energies and line strengths. Four lines were measured simultaneously: a doublet at 1278.5653/1278.5662 cm^{-1}, another doublet at 1278.7432/1278.7466 cm^{-1}, a strong singlet at 1278.9637 cm^{-1}, and a weaker singlet at 1279.1099 cm^{-1}. I used pressures of 12 and 20 mb, and amplitudes of 0.5, 1.2, 1.75 and 2.3 mm. Insufficient resolution prevented fast scans being done of all four lines; only the strongest was measured in this way.

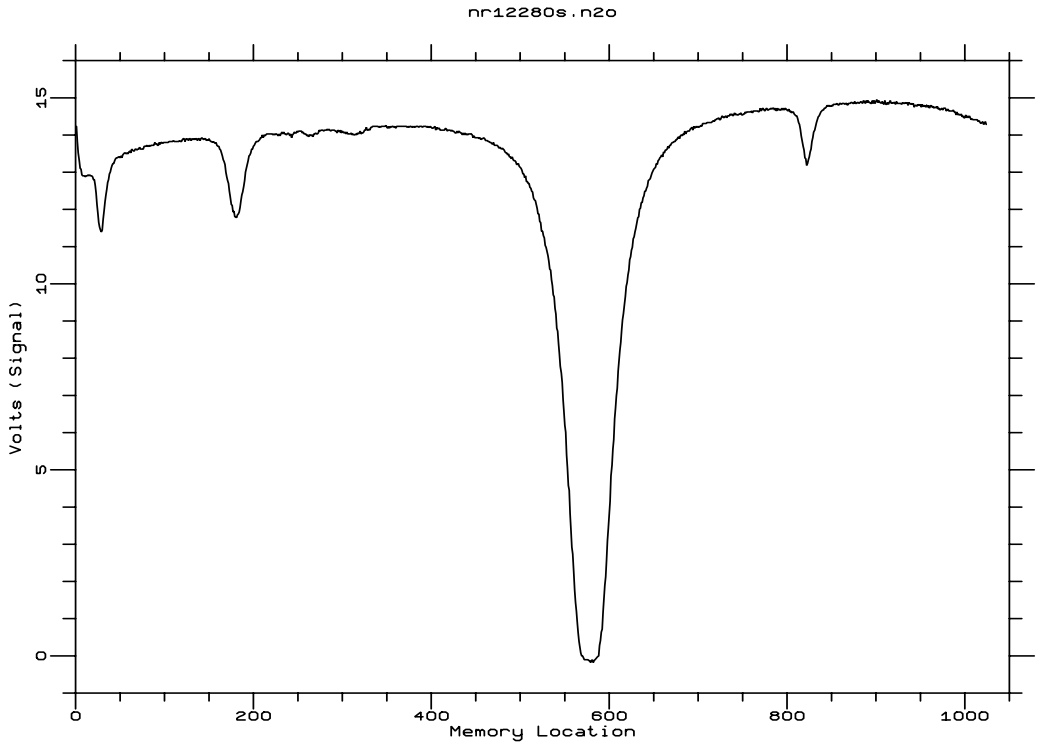

Figure 6.2: Wideband spectrum measured for the brassboard modulator, near 1279 cm^{-1}. The modulator contained N_2O at 29.7 mb, in a 4 cm cell.

The slow scans show an asymmetry in the peak heights (figure 6.3) the lower-frequency peak being larger. This was independent of the direction and magnitude of the current ramp, the modulator frequency, and the time constant on the phase-sensitive detector used for the pressure modulator, and thereby not an artifact of the measurement system, but a spectroscopic feature. It is discussed more in chapter 7.

6.5 Water vapour

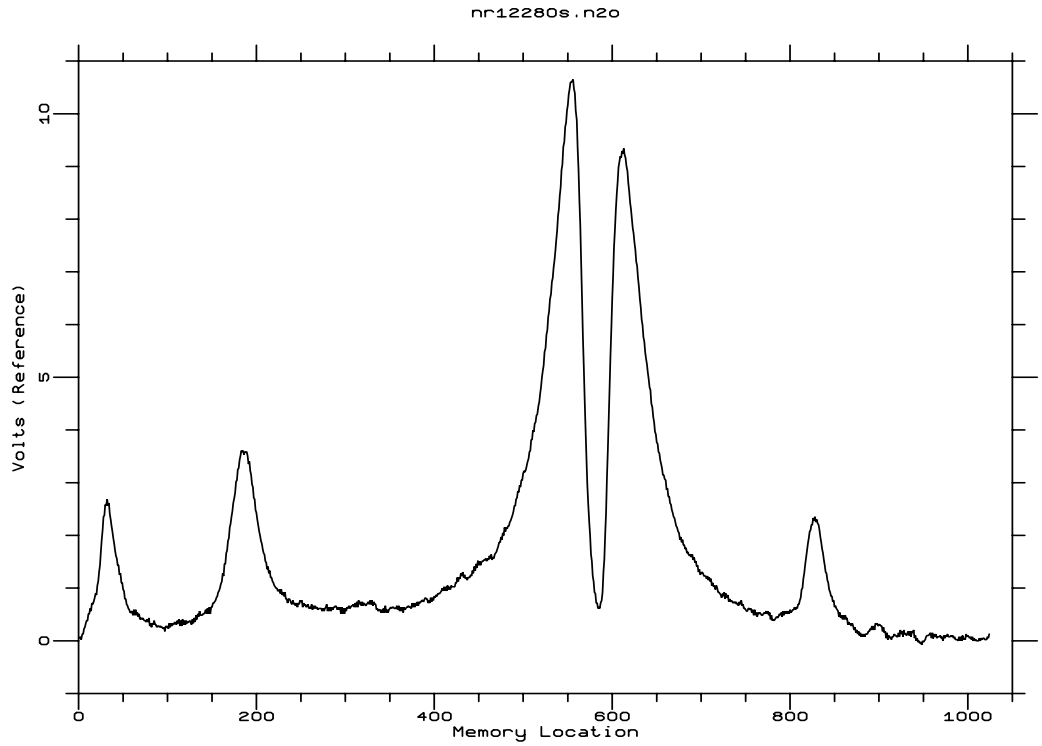

Figure 6.3: Pressure-modulator spectrum of N_2O near 1279 cm^{-1}, for a piston amplitude of 2.3 mm and a mean pressure of 29.7 mb. The asymmetry in the central profile is not an artifact of the measurement process, but a spectroscopic feature, possibly caused by pressure shifting.

6.5.1 Results – February 1988

A diode had been selected to operate in the spectral region of the PMIRR H_2O channel. However, it had not previously been tested, and when first switched on was found to have a very small tunable range between mode hops, rendering it almost unusable. Therefore the diode that had been used for N_2O measurements was retained. The strongest water line measured was at 1271.79 cm^{-1}, with a strength of 1.6×10^{-21} mol cm^{-1}, about two orders of magnitude weaker than the line originally chosen. This also reduced the absorption by water vapour in the atmosphere, which for the strong lines in the PMIRR channel 4 spectral bandpass absorbed most of the light near the centres of the lines.

Mean pressures of 6.5, 13 and 19.5mb, and amplitudes of 0.6, 1.2, 1.8 and 2.25 mm were used for these measurements. The profiles recorded using phase-sensitive detection were very noisy, with inconclusive shapes even after averaging for more than an hour, as each successive scan was different in magnitude and form. This occurred because of large variations in absorption by atmospheric water vapour. The best result is shown in figure 6.4. A typical sweep-integrated scan is shown in figure 6.5. These measurements were degraded in the same way although the effect was less obvious, as the time for each sweep was smaller than the time over which the atmospheric path varied. The atmospheric water vapour was responsible for producing a baseline which was not only highly-curved (see figure 6.5), but also, because of instantaneous variations in the optical path of the spectrometer, varied over the course of each measurement, preventing interpretation.

6.5.2 Modifications to the spectrometer

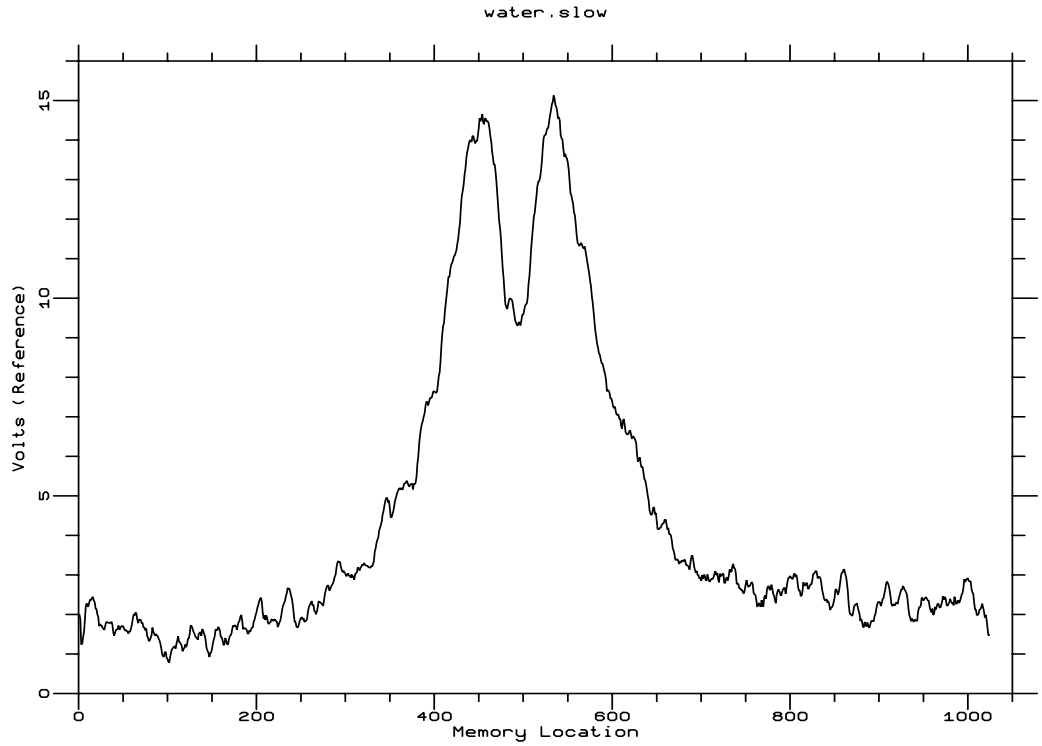

Figure 6.4: Pressure modulated water vapour transmission lineshape for the line at 1271.79cm^{-1}, The mean pressure was 13.5 mb, and the piston amplitude 2.2 mm. This figure shows the least noisy measurement made before the spectrometer was purged.

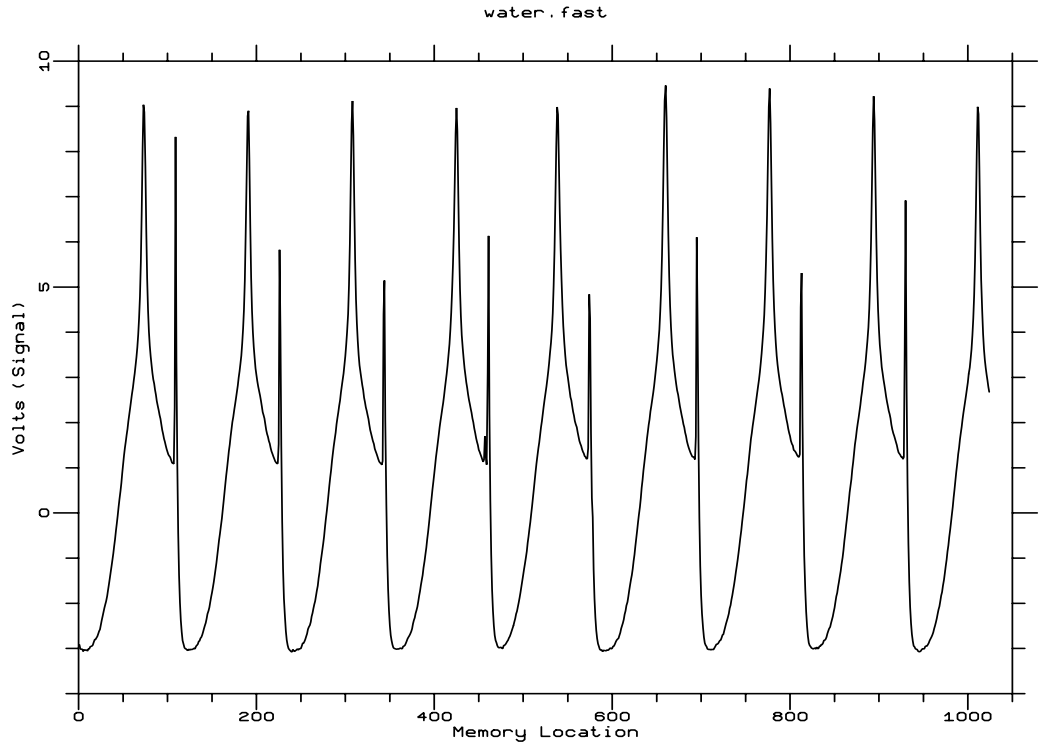

Figure 6.5: Sweep-integrated spectrum of water vapour in the pressure modulator at 12.9 mb pressure, 2.2mm piston amplitude. The transmission maxima are towards the bottom. Eight 'snapshot' spectra were measured per modulator cycle. The offset between the beginning and end of each spectrum arose mainly as a result of absorption by atmospheric water vapour. This was removed in subsequent measurements by purging the spectrometer.

To remove the atmospheric water vapour from the optical path of the spectrometer, an air-tight box of half-inch perspex, which surrounded the entire spectrometer system, was constructed by the Atmospheric Laser Spectroscopy Group at JPL in 1989. All the holes in the optical table were blanked off, and all the screw holes in the box were taped over, to prevent ingress of water vapour. Dry nitrogen was fed into one end through a manifold and pumped out at the other. Indicating dessicant was also placed on trays in the purging box. Measurements when the box was made showed that the water concentration approached 2 ppm after several days. This measurement was deliberately made on a wet day, when the water vapour concentration was expected to be at its largest.

Two circular panels, fifteen inches in diameter, were bolted to the lid of the box for access, and an aluminium plate through one of the sides was used for electrical connections. To insert the modulators into the spectrometer, it was necessary to remove the whole lid of the box. This injected a lot of water vapour into the system, preventing measurement until after about ten hours of purging. Removal of one of the access panels necessitated about an hour of purging before the spectrometer was usable for water vapour measurements.

To reduce the number of times the lid had to be removed, both modulators were inserted into the spectrometer simultaneously, along with the chopper. The brassboard modulator was evacuated to prevent any contribution to the signal. To use the brassboard PMU the flight modulators were moved out of the optical path.

6.5.3 Results – December 1989

A new, high-quality diode was used for the new water vapour measurements in the spectral region 1471.2 to 1472.4 cm^{-1}. Two lines predominated in the measurements: one at 1471.4817 cm^{-1}, with a strength of 7.380E-21 cm^{-1}/(molecule cm^{-2}) and a self-broadening coefficient of 0.3587 cm^{-1}/atm, and the other at 1472.0512 cm^{-1}, with a strength of 1.030E-19 cm^{-1}/(molecule cm^{-2}), and a self-broadening coefficient of 0.4839 cm^{-1}/atm.

The wideband and pressure modulator spectra, as measured, are shown in figures 6.6 and 6.7 respectively. Each sweep of the spectrum took 100 seconds. The general improvement in SNR over the previous measurements is readily observed. Figure 6.8 shows the measured frequency reference spectrum. For the flight model H_2O modulator, with its fixed pressure and higher attainable amplitude, five amplitudes were measured, at 2.12, 2.75, 2.89, 2.93 and 2.97 mm peak to peak per piston. This was a subset of the amplitudes measured during filling at Oxford, but the principle of using a finer mesh near the flight operating conditions was maintained. At the nominal flight amplitude (2.9mm), 23 spectra were averaged, compared with 7 at other amplitudes, to obtain a higher SNR. For the brassboard PMU, lower amplitudes were used, *viz.* 1.45, 2.09, and 2.65 mm peak to peak per piston. The pressures were chosen to cover both sides of the flight modulator filling pressure. First the modulator was filled to as high a pressure as could be achieved (25 mb), to allow investigation of the effects of saturation during part of the cycle, and then the measurements were repeated at lower pressures.

Some of the brassboard measurements were extremely noisy. This was partly because of a lack of time in which to integrate the measurements, and also because the diode power near the line at 1471.483cm^{-1} was low, since the mode was operating near its threshold. Since only one pressure could be used for the flight modulators, more time was available for good measurements.

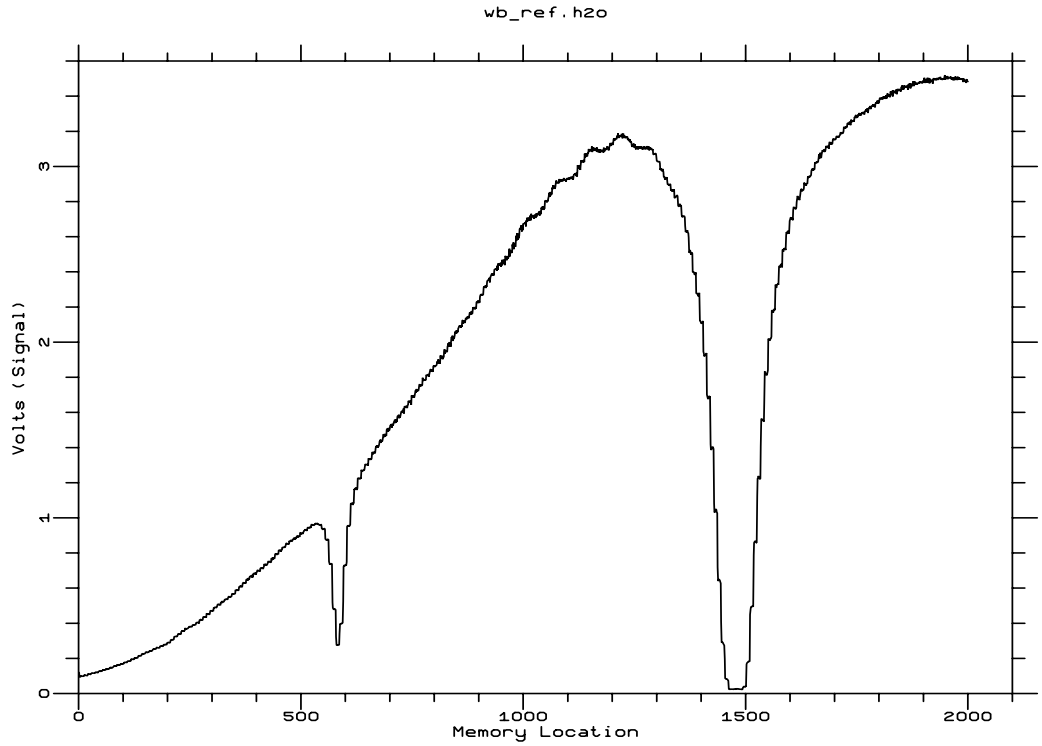

Figure 6.6: Wideband transmission spectrum measured for the flight H_2O PMU. The cell length is 10 cm and the mean pressure is 16.81 mb. The vertical steps in the trace arise from the synthesized ramp generator. The power envelope is clearly visible, as are fringes between the two absorption lines.

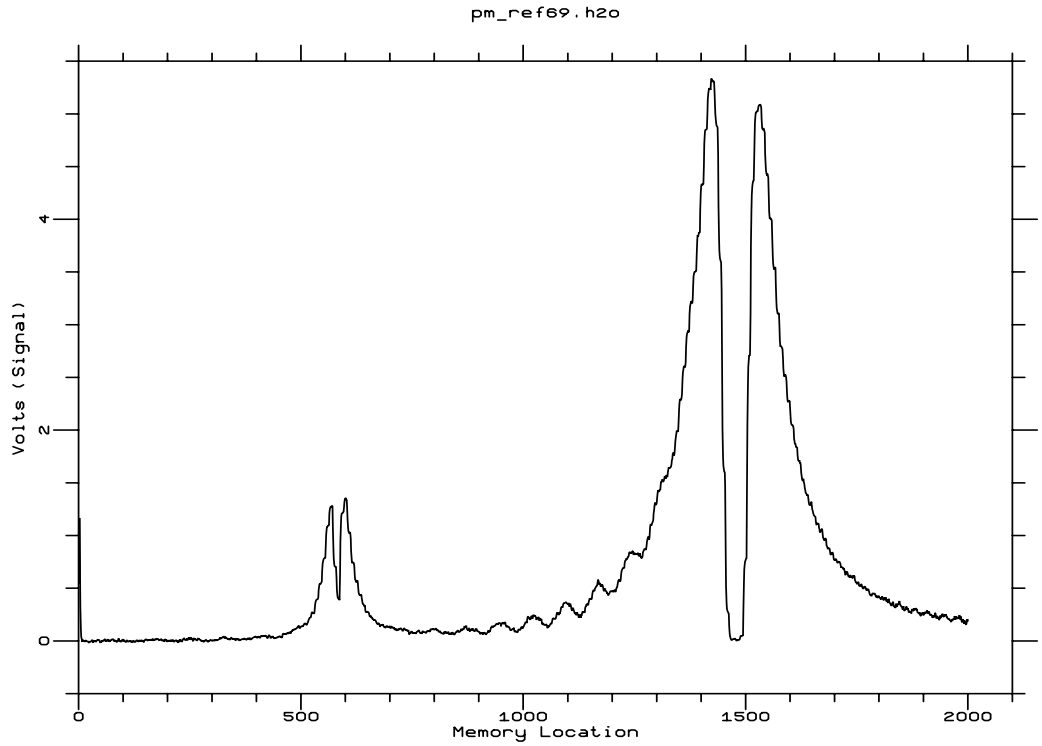

Figure 6.7: Pressure modulator transmission spectrum measured for the flight H_2O PMU. The piston amplitude is 2.9 mm; other conditions are the same as given for figure 6.6, although the power envelope is not obvious in this case.

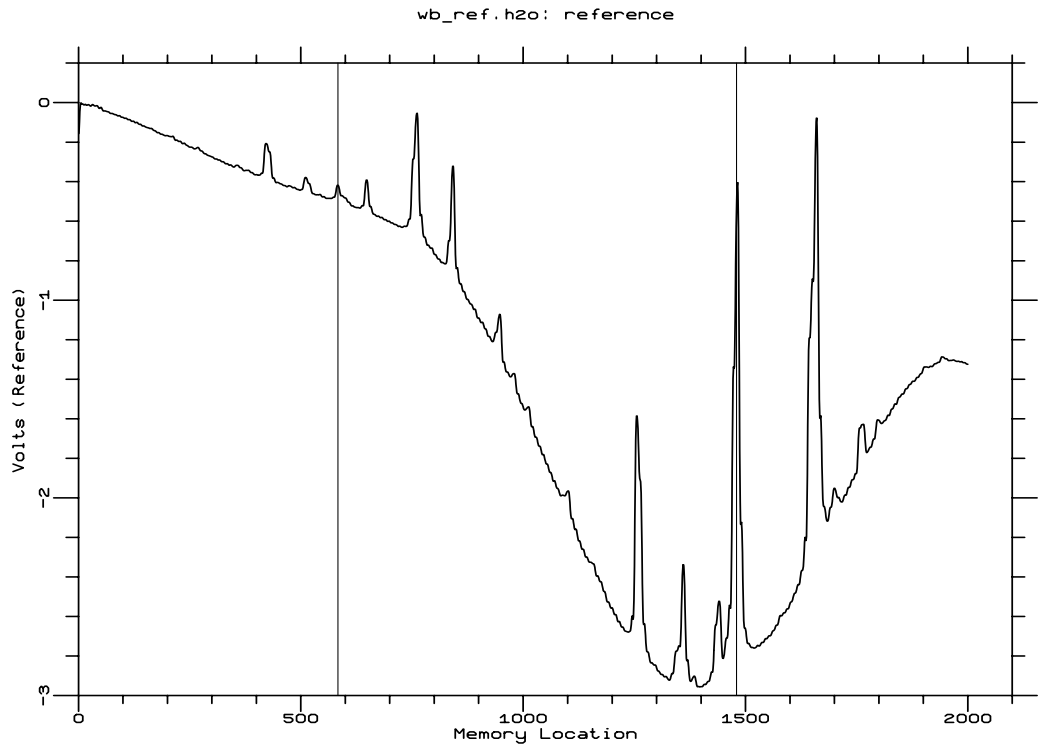

Figure 6.8: Frequency reference spectrum, measured simultaneously with figure 6.6. The gas path is a 15cm cell containing CS_2 and H_2O at low pressure. The two marked lines are the stronger water lines at 1471.48 and 1472.05 cm^{-1} discussed in the text. Most of the rest are CS_2 lines. Absorption increases towards the top of the figure.

Both the wideband and pressure modulated spectra appear to have steps in the signal output. This was because the ramp generator was a digital synthesizer, and so the ramp was really a staircase. The resolution of the signal averager screen was insufficient to show these steps, so that nothing untoward was noticed until my return to Oxford. The steps are effectively filtered out during the analysis but still reduce the frequency resolution of the measurements. More serious are the ripples at the low frequency end of the pressure modulator and wideband spectra, which arose from multiple reflections from the parallel cell windows. This effect decreases in magnitude as the solid angle of the beam increases (see Appendix D), and so is unlikely to be as marked in the PMIRR instrument.

The measured water vapour transmission profile in figure 6.7 is asymmetric, like that for N_2O shown in figure 6.3. This asymmetry is thought to be caused by the pressure dependence of the line centre frequencies (described in Appendix A). It is discussed in more detail in chapter 7.

6.6 Carbon Dioxide

For the measurements in February 1988, the flight mixture of isotopic CO_2 was not available, so that some isotopically-enriched CO_2 (produced by ISOTEC) was used. The gas contained (by volume) 31% 626, 48% 628, and 20% 828; the missing 1% was probably 636. This isotopic mixture was not exactly in the 4:4:1 ratio which has been used for the flight instrument. The spectral lines of the 628 isotope, used in channel 1 of PMIRR, are very weak, and a direct transmission spectrum showed a peak absorption of only about 2 % with 80 mb in the 4 cm cell attached to the modulator.

Measurements of the lineshape were made with the modulator running at a mean pressure of 85 and 42.5 mb, and at amplitudes of 0.6, 1.2, 1.8 and 2.4mm. 25 scans were averaged for each spectrum. The direct-absorption trace showed the noise typical of the short time constant on the lock-in amplifier, while the PMU trace contained a slow sinusoidal background signal not traceable to aliasing or interference between the ramp and modulator frequencies (they were altered without effect).

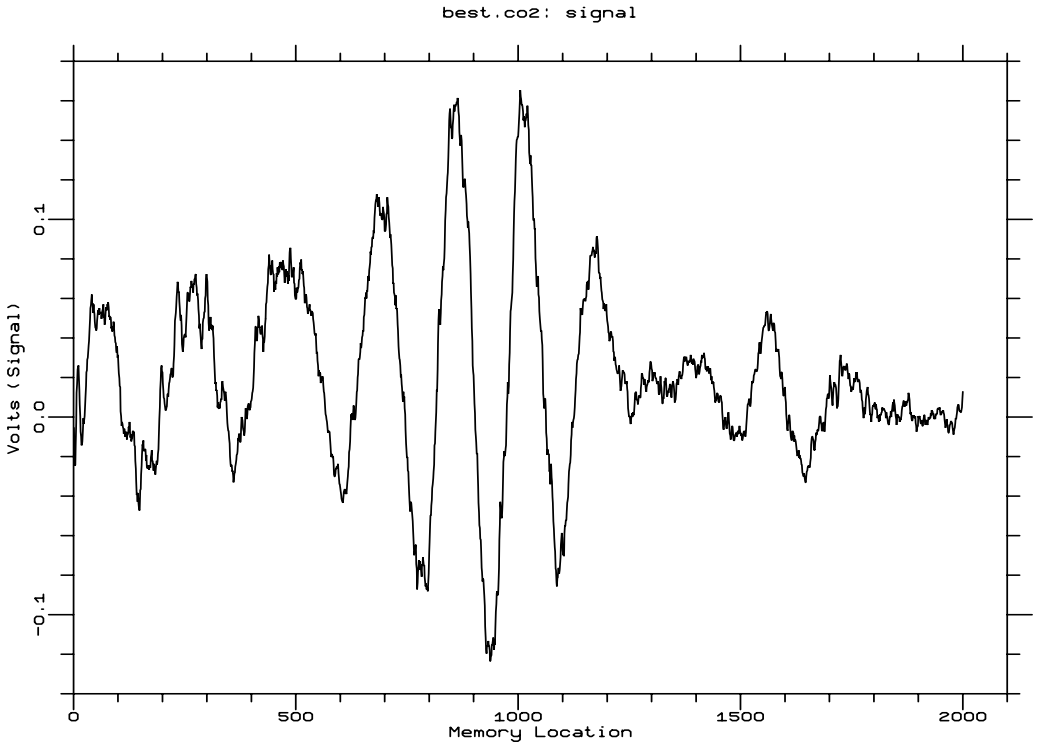

Figure 6.9: measured spectrum of flight pressure modulator at about 1300 cm^{-1}. The isotopic spectral lines are very weak, and are obliterated by coherent fringing caused by multiple reflections between the inner faces of the cell windows (see Appendix D).

When the measurements were repeated for the flight modulators in December 1989, sinusoidal fringes, as instanced in figure 6.9, were observed. These completely obliterated the signal from the absorption lines, and so no useful spectral measurements on the 628 band could be made. The free spectral range of the fringes, 0.125cm^{-1}, was characteristic of a 4 cm étalon, indicating that they arose from interference between multiple reflections from the parallel windows of the pressure modulator cell, as also occurred for H_2O. By altering the alignment of the PMUs, it was possible to move the position of the fringes, but not their FSR. The dependency of such fringes on beam angle is discussed in Appendix D. Since the beam through the CO_2 cell is divergent in the optical configuration of PMIRR, the effect of the multiple reflections will be very much attenuated.

As measurement of the transmission profile for water vapour was more important than that for CO_2, I spent almost all of the time available making measurements with H_2O. As a result, it was not possible to change the diode in order to measure the pressure modulated lineshape of the flight CO_2 PMU in a '626' band, for which the signal to noise ratio should be larger, because the lines are very much stronger.

This page intentionally left blank.

Chapter 7:
Analysis of Spectroscopic Measurements

In order to draw inferences from the data described in chapter 6, they must be converted into a form which gives the transmission as a function of frequency. The converted data may then be used directly, and scaled for different line parameters to predict a whole channel, or may be compared with the predictions of a pressure modulator model. Agreement of the calculated and measured transmission within the expected uncertainty gives confidence in the ability of the model to predict the pressure modulated transmission for other operational régimes on which direct measurements were not made, and is therefore the more generally useful approach.

This chapter describes the processing scheme for the measurements, with particular emphasis on the H_2O measurements as these are of direct importance to PMIRR, and assesses some of the limitations arising from the measurement techniques. Some suggestions for possible improvement of these are given in chapter 8. Since the analysis relies heavily on the calculation of transmission from the model output, this is described first.

7.1 Calculated spectra

The traditional method of calculating pressure modulator lineshapes [*e.g.*, Schofield 1980] uses the so-called 'two-pressure' approximation, which considers the PMU pressure variation as a square wave at the PMU frequency between two characteristic pressures. When the temperature cycle amplitude is large, as is the case for both PMIRR modulators, a two-pressure, two-temperature representation may be used. This technique significantly reduces the time taken to compute the transmission of a whole instrumental signal channel, at the cost of fitting the two pressures to the calculated transmission for a few representative lines, or, to lesser accuracy, fitting square waves to the cell pressures and temperatures. However, the short spectral ranges used in the tunable diode laser experiments, and the speed of present-day computers, render this procedure unnecessary in this instance.

The pressure modulated and wideband transmissions are calculated in a manner analogous to their measurement. Each set of operating conditions used for the measurements was input to the modulator model to provide the cell temperature and pressure cycles, and their variances and covariances. The instantaneous transmission profile over the spectral region of interest is calculated at each of sixteen points, equally spaced around the cycle, using the theory outlined in Appendix A. Four such transmission spectra are shown in figure 7.1. The sine and cosine components of the varying transmission at the PMU frequency are derived from the sixteen transmissions, and combined, in a phase which gives zero frequency-integrated quadrature, to produce the pressure-modulated transmission. The sine and cosine components, and the resulting pressure modulated and quadrature transmission spectra, are shown in figure 7.2. The wideband transmission is obtained from the mean of the instantaneous profiles.

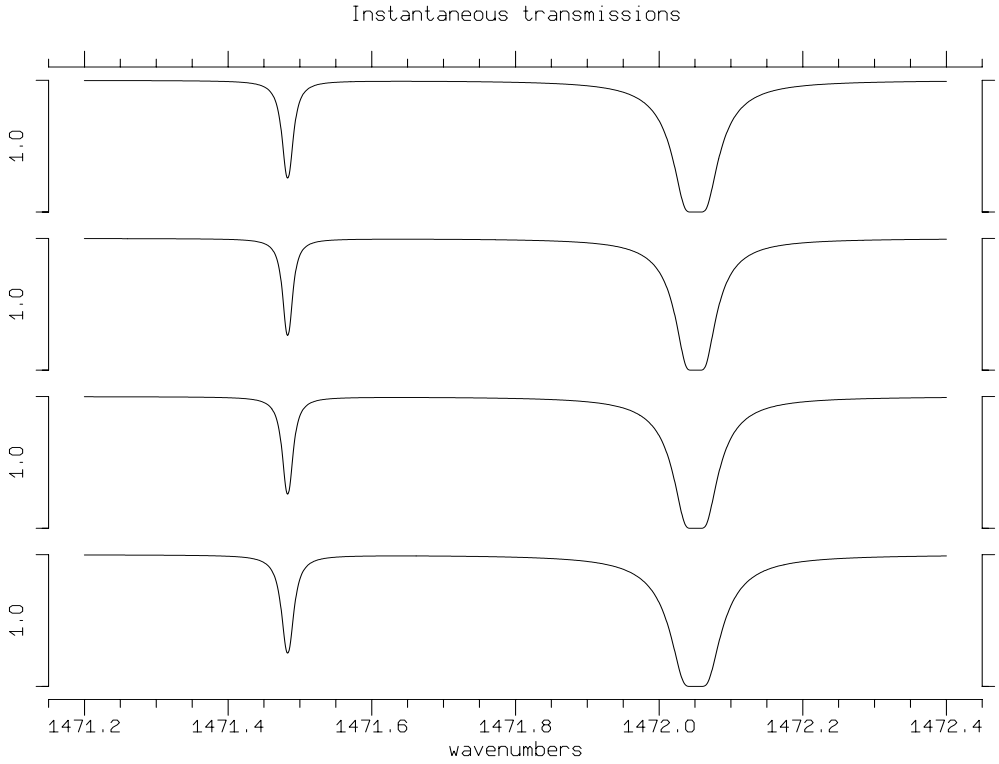

Figure 7.1: Calculated instantaneous transmissions through the cell of the PMIRR H_2O pressure modulator, at 16.81mb pressure and 2.9mm piston amplitude, at four points equally spaced around the modulator cycle. The cycle phase increases from the bottom of the figure to the top.

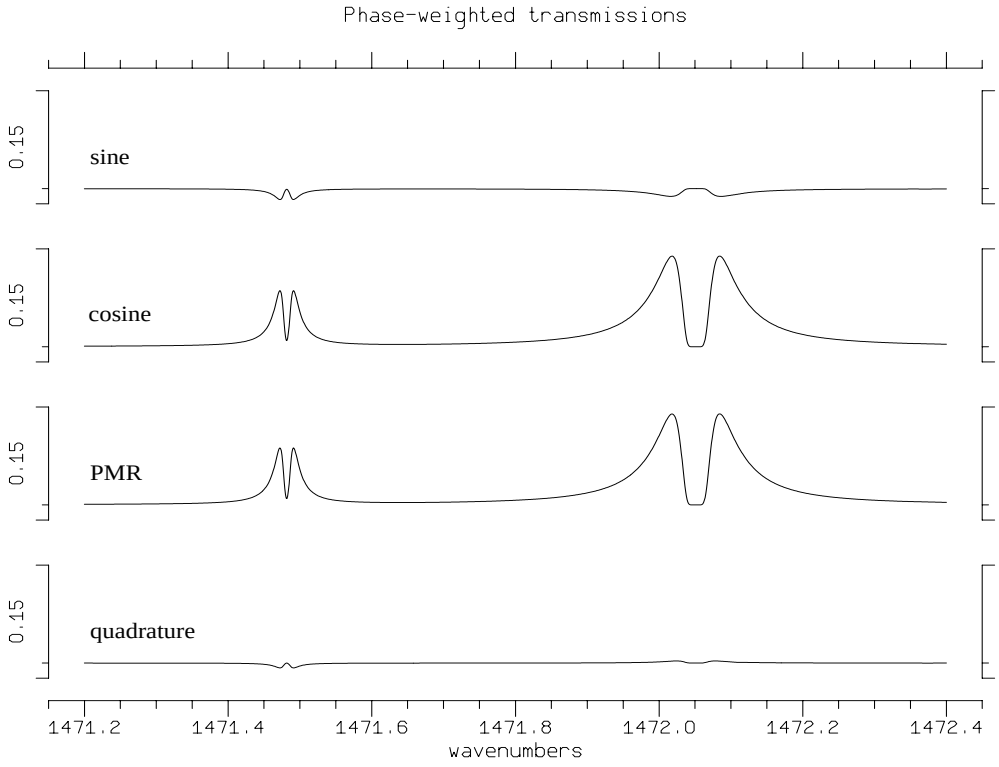

Figure 7.2: Sine and cosine transform spectra calculated for the PMIRR H_2O pressure modulator, at 16.81mb pressure and 2.9mm piston amplitude, and their combination into pressure modulated and quadrature transmissions.

Strong lines from outside the region of interest contribute to the absorption within it. In the case of the water vapour measurements, the lines with centre frequencies within only 2cm^{-1} of the measurement interval were used, since exclusion of lines further away in frequency results in a change of integrated transmission of less than 0.01%. On the same principle, the N_2O lines used overlapped the measurement interval by 5cm^{-1}.

7.1.1 Pressure shifts

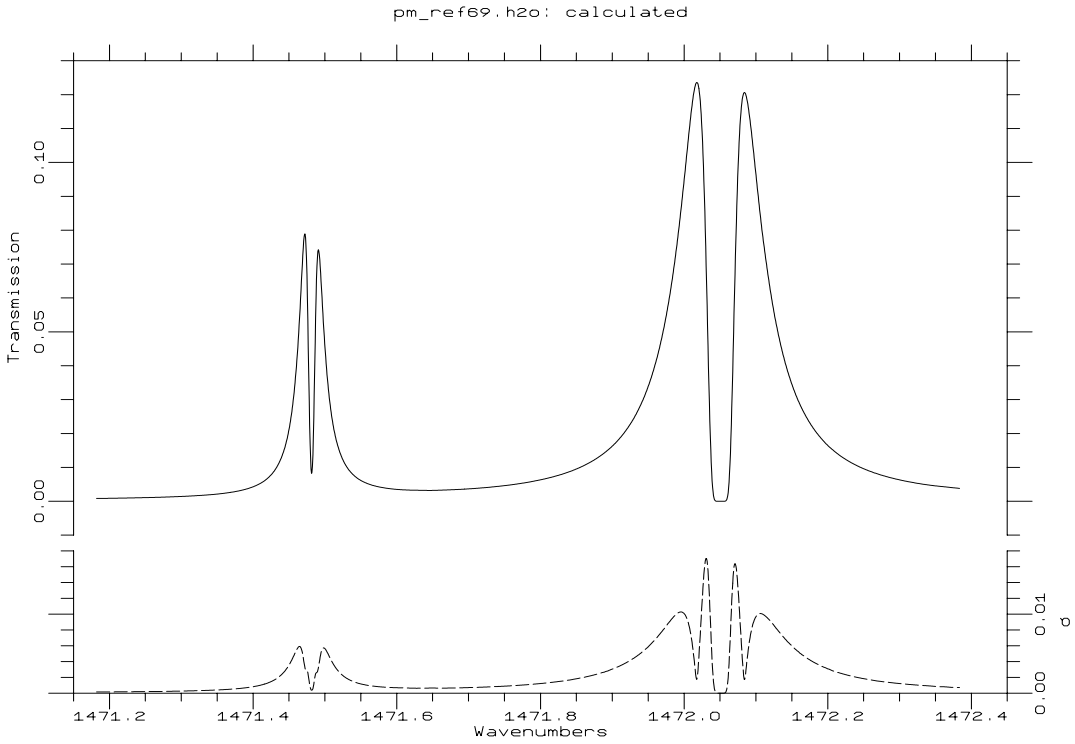

Figure 7.3: The effect of pressure shift on PMU transmission, calculated from the model predictions for the PMIRR flight model H_2O PMU under its nominal operating conditions. The dashed curve is the expected transmission uncertainty, calculated from the uncertainties in the line data and cell pressure and temperature cycles (see next section).

Pressure shifts cause the frequency of the absorption peak for each line to vary throughout the cycle. This leads to an asymmetry in the pressure-modulated lineshape, as shown in figure 7.3, which shows the calculated pressure-modulator lineshape for the nominal flight H_2O PMU conditions of 16.81mb and 2.9mm piston amplitude, assuming that the pressure shift coefficients are 0.05 times the self-broadening coefficients for all the lines used (although the two are not necessarily expected to be proportional to each other – see Appendix A). The resolution of the measurements was insufficient to calculate pressure shift coefficients even for the two strong water lines, since the effective instrumental resolution was equivalent to a shift of 5% of the self-broadening coefficient at 40mb. However, an asymmetry similar in magnitude was observed in the measured lineshapes for water vapour (see chapter 6), and so all the water vapour calculations assumed a pressure shift coefficient 5% of the self-broadening coefficient.

7.1.2 Uncertainties in the calculated transmission

The uncertainties in the calculated transmission arise predominantly from the uncertain-

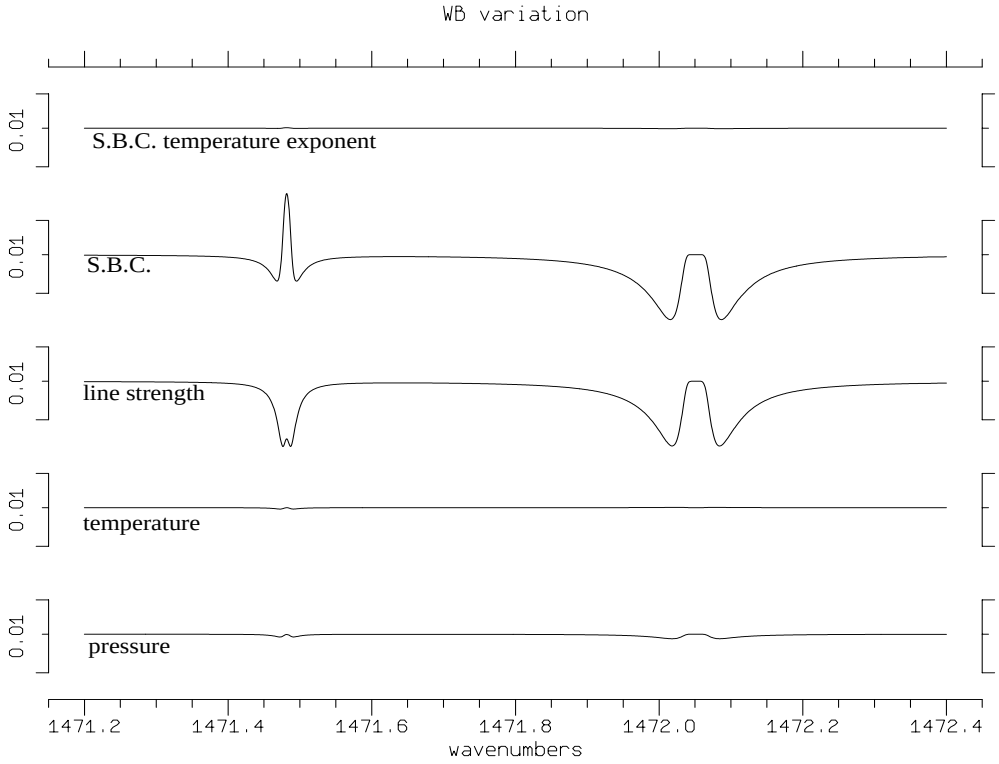

Figure 7.4: Contributions to the uncertainties in the wideband transmission of the PMIRR flight model H_2O PMU, over the spectral interval of the tunable diode laser measurements, arising from errors in the line data and modelled temperature and pressure cycles. The lowest two curves are derived from orthogonal combinations of the pressure and temperature cycle uncertainties, and are marked with the predominating component. S.B.C. is the self-broadening coefficient.

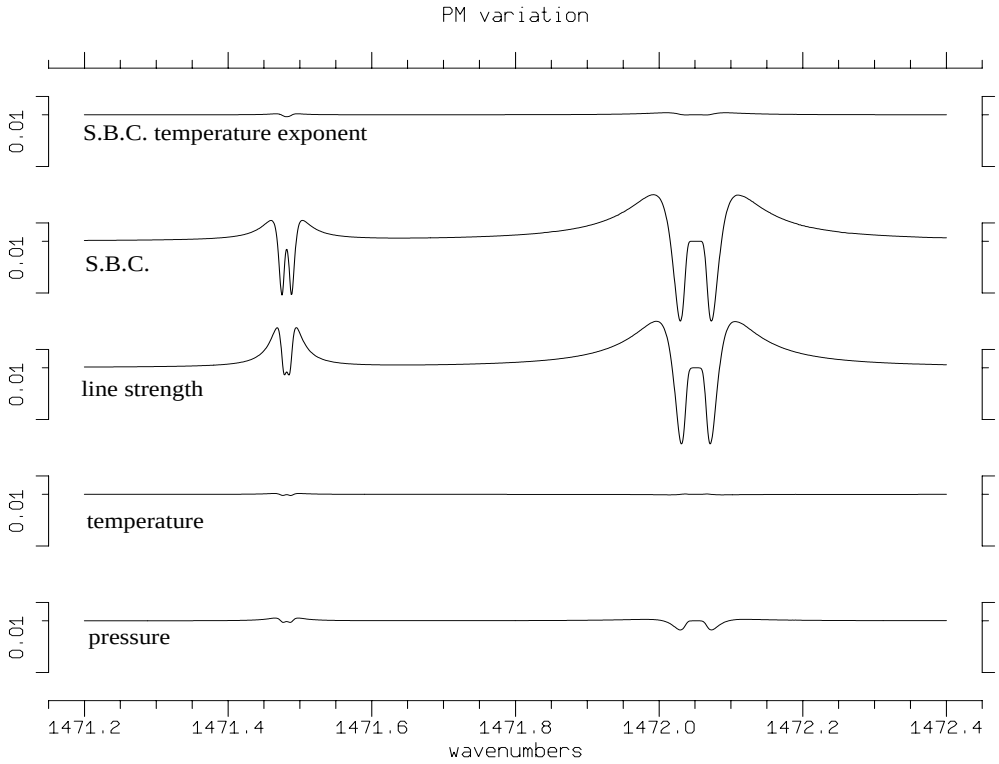

Figure 7.5: Contributions to the uncertainties in the pressure modulated transmission of the PMIRR flight model H_2O PMU, over the spectral interval of the tunable diode laser measurements, arising from errors in the line data and modelled temperature and pressure cycles. The lowest two curves are derived from orthogonal combinations of the pressure and temperature cycle uncertainties, and are marked with the predominating component. S.B.C. is the self-broadening coefficient.

ties in the line data and, to a lesser extent, those in the modelled pressure and temperature cycles. To calculated the uncertainties, the line strengths, self-broadening coefficients and their temperature exponents were perturbed in turn, for all the lines simultaneously. The uncertainties in the pressures and temperatures are correlated, so their contributions were generated by perturbing the pressure and temperature at each of the sixteen cycle points by each of the two orthogonal error vectors calculated from the 2×2 pressure-temperature covariance matrix provided by the model. For each perturbation, a variance contribution of optical depth was calculated; these were summed and square-rooted before conversion into transmission uncertainty. Figures 7.4 and 7.5 show the relative contribution to the total uncertainty in the wideband and pressure modulated transmissions of the flight model H_2O modulator, at its nominal operating amplitude of 2.9mm, over the spectral interval used in the tunable diode laser measurements. It is evident that the errors in the line data, particularly in the line strengths and self-broadening coefficients, dominate. The accuracies in the HITRAN 91 database in the spectral region for water vapour are about 13% for line strength, 15% for self-broadening coefficient, and 9% for its temperature exponent. The relative (line-to-line) errors are smaller than the systematic uncertainty in the band as a whole [J J Remedios, private communication].

The pressure cycle comparisons described in chapter 5 indicate, for water vapour at least, that the model underestimates the magnitude of the pressure cycle in the brassboard modulator by 15% at moderate pressures, and up to 24% at high pressures. The uncertainty in the modelled pressure cycle, and that used to calculate the contributions shown in figures 7.4 and 7.5, is approximately 4% of the pressure cycle amplitude, for the operating conditions of the PMIRR flight model H_2O modulator (16.81mb mean pressure, 2.9mm piston amplitude). Thus in reality the contribution of the pressure cycle to the transmission uncertainty should be about 5 times larger than shown in these figures. This is still smaller (by about a factor of 2) than the contributions of the self-broadening coefficient and line strength uncertainties, and so the modelled pressure cycle was used unadjusted for the calculations, rather than correcting it to conform with the pressure cycle measurements. This approach allows the use of the spectroscopic measurements as a direct and independent diagnostic of the model, and avoids some of the possible difficulties associated with transferring the measured pressure cycle errors from the brassboard to the flight modulator, which operated at a larger amplitude, a considerably higher frequency, and a higher temperature, the last of which greatly affects the amount of gas exchanged between the walls and the cell by adsorption.

7.2 Processing scheme

The wideband and PMR measurements were all processed in the same way. The sweep-integrated scans were first split into individual transmissions, by identifying the sweep flyback lines, and then treated as wideband spectra. The signal averager memory locations were converted to frequency, and the signal voltage in each memory cell was divided by the full scale signal voltage ('baseline') to convert it to transmission. The noise level was estimated by calculating the RMS deviation of the signal from a filtered mean, in a region without strong lines, and scaled by the baseline to give the random component of the transmission uncertainty. Allowance was also made for a combined non-linearity of 2% in the detector and preamplifier. The expected transmission was calculated, as described above, and the calculated and measured transmissions were compared, providing a quality-of-fit indicator, from which the probability that the measurements conform to the model could be derived, as in chapter 5. These processes are discussed below.

7.2.1 Frequency axis

All the measured spectra contain one channel of frequency reference, as described in chapter 6. This consists of either a trace of the transmission through a confocal étalon with approximately 300 MHz FSR, or a spectrum of CS_2 and H_2O at low pressure. The signal averager memory location was converted into a frequency scale by identifying spectral features with known frequencies in the reference channel, and fitting a quadratic curve around the local maximum or minimum in the data. In the case of water vapour, since the reference cell contained a small amount of water vapour at low pressure, its line centres are very distinct, and so a very small frequency error results. The diode frequency/current characteristic, was very linear (as étalon fringes were measured to have a constant separation in diode current), so that only two lines were required to convert the memory location into frequency. For the N_2O measurements, the diode current-tuning rate was non-linear, so that a quadratic fit is used. No low-pressure lines were available, so the three weakest lines have been used as frequency references.

The CS_2 lines or etalon fringes, the positions of which were known to high precision, could have been used to identify the frequency scale. However, fitting the frequency to these data leads to an inconsistency between the measured and calculated spectra, mostly from pressure shifting but partly from errors in values given in the line database for line centres. To avoid an apparent mismatch between the measured and calculated spectra arising from a frequency offset, my analysis is based on the assumption that the measured line centres have frequencies as published. The etalon and CS_2 measurements have been used as spot checks to verify the linearity of the resulting frequency scale.

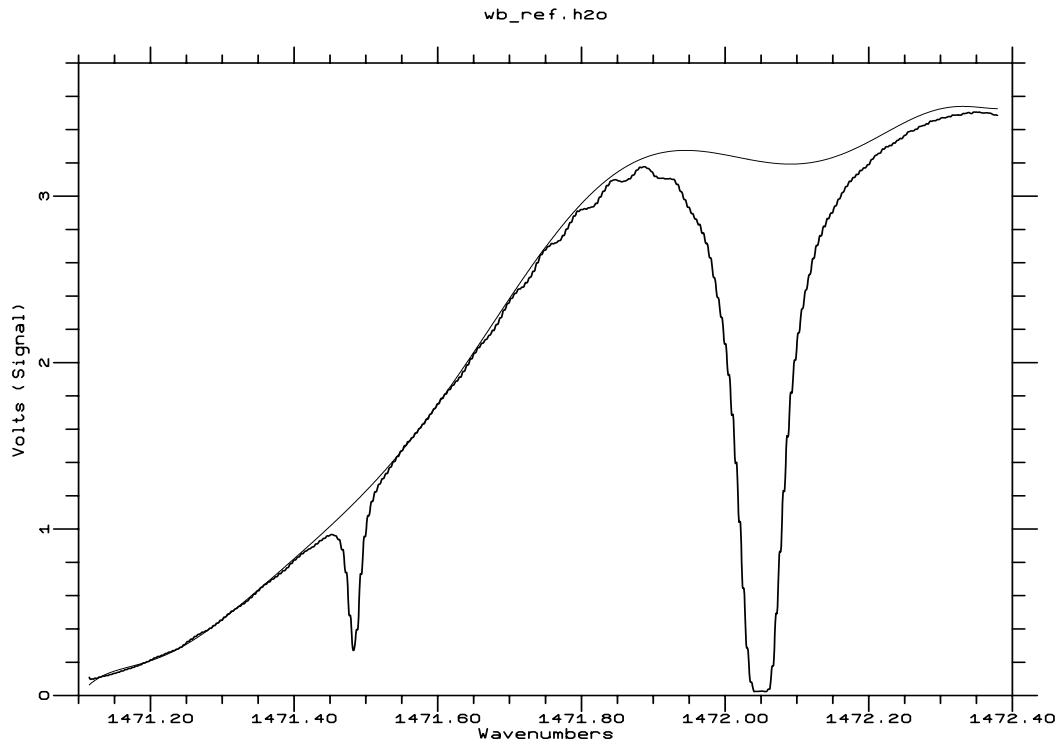

Figure 7.6: Measured (raw) spectrum and fitted full scale signal (baseline) for the PMIRR flight model H_2O PMU, which contains 16.8mb of water vapour. The fitted baseline is a ninth-order polynomial.

7.2.2 Baseline calculation

The variable full-scale signal or 'baseline' has been described in chapter 6. As stated above, knowledge of the baseline is required to convert the measured detector output to transmission. Since the attempted direct measurement of the baseline (described in chapter 6) was unsuccessful, it was derived indirectly, from the ratio of the measured signal to the calculated wideband transmission. The ratio of the measured spectrum to the calculated transmission was fitted to a polynomial curve, the degree of which was increased until the fit probability was close to unity. The fit extended into the near wings of the lines, but ignored the line centres, by rejecting points for which the calculated transmission was less than 0.95. An example is shown in figure 7.6, which shows the measured wideband spectrum and fitted baseline for the PMIRR flight H_2O PMU at its normal pressure. The fitted curve in this figure is a ninth-order polynomial.

The shape of the baseline for the pressure modulator lineshapes is identical to that of the wideband measurement under the same conditions. However, because the relative gain difference between the wideband and pressure modulator spectra is uncertain, a scaling factor between the wideband and pressure modulator full scale signals was derived by fitting the measured spectrum, divided by the baseline polynomial, to the predicted pressure modulator spectrum. Where the baseline is very small, the measured transmission spectrum is noisy, so the square of the baseline value at each spectral point was used as a weight for the fit. A consequence of this method is that only width information can be inferred from a single line – two or more lines of different strengths must be used in order to compare absolute transmission measurements, and the quality of fit judged from all of them.

7.2.3 Sweep-integrated spectra

The sweep-integrated scans, which consisted of wideband transmission measurements at eight PMU cycle phases, were first divided into individual spectra, by identifying the flyback lines which arose from the rapid backwards scanning of the spectral region on the falling ramp edge. In principle, each individual transmission spectrum could be compared to that predicted by the model for the range of the cycle that it represents, as for other wideband spectra. However, there are a number of limitations in the data which prevent this, and so the sweep integrated measurements could not be compared quantitatively with calculations.

The spectral resolution of the measured spectra was determined by the number of sampled points across a line. When the sweep-integrated scans were divided into individual measurements, there were generally only about 20 samples per line, which was not adequate to infer information about the pressure cycle.

Because no chopper was used in the technique, the signal chain was AC coupled to remove the large DC ambient light signal. This meant that there was neither a zero level nor a 'full scale' signal, and to convert the measurements into transmission it was necessary to infer both of these from comparisons of the calculated and measured spectra, to assume that the transmission for one of the spectra is what it is calculated to be, and to compare the other spectra against it. This is done below, but is not very satisfactory.

An additional problem arises from the emission of the gas in the PMU, which varies at the modulator frequency as the temperature of the gas is cycled. At 300 K, the Planck function peaks at around 10 μm, which is within the sensitive wavelength range of the detector, and the total (frequency-integrated) emission from 1 cm^2 of the cell is about 46 mW, varying by about 0.6 mW/K. Thus a cell temperature cycle amplitude of 3K has a maximum varying component of emission of 1 mW. This is comparable to a typical laser output power, so that, depending on the emissivity of the gas around the Planck function peak, the emission signal may appear on spectra measured without a chopper. In principle, the emission signal could be inferred from measurements of the detector signal without the TDL operating, but while the laser output is highly collimated, the emission of the modulator cell is not, so the results would depend on the detector optics, and the experimental techniques are not entirely straightforward. (For further details, see Irwin [1991].)

7.3 Results for nitrous oxide

For N_2O, two pressures (13 and 20–21mb) and four amplitude settings (0.57, 1.2, 1.75 and 2.32 mm) were used; the results are shown in figures 7.7 – 7.15.

These measurements were among the first PMR transmission measurements made using the tunable diode laser technique. However, as the technique was at the beginning of its development, the measurements are deficient in many ways. The wideband and pressure modulated transmissions were recorded simultaneously, and the confocal étalon trace was measured separately. This implied that to fit the frequency scale to these measurements, lines had to be located in the wideband transmission spectrum. This was not entirely satisfactory, as the lines were broad, and so provided a poor frequency reference. Furthermore, the frequency / current characteristic of this diode appears from the étalon traces to have been non-linear.

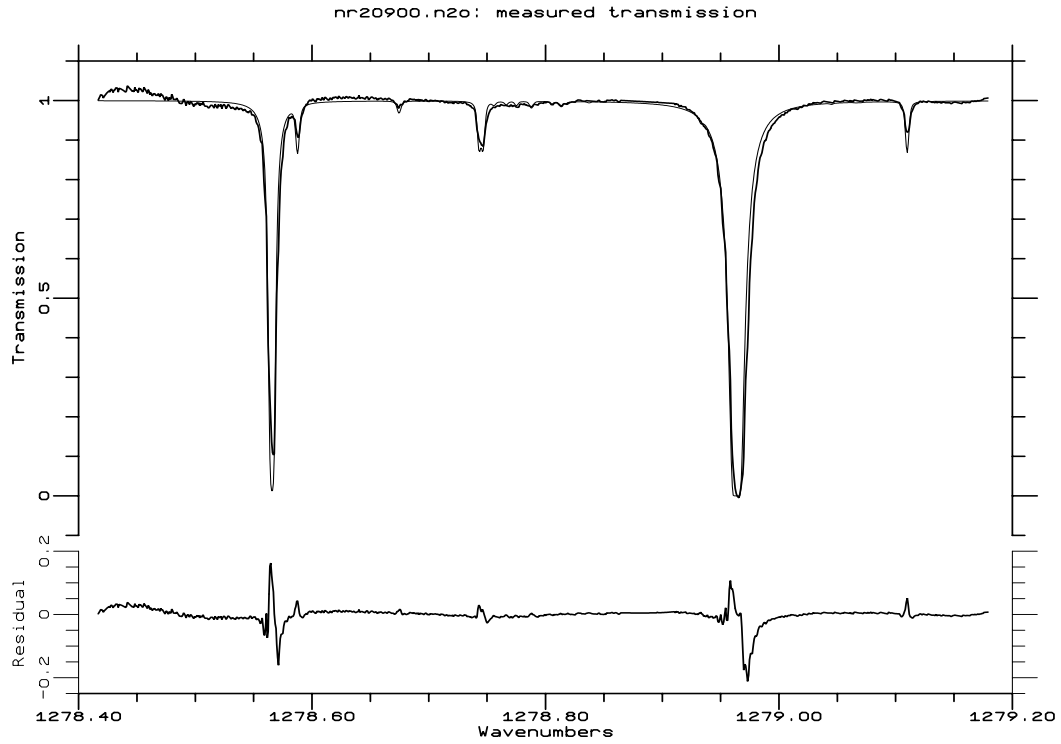

Figure 7.7: Measured (thick line) and calculated (thin line) wideband transmission of the brassboard PMU, containing N_2O at 12.95 mb.

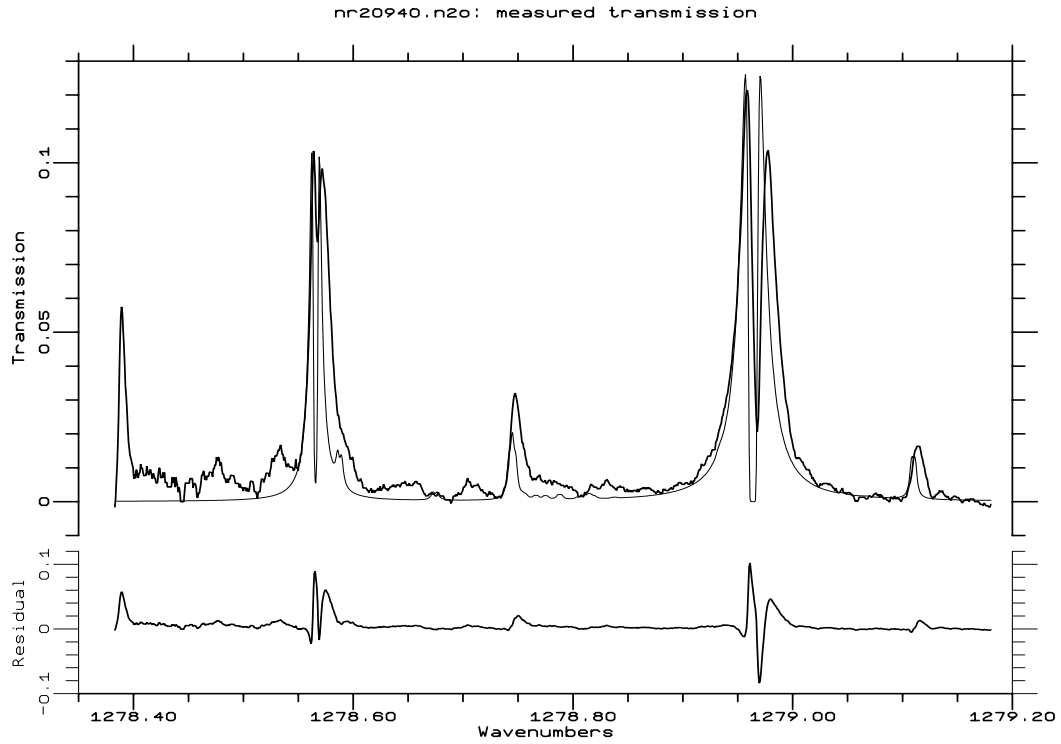

Figure 7.8: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 12.68 mb, 1.20 mm piston amplitude. These measurements have been filtered.

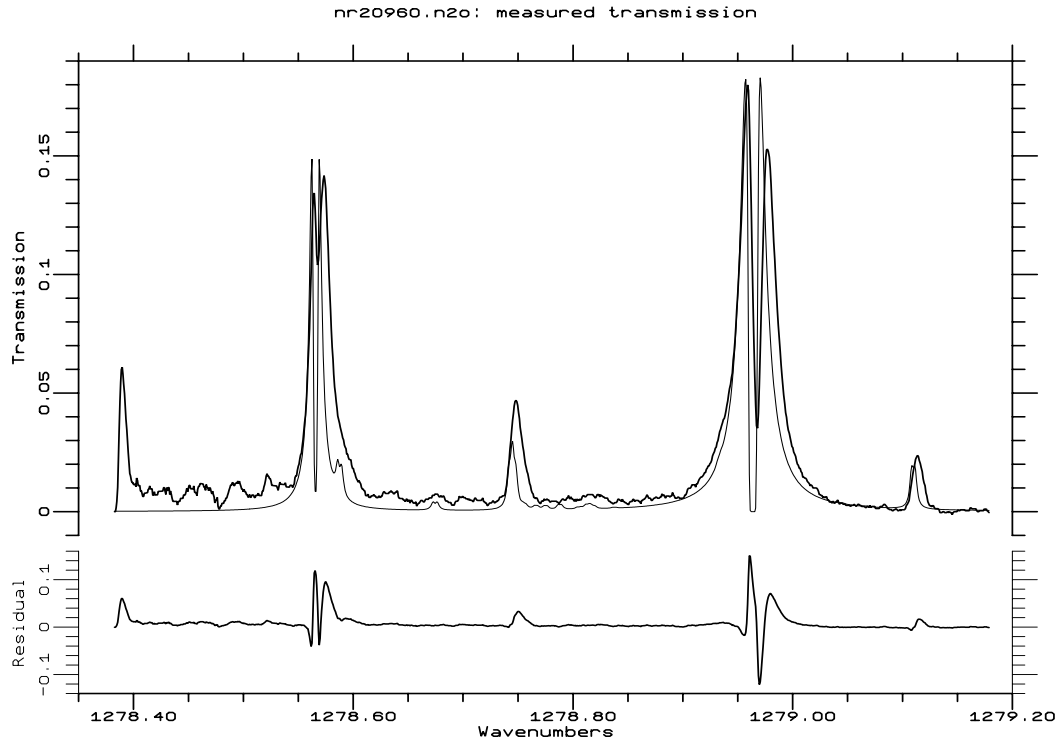

Figure 7.9: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 12.59 mb, 1.75 mm piston amplitude. These measurements have been filtered.

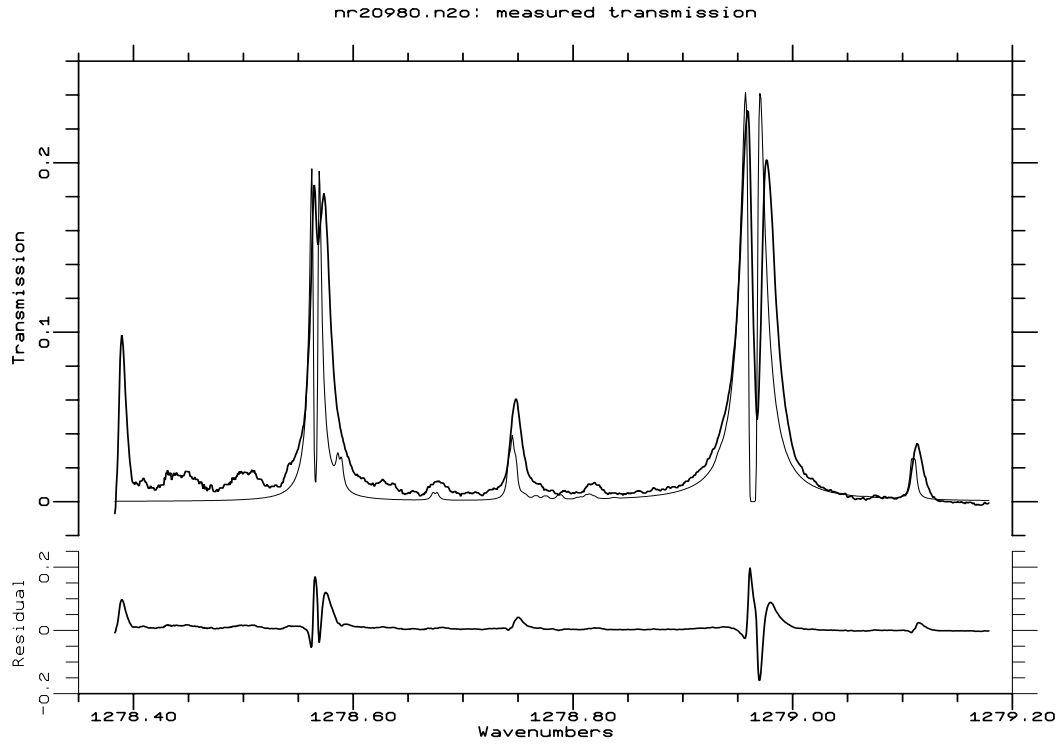

Figure 7.10: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 12.45 mb, 2.32 mm piston amplitude. These measurements have been filtered.

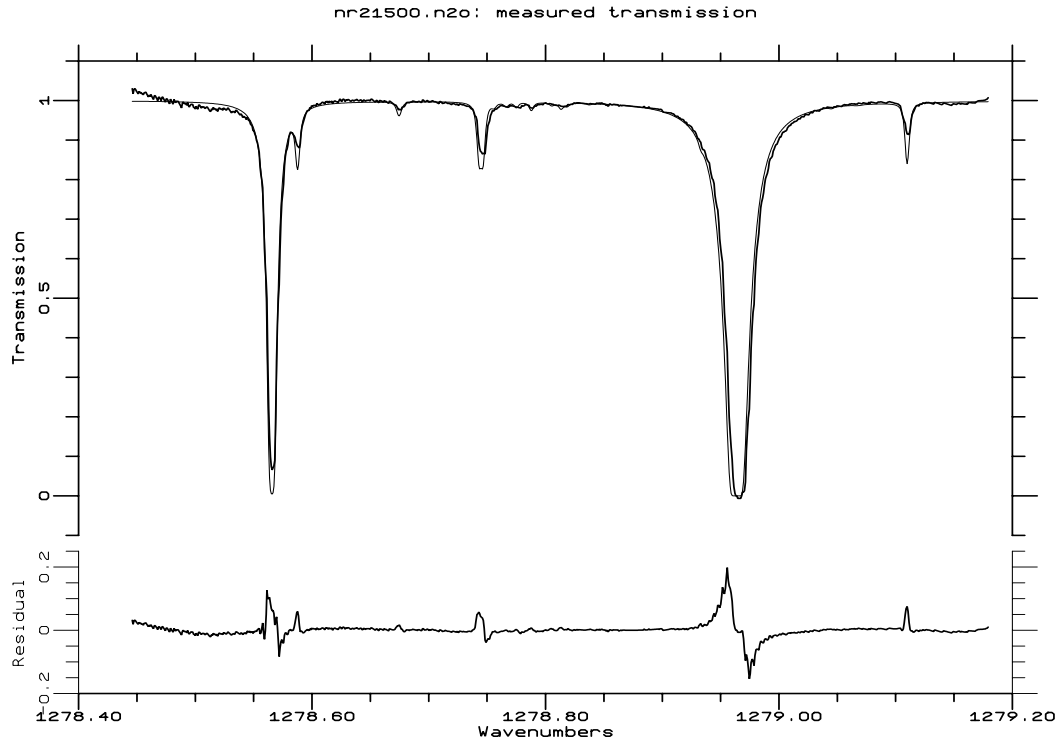

Figure 7.11: Measured (thick line) and calculated (thin line) wideband transmission of the brassboard PMU, containing N_2O at 20.53 mb.

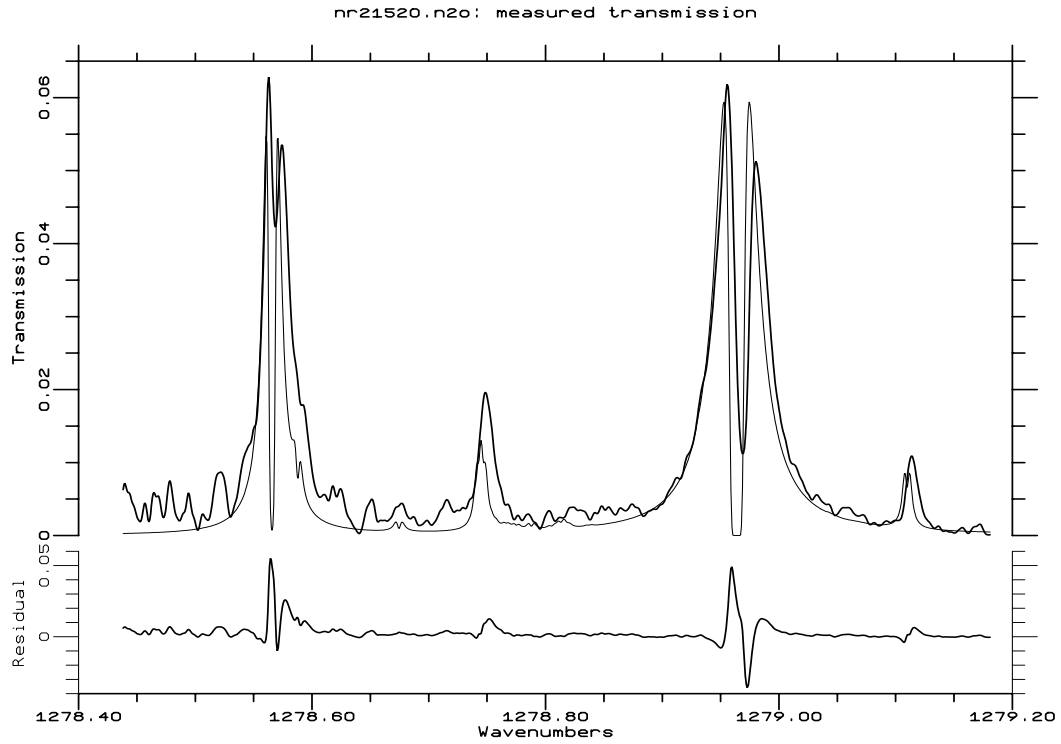

Figure 7.12: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 20.67 mb, 0.57 mm piston amplitude. These measurements have been filtered.

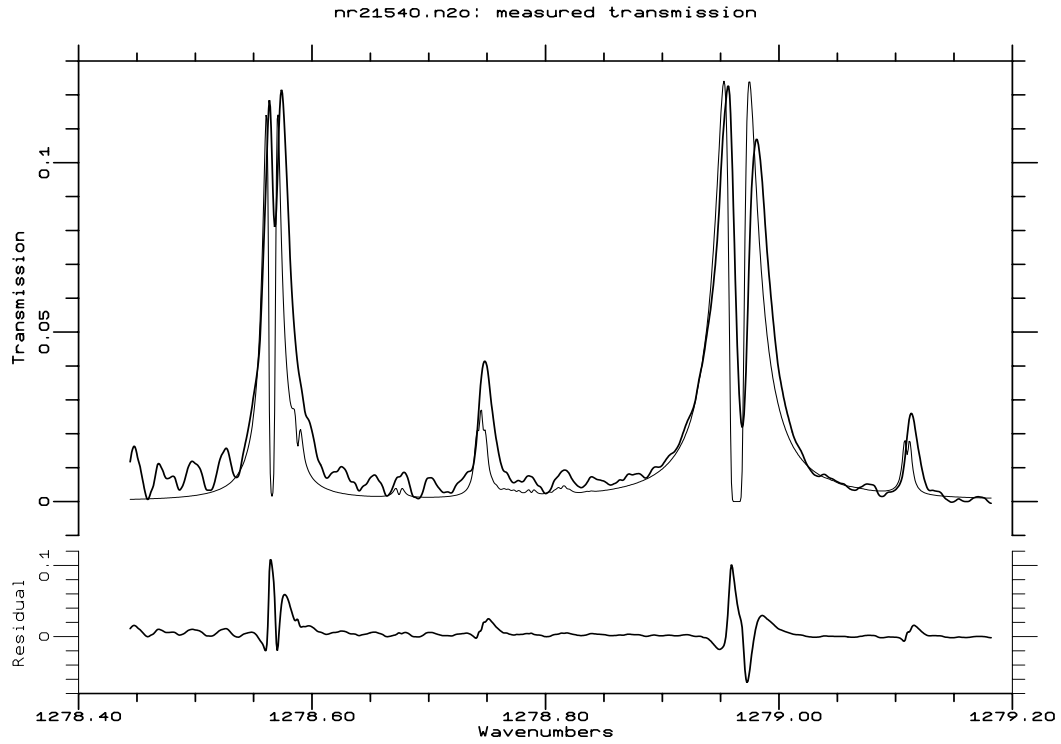

Figure 7.13: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 20.67 mb, 1.19 mm piston amplitude. These measurements have been filtered.

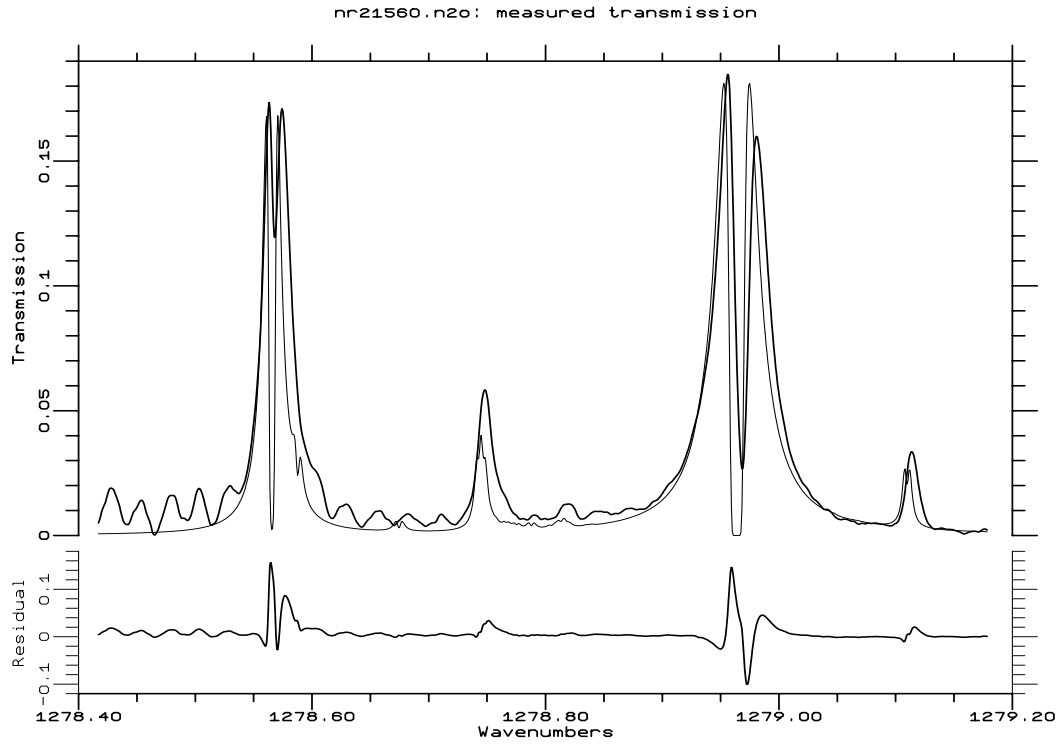

Figure 7.14: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 20.80 mb, 1.75 mm piston amplitude. These measurements have been filtered.

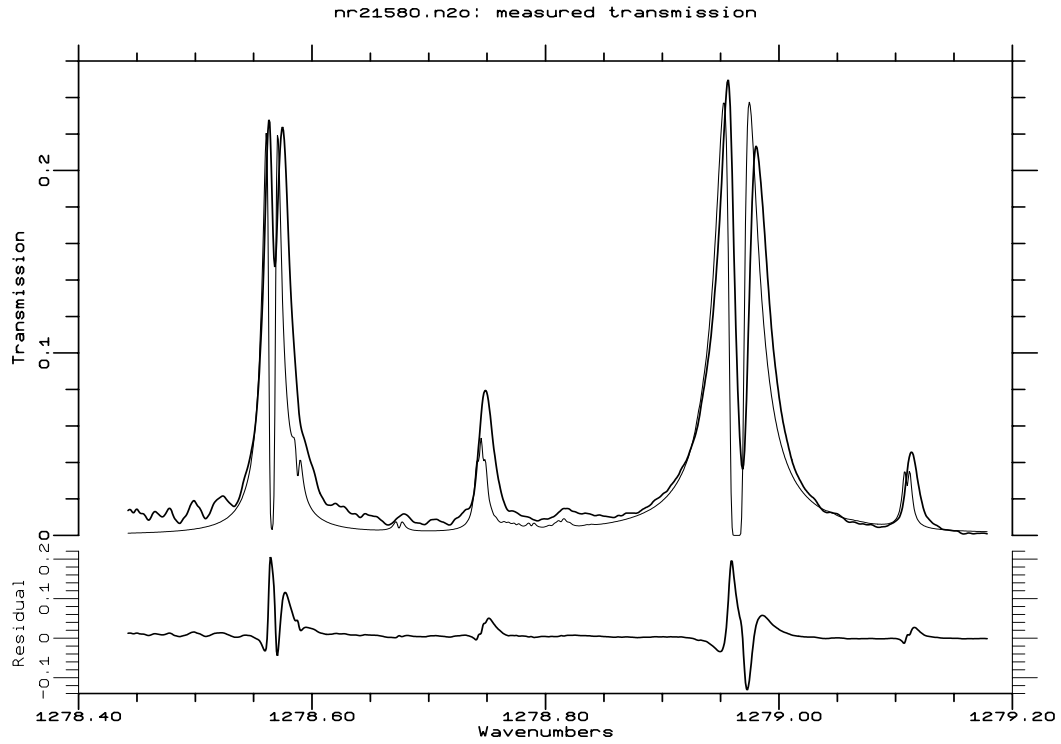

Figure 7.15: Measured (thick line) and calculated (thin line) pressure modulated transmission of the brassboard PMU, containing N_2O at 20.93 mb, 2.31 mm piston amplitude. These measurements have been filtered.

The modulator was not well-characterized at the time of these measurements. For this reason the uncertainty in the calculated transmission is enormous and has been omitted. For all these measurements, the calculated spectra are too narrow because of the absence of self-broadening parameters from the line data. The combined effect of the pressure shift of the lines and the uncertainty in fitting the line centres to derive the frequency scale is evident; there is a noticeable frequency offset in all the plots of pressure modulator transmission. The signal averager used for the N_2O measurements had less memory than that used for H_2O measurements, so the effective spectral resolution of the TDL/averager combination was smaller, and was insufficient to resolve the many weak lines predicted.

These comments notwithstanding, several inferences may be made from the N_2O measurements. The measured form of the strong singlet at 1278.9637cm^{-1} shows no evidence of saturation; the central null does not reach zero at any pressure or amplitude. The null is even smaller for the doublet at 1278.5653/1278.5662 cm^{-1}. The ground state energies of these lines are 23.5cm^{-1} and 1372cm^{-1} respectively, so the difference between their measured transmission profiles indicate that the magnitude of the temperature cycle in the cell was considerably larger than that predicted by numerical model.

The channel spectra, evident in figure 7.14, have a free spectral range of 0.025 cm^{-1}. These probably arise from interference between two surfaces in the spectrometer, rather than from the cell, because the characteristic length (separation of the reflecting surfaces) to produce this fringe spacing is 20 cm.

One example of a sweep-integrated N_2O spectrum is given in figure 7.16. In this figure the middle curve around the strong line is the transmission calculated (dashed line) for the mean pressure, and the two extreme curves are the instantaneous transmission spectra,

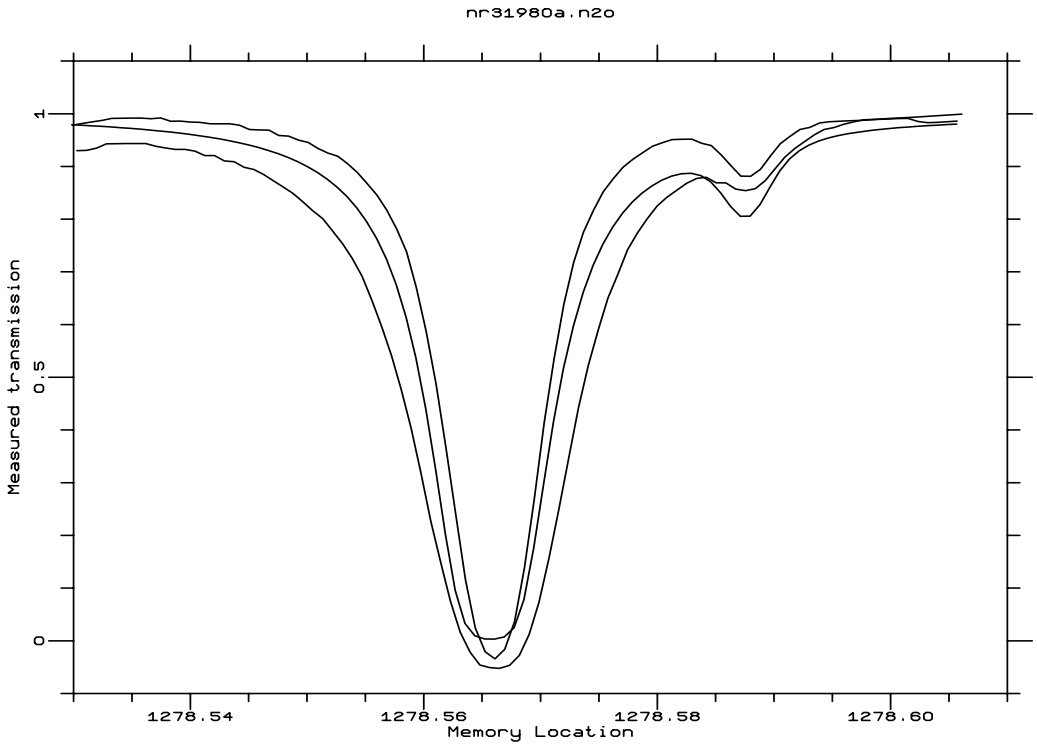

Figure 7.16: sweep-integrated spectrum of N_2O at pressure extremes (24.8mb, 2.33mm)

measured as described above, for the extremes of the pressure cycle. The modulation of the width of the line is evident; however, the absolute values of the transmission are uncertain because of the difficulties (discussed above) in ascribing full and zero scales to the raw measurements. As was subsequently the case for the water vapour sweep-integrated spectra, the need to fit the zero level prevented this being compared with the calculated pressure modulator spectra.

7.4 Results for water vapour

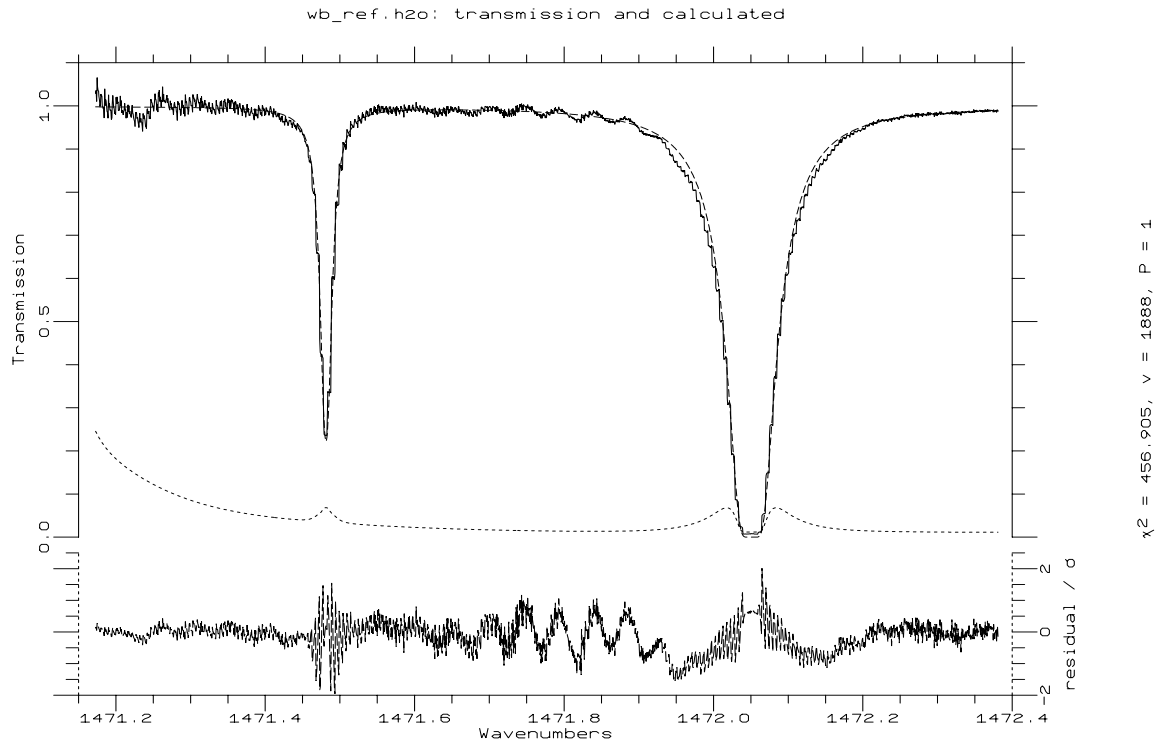

Figure 7.17: Measured (solid line) and calculated (dashed line) transmissions for the PMIRR flight model H_2O PMU, which contains 16.81mb of water vapour in a 10cm cell. The dotted line is the root sum squared transmission uncertainty of the calculated and measured transmissions combined.

The wideband transmission of the flight H_2O PMU is shown in figure 7.17. Measurements of the pressure modulated transmission were recorded for five modulation amplitudes, and the resulting measured spectra are shown in figures 7.18–7.22. Because of the small diode output power at the low frequency end of these measurements, the transmission is noisier at lower frequencies. These figures also clearly show the channel spectra described in chapter 6.

The measured and calculated spectra agree well, all having a fit probability (defined in chapter 5) of unity. The calculated profile is slightly too narrow, particularly at higher amplitudes, which, taken in conjunction with the pressure cycle measurements of chapter 5, indicates that the calculated cell pressure cycle is too small.

The brassboard measurements covered a wider range of pressures and amplitudes, and the resulting measured and calculated PMU transmission profiles are shown in figures 7.23–7.34. At 16.9 and 20.5mb, the agreement between the calculated and measured transmissions is good. At low pressures and low amplitudes, however, the SNR of the measurements is poor, and this causes uncertainty in the scaling, which leads to the poor agreement at 8.9mb. At 25mb, the underestimation of the weaker line at 1471.4817cm^{-1} relative to the stronger one at 1472.0512cm^{-1}, and the obvious width discrepancy between the calculated and measured transmissions indicates that the amplitude of the pressure cycle is underestimated by the model, in agreement with the pressure cycle measurements described in chapter 5.

The sweep-integrated spectra for water vapour were generally poor in quality, although

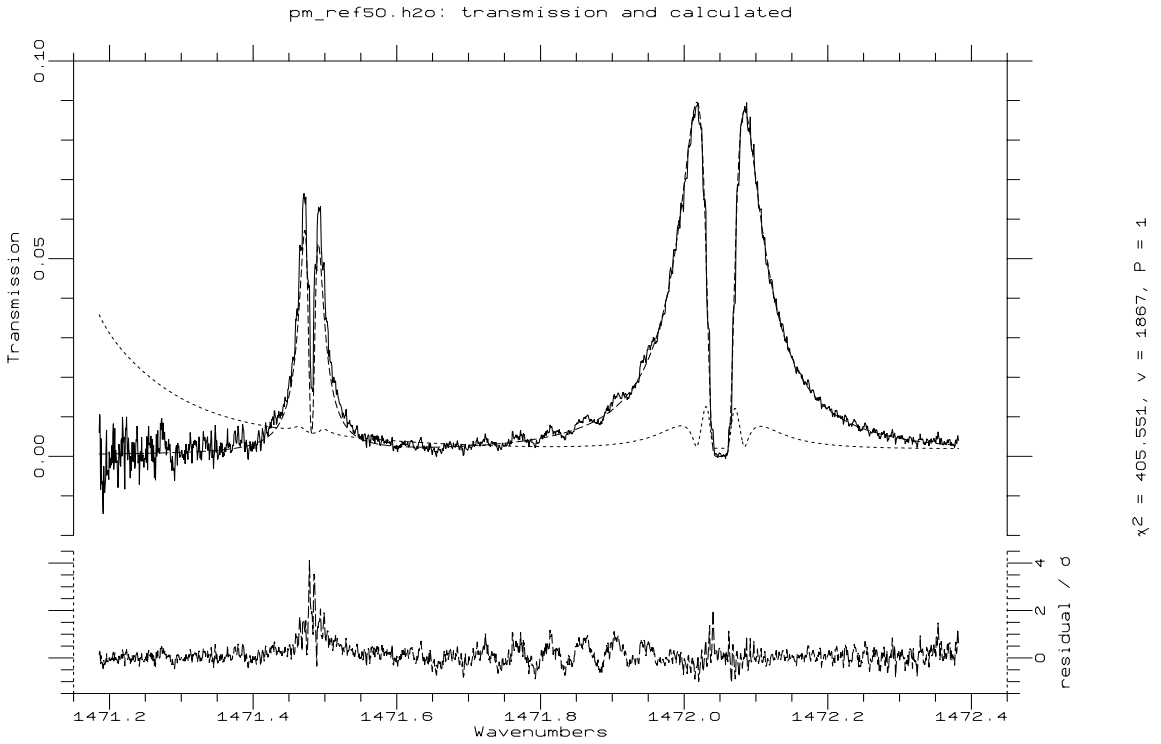

Figure 7.18: Measured (solid line) and calculated (dashed line) pressure modulated transmission of the flight model H_2O PMU (filled to 16.81mb) at 2.12 mm piston amplitude. The dotted line is the uncertainty in the transmission.

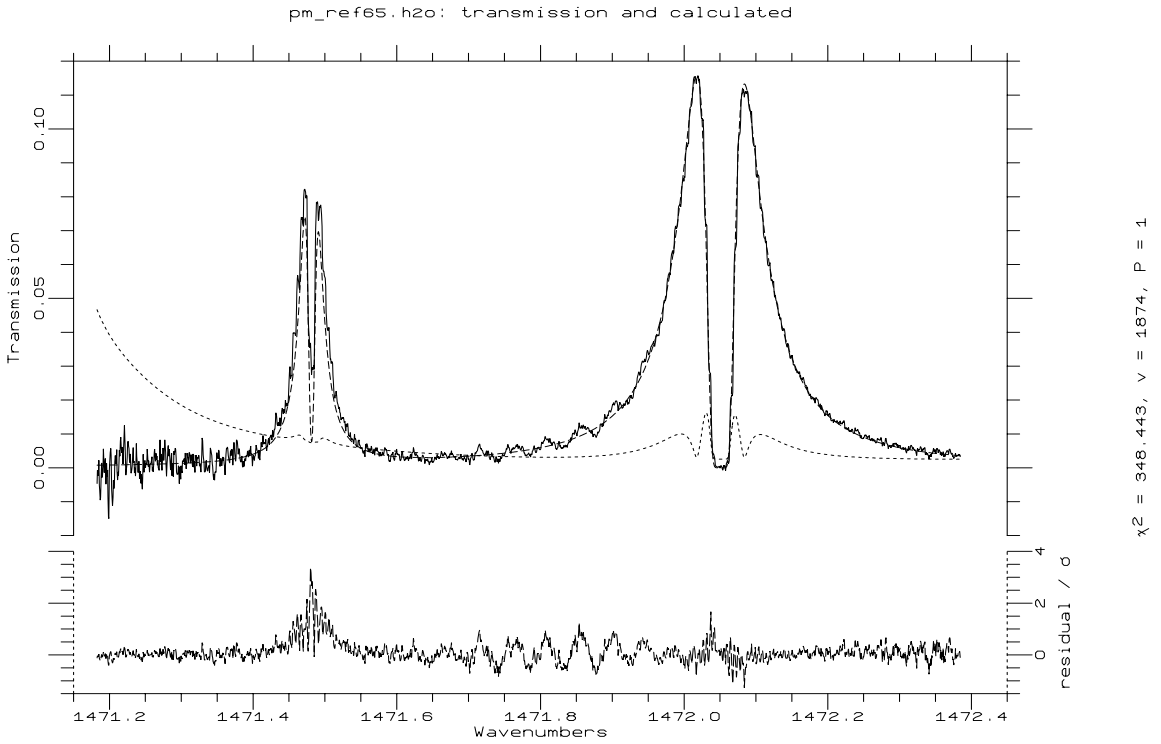

Figure 7.19: Measured (solid line) and calculated (dashed line) pressure modulated transmission of the flight model H_2O PMU (filled to 16.81mb) at 2.75 mm piston amplitude. The dotted line is the uncertainty in the transmission.

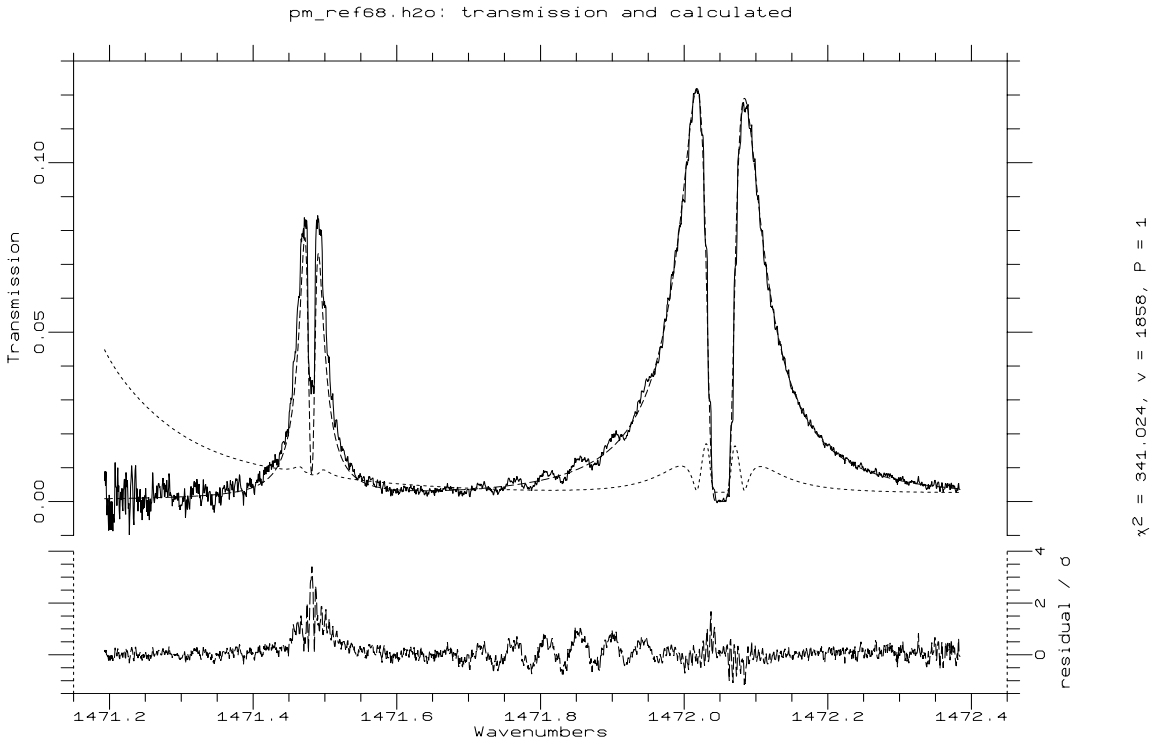

Figure 7.20: Measured (solid line) and calculated (dashed line) pressure modulated transmission of the flight model H_2O PMU (filled to 16.81mb) at 2.89 mm piston amplitude. The dotted line is the uncertainty in the transmission.

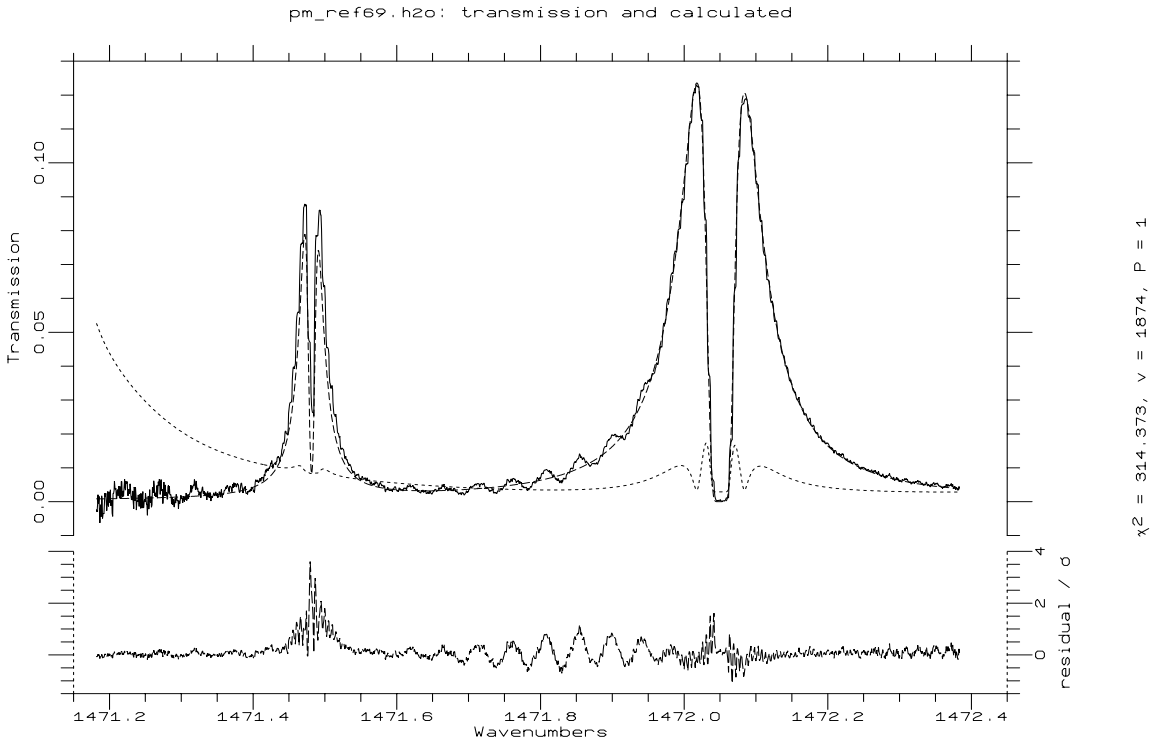

Figure 7.21: Measured (solid line) and calculated (dashed line) pressure modulated transmission of the flight model H_2O PMU (filled to 16.81mb) at 2.93 mm piston amplitude. The dotted line is the uncertainty in the transmission.

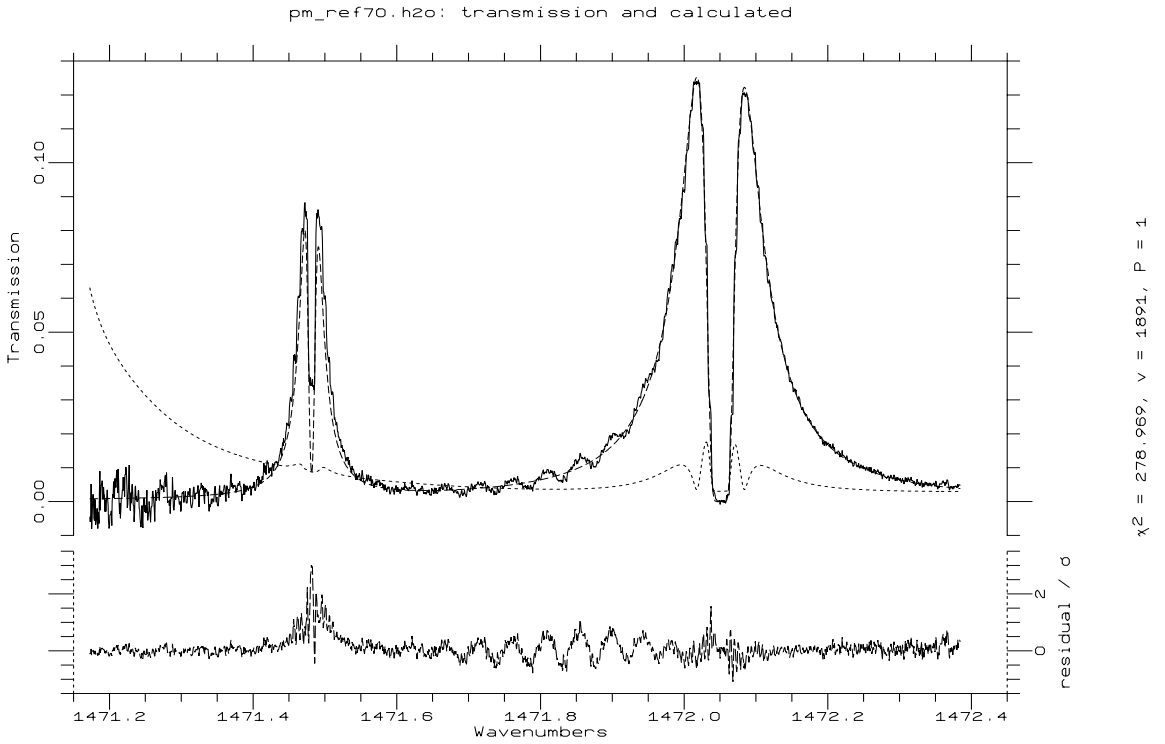

Figure 7.22: Measured (solid line) and calculated (dashed line) pressure modulated transmission of the flight model H_2O PMU (filled to 16.81mb) at 2.97 mm piston amplitude. The dotted line is the uncertainty in the transmission.

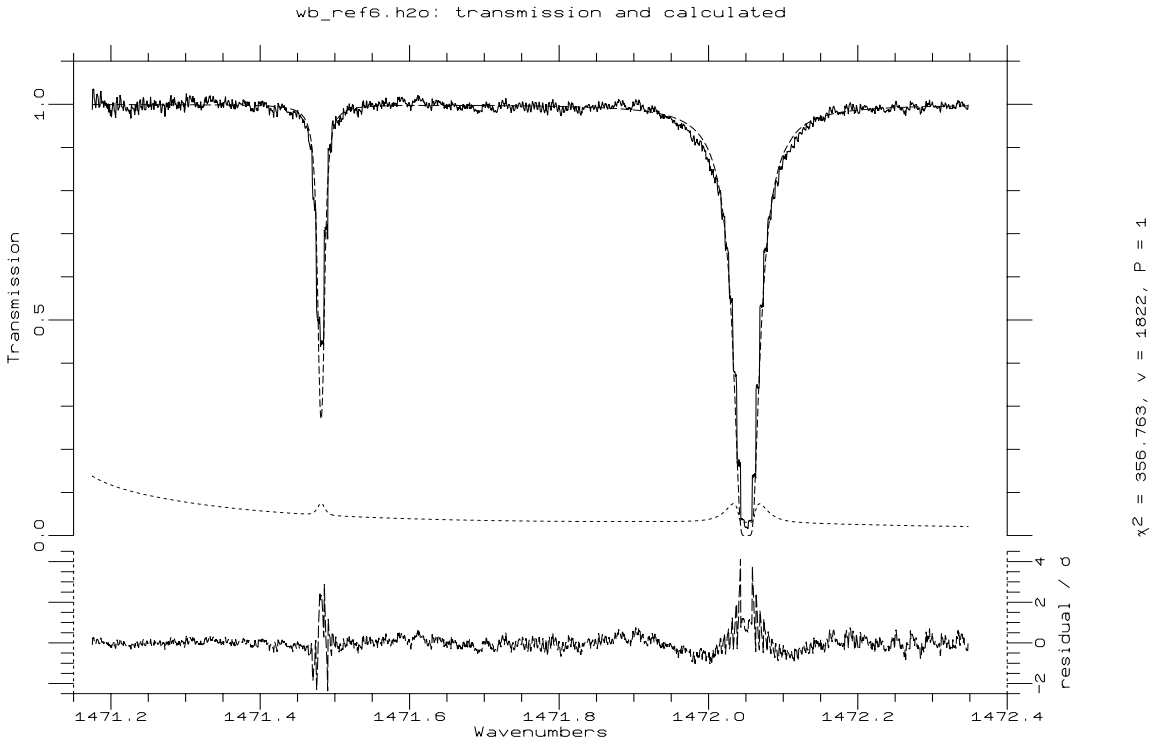

Figure 7.23: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU containing H_2O at 8.77 mb.

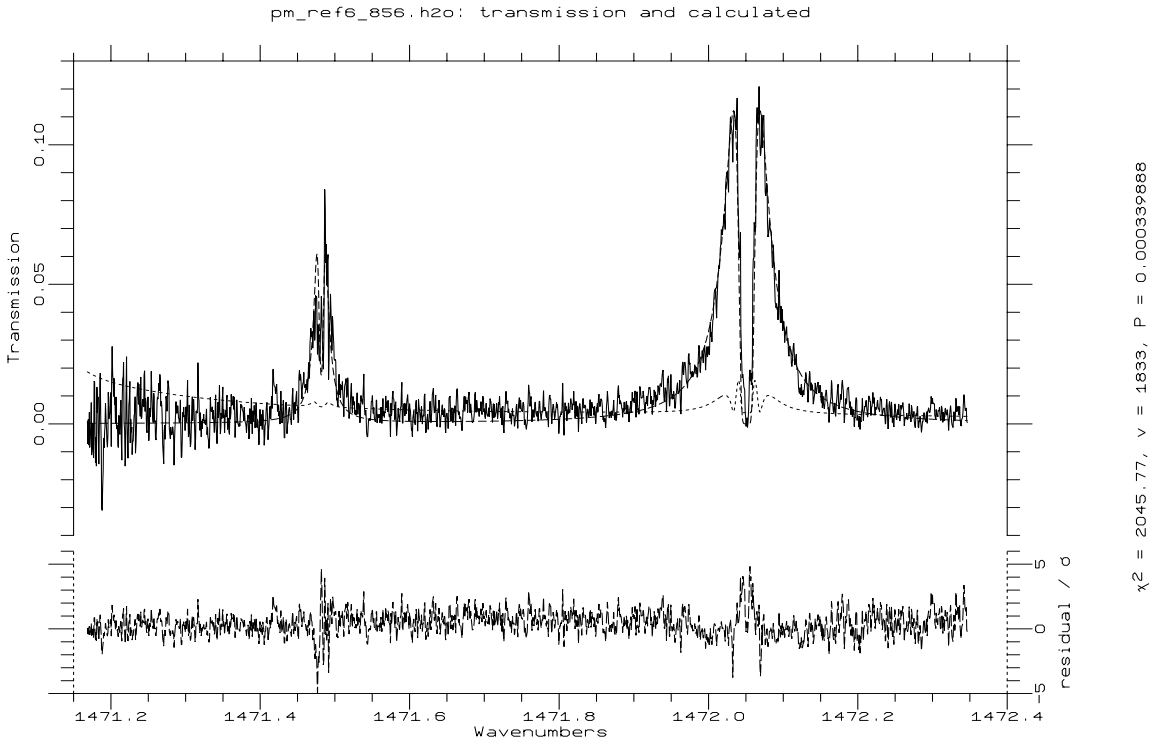

Figure 7.24: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU containing H_2O at 8.88 mb, with 2.64mm piston amplitude.

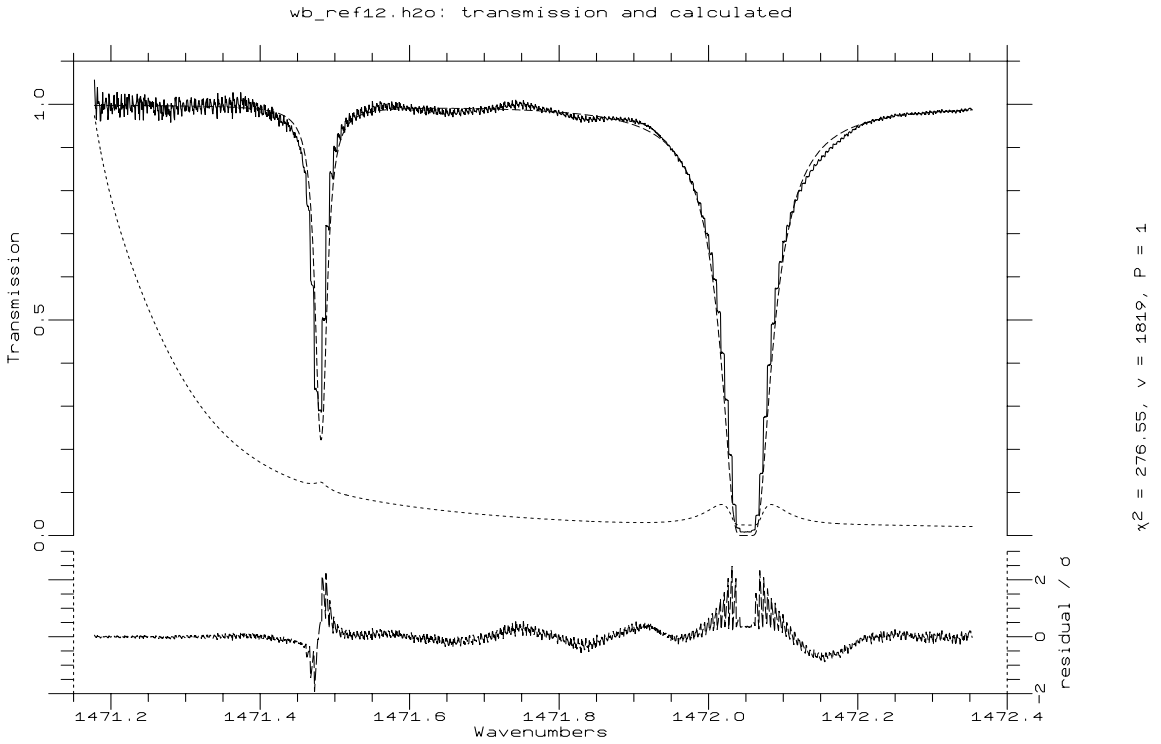

Figure 7.25: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 16.92 mb. The dotted line is the uncertainty in the transmission.

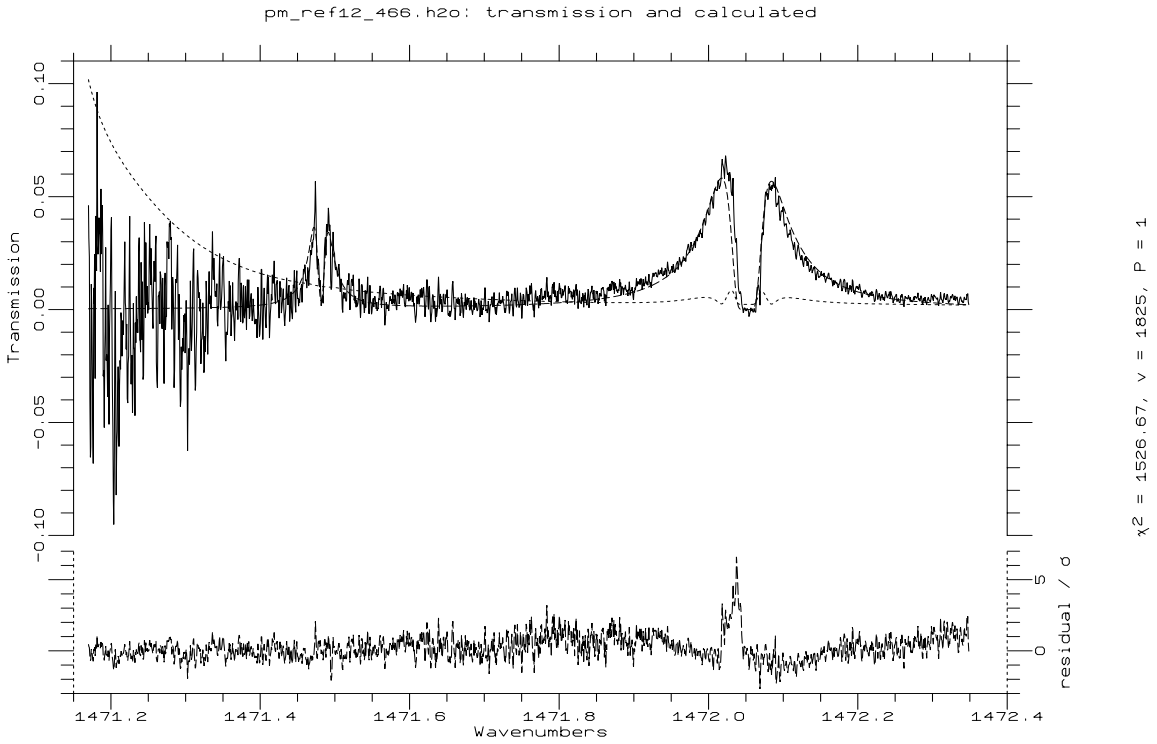

Figure 7.26: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 16.91 mb, 1.45mm piston amplitude. The dotted line is the uncertainty in the transmission.

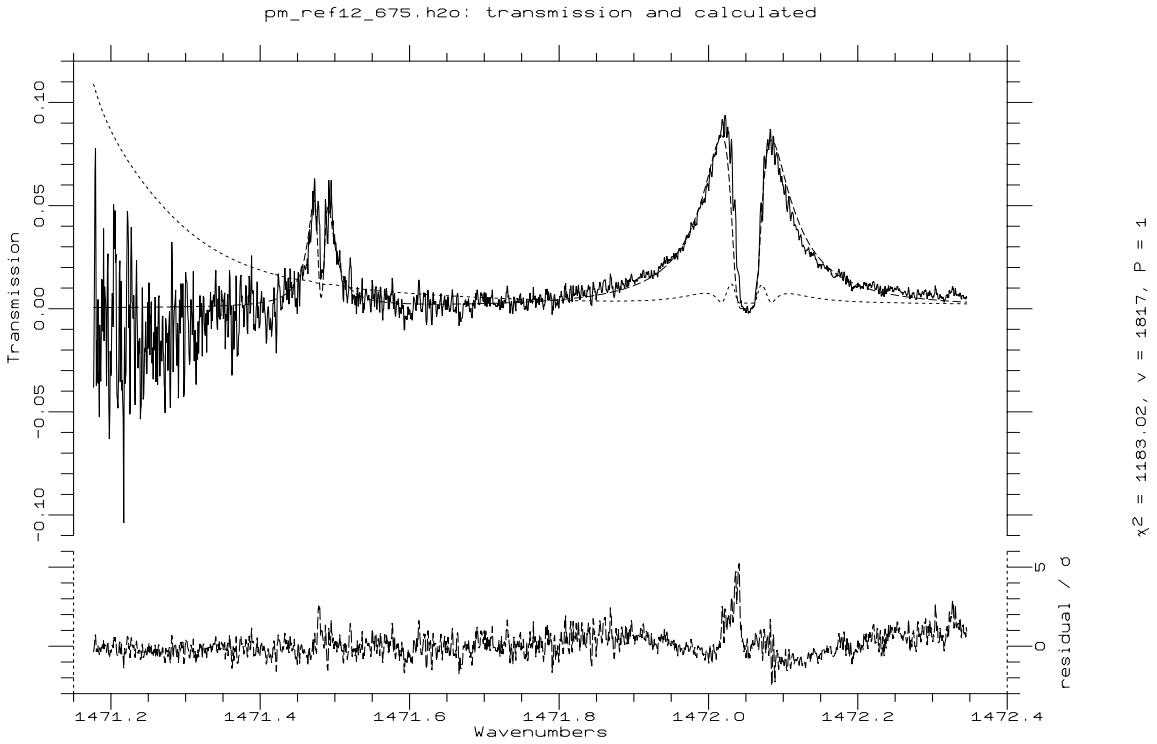

Figure 7.27: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 16.95 mb, 2.09mm piston amplitude. The dotted line is the uncertainty in the transmission.

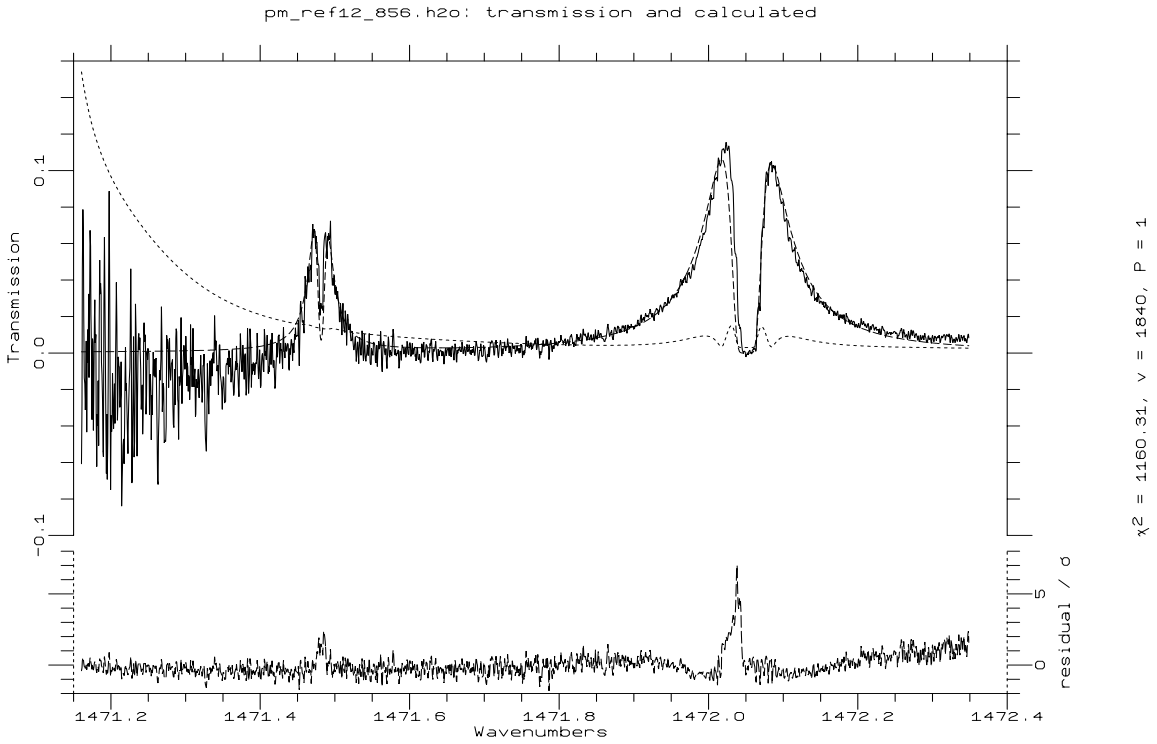

Figure 7.28: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 16.95 mb, 2.65mm piston amplitude. The dotted line is the uncertainty in the transmission.

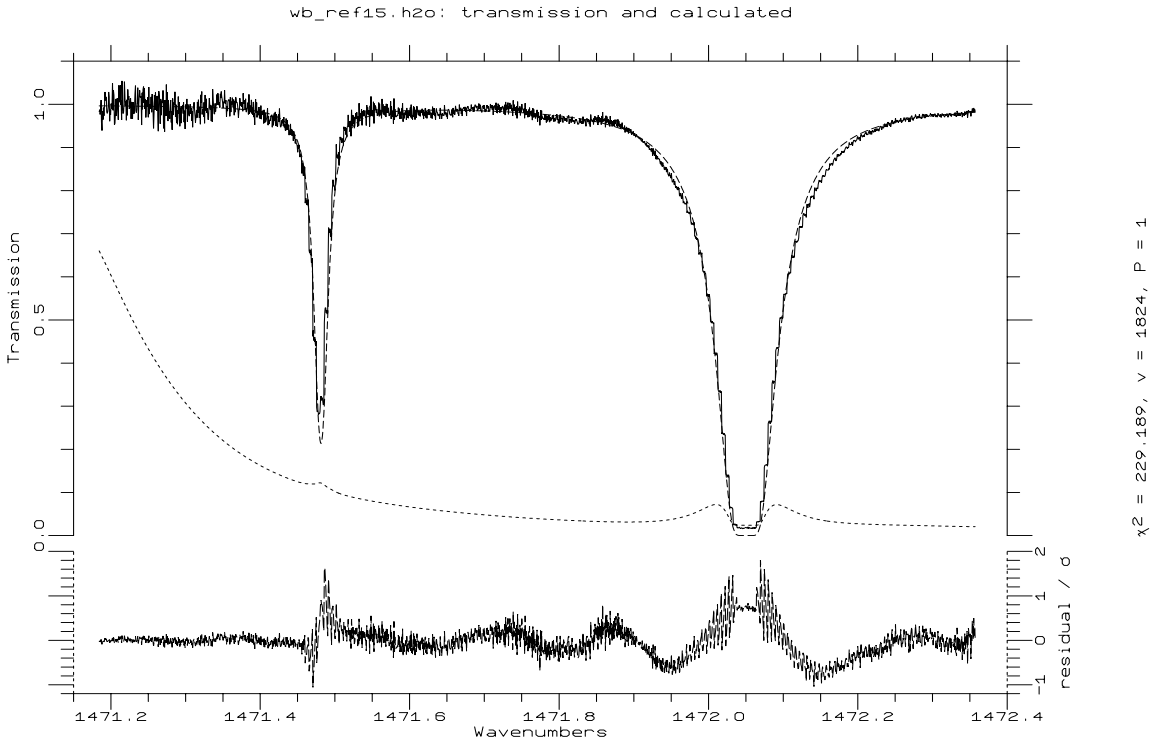

Figure 7.29: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 20.52 mb. The dotted line is the uncertainty in the transmission.

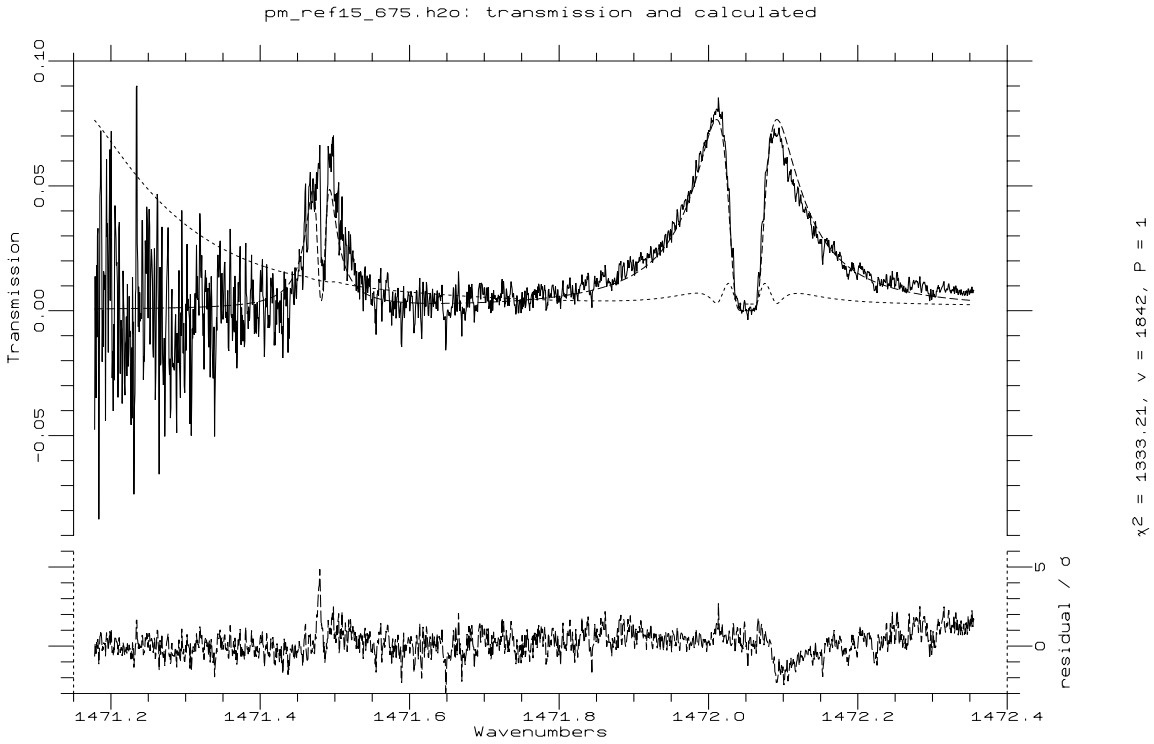

Figure 7.30: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 20.45 mb, 2.09 mm piston amplitude. The dotted line is the uncertainty in the transmission.

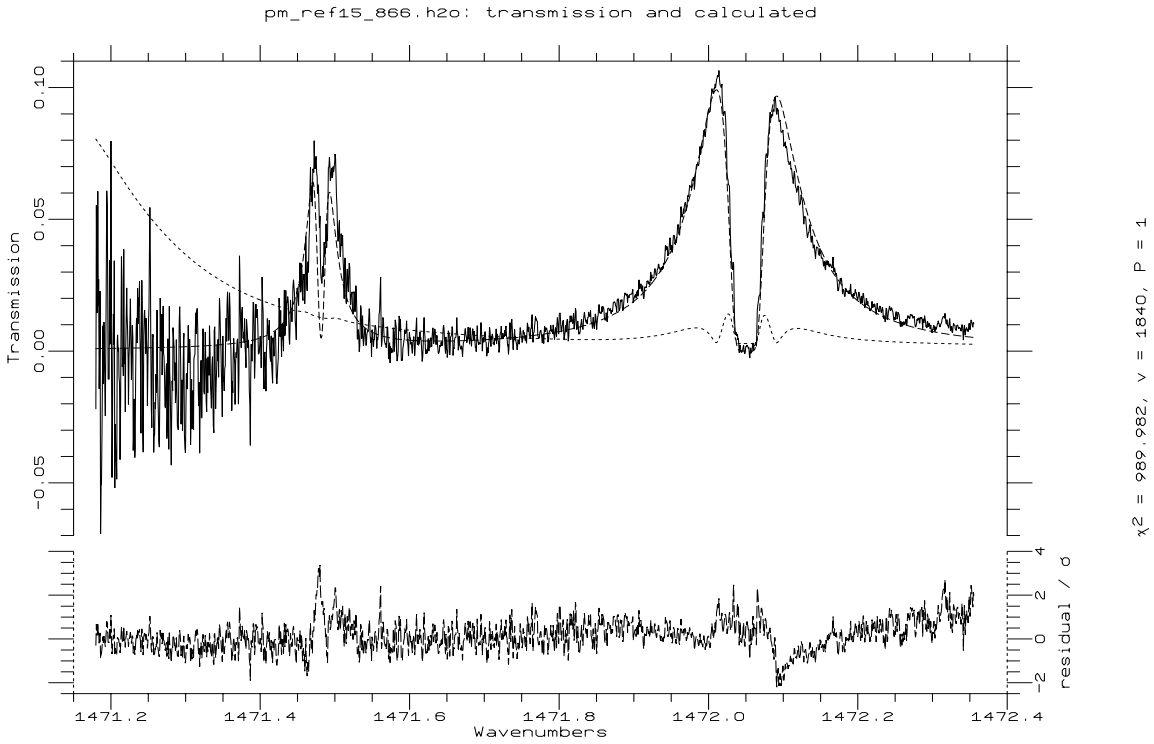

Figure 7.31: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 20.47 mb, 2.68 mm piston amplitude. The dotted line is the uncertainty in the transmission.

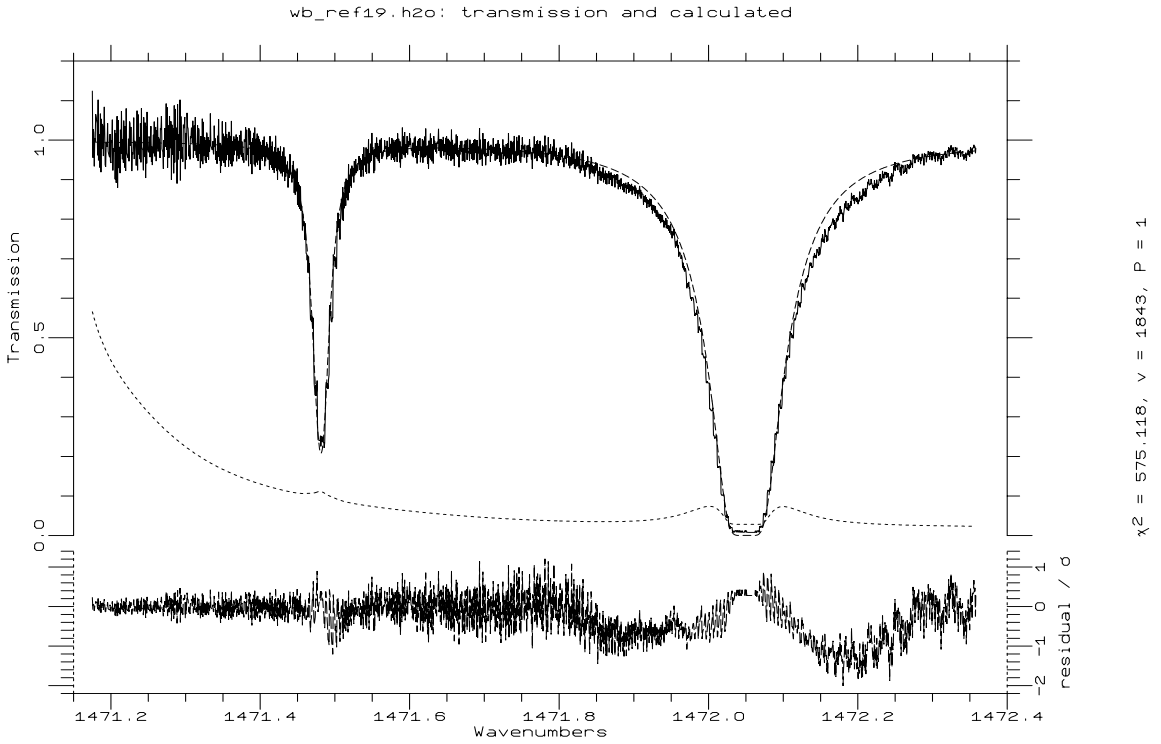

Figure 7.32: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 24.99 mb. The dotted line is the uncertainty in the transmission.

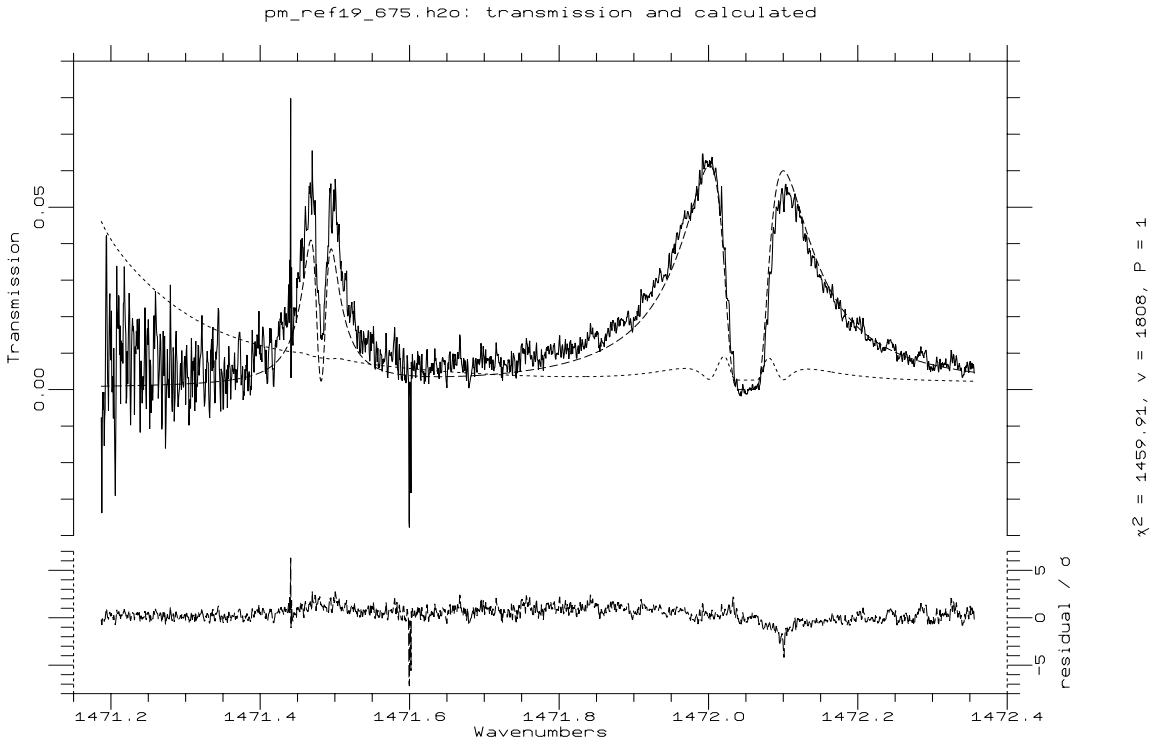

Figure 7.33: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 25.09 mb, 2.09 mm piston amplitude. The dotted line is the uncertainty in the transmission.

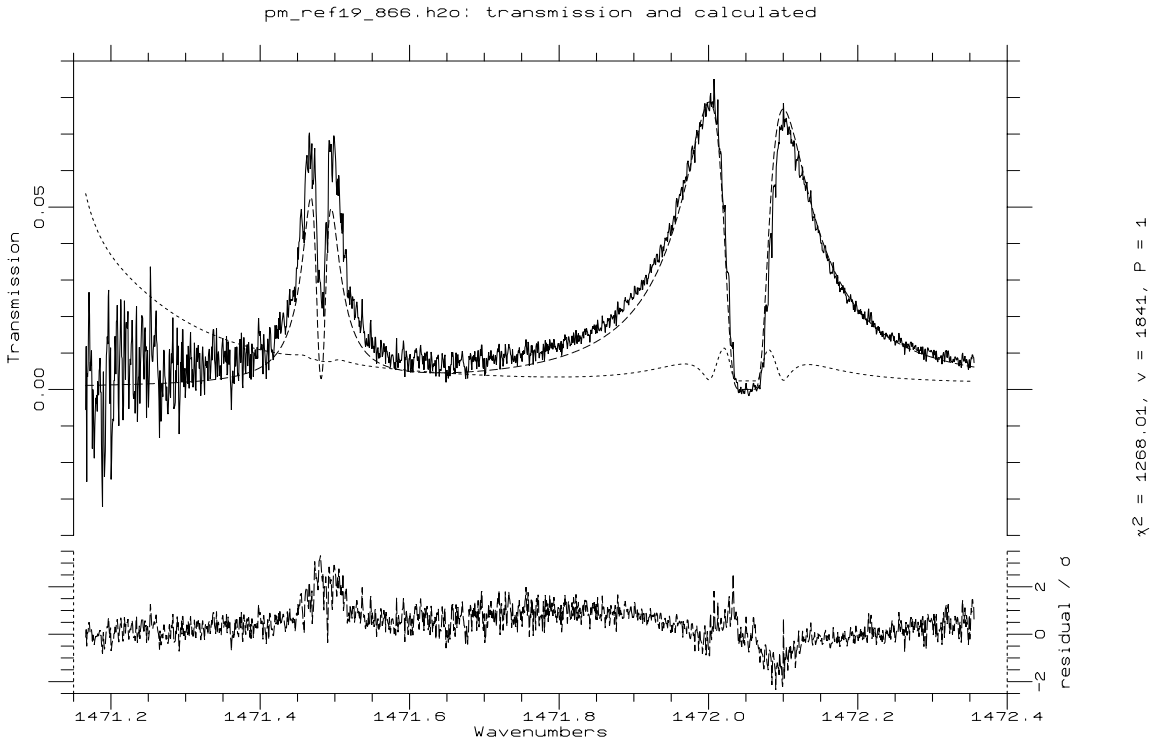

Figure 7.34: Measured (solid line) and calculated (dashed line) transmission of the brassboard PMU, containing H_2O at 25.13 mb, 2.68 mm piston amplitude. The dotted line is the uncertainty in the transmission.

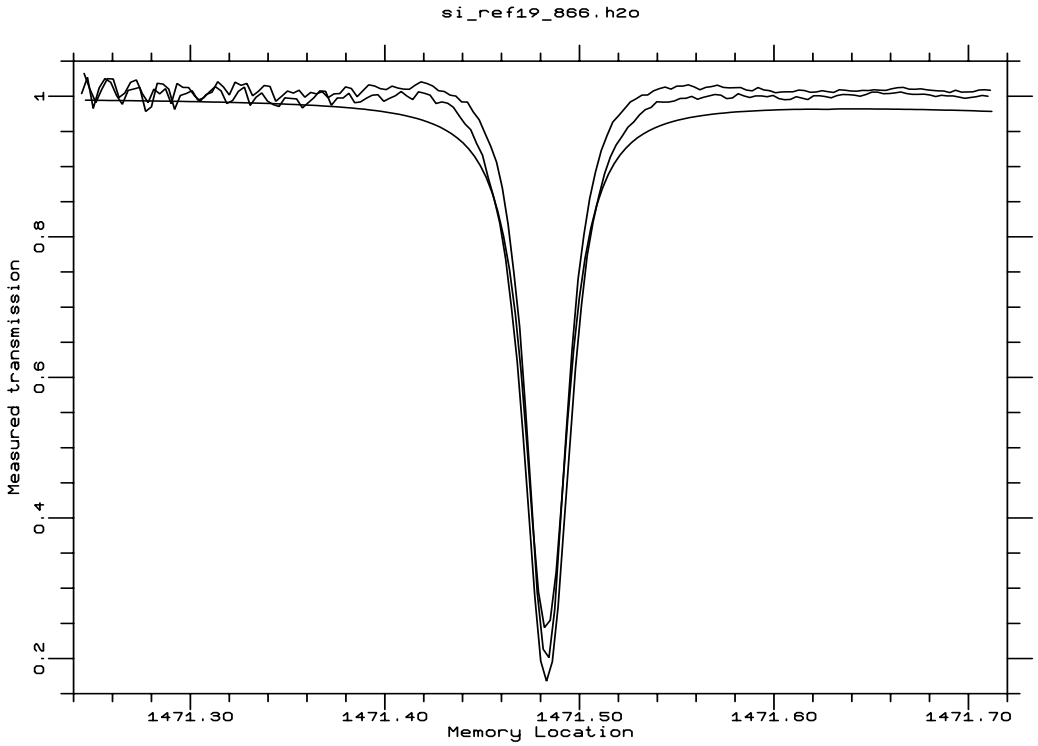

Figure 7.35: Sweep-integrated spectrum of H_2O at pressure extremes (25.5 mb, 2.68mm). The smooth curve is the calculated transmission at the mean pressure.

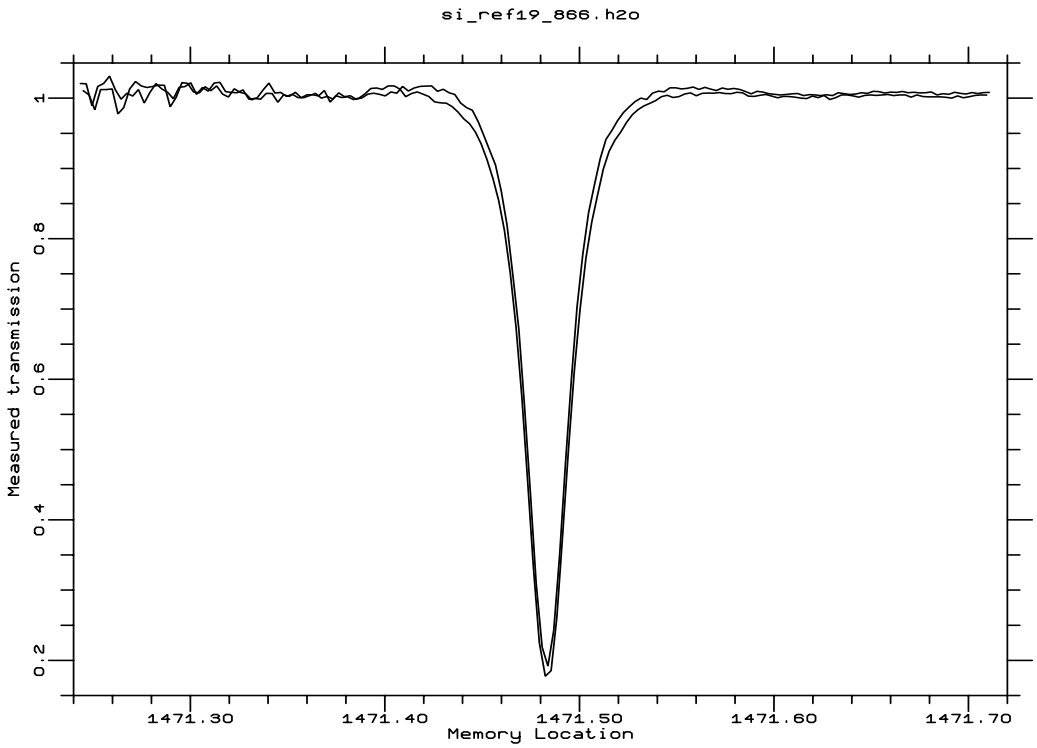

Figure 7.36: Sweep-integrated spectrum of H_2O at pressure coincidences (25.5mb, 2.68mm).

it is possible to observe the pressure cycle qualitatively from the widths of the spectral line around the cycle. Figure 7.35 shows the measured instantaneous line shape for the weak line at 1471.483 cm^{-1}, processed as described above. The extremes of the line width correspond to the extremes of the pressure cycle. The transmission calculated for the mean pressure and temperature is also included for comparison. Figure 7.36 shows the same line for the cycle points 90° from those in 7.35. The effect of the pressure cycle can be seen in the variation of the line width. Because of the difficulty in deriving the zero signal level, the agreement between the measured and calculated spectra in the wings of the lines is poor.

7.5 Summary

The analysis of pressure modulated transmission lineshapes measured using a tunable diode laser necessitates calculation of the transmissions from the predictions of the model. Using present-day computers, it is computationally possible to achieve this using sixteen points around the cycle, rather than two, as in previous calculations, is computationally feasible. The uncertainty in the calculated spectrum is limited by the accuracy of the line data far more than by that of the modelled pressure cycle.

The analysis also depends strongly on the ability to transform the raw signals into transmission as a function of frequency. In practice this requires knowledge of the zero level in the signal, and preferably of the full-scale level as well. Absence of the zero level prevented proper analysis of the sweep integrated measurements of both H_2O and N_2O, and while the full-scale pressure modulated transmission level can be inferred by scaling from measurements of wideband transmission, it is difficult to draw quantitative conclusions about

the thermodynamic cycling in the cell without this information. The transformation from signal averager memory location into frequency was facilitated in the case of H_2O by the simultaneous measurement of the lines at low pressure. Conversely, the transformation for N_2O was not satisfactory, as no low-pressure lines were available, and the cell lines exhibit large pressure shifts. This, and the absence of self-broadening coefficients from the line data, prevented quantitative comparisons between the calculated and measured transmissions.

Comparisons between the measured and calculated transmissions for the PMIRR H_2O modulator are encouraging. The agreement between the measured and calculated lineshapes is good, especially at moderate pressures (16–20mb). The agreement indicates that the model may validly be used to calculate the transmission of the PMIRR H_2O PMR channel. The observed asymmetry in the pressure modulated lineshapes of two strong water lines measured was broadly consistent with their pressure shift coefficients being 0.05 times their self-broadening coefficients. It seems probable, from comparisons of the measured and calculated line widths for water vapour, that the pressure cycle has been underestimated in the model, perhaps for the reasons discussed in chapter 5. The measurements made using the brassboard modulator also show good agreement with calculations at moderate pressures, but have poor SNR at low pressures and amplitudes, so the comparison is not as favourable. At 25mb, the modelled and measured line widths are noticeably different, indicating either that the line strengths given in the HITRAN database are too small, or that the pressure cycle predicted by the model has too small an amplitude. Since the latter is in agreement with the pressure cycle measurements described in chapter 5, it seems a more likely cause of the discrepancy.

Chapter 8:
Conclusions and Further Work

8.1 General

Pressure modulator radiometry offers the means to make important new measurements of the temperature and composition of the atmosphere of Mars. As part of a project to obtain such measurements, the PMIRR experiment on the forthcoming *Mars Observer* spacecraft, three areas have been addressed. A novel, reliable and efficient drive circuit for pressure modulators has been designed and built. It meets the demanding power requirements of the *Mars Observer* mission, and consequently has been used as the core of the PMIRR pressure modulator drive circuit. A steady-state numerical model of the PMIRR pressure modulators has been constructed. It simulates the mechanical and thermodynamic behaviour of the pressure modulators, and includes the first representation of adsorption. Pressure modulator lineshapes for water vapour and nitrous oxide have been measured under a variety of conditions, verifying that the spectroscopic behaviour of the PMIRR flight H_2O PMU conforms largely with expectations.

8.2 Drive circuit

The pressure modulator drive circuit worked well and required few modifications for flight model use; the most notable was J R Locke's change in the location of the current-sensing resistors to reduce the common-mode signal to which the current sensing is particularly sensitive. There is, however, a frequency dependence of the ratio of physical to demanded piston amplitude which should be removed. The drive circuit does not function well when the modulator is nearly empty, because the resonance in the PMU is then so narrow that the amplitude and phase control loops interact, causing instability, and because the small amount of current required to drive the modulator (because of the small damping) prevents an accurate current reference phase input into the PLL. This is not a problem for PMIRR, in which the modulators are run at constant pressure, but may require modification in a more general context. The circuit could be made more robust in the event of the pistons hitting, by sensing an over-amplitude on the two pistons individually. At present, since a piston collision causes the two sides of the modulator to run out of phase with each other, the differential amplitude sensed by the drive circuit decreases, so that the amplitude control loop increases the driving force, preventing a recovery.

8.3 Pressure Modulator Model

The pressure modulator model converges quickly to the PMU steady state. Gas flow, heat conduction, adsorption and condensation are all represented in it. The final state is generally independent of the initial settings, although slightly sensitive to the first choices of drive and DC correction level. The coordinates used, where one cycle of data is always

produced with frequency as an eigenvalue, are particularly convenient for calculation of the pressure modulated transmission. An adaptive timestep method is used to accelerates the calculations.

The inclusion of adsorption in the model has allowed the first successful prediction of PMU frequencies for water vapour, although the pressure cycle magnitude for the brassboard modulator at 298.4K is significantly underestimated. The theory should be corroborated by an extensive investigation into the operating frequency of a modulators by measurement at different pressures, amplitudes and body temperatures and with cells of different dimensions. Experiments to determine the sticking factors and latent heats of adsorption for CO_2 should also be performed.

For the purposes of this thesis, the modulator spring strengths have been derived indirectly, from least-squares fits to frequency measurements. It would have been better if all the strengths had been measured directly, and, by implication, such measurements should be made during the assembly of future modulators.

The temperature swings in the cell were expected to be large, because of the large radius of the cell and its correspondingly long thermal relaxation time. However, calculation of the detailed temperature structure inside each volume of the modulator, in addition to all the other dynamic processes, is at present computationally impractical. The simplest axial and radial mode was used, because the use of just one mode greatly simplifies the expression for the heat transferred to or from the walls of the cell and other volumes, and because it has the longest lifetime. Irwin [1991] considered the relaxation of 160 modes in the cell following an adiabatic temperature rise, in an otherwise analytical model. With the next generation of computers, it may be possible to include this more detailed representation. The model could also be improved by considering the finite thermal conductivity and thermal masses of the modulator body, pistons and cell, and by re-consideration of the representation of the tube, to include it as a separate volume, and considering adsorption on its walls.

The uncertainty in the spectroscopic predictions of the model is dominated by errors in the line strengths and widths given in the databases, because of the unique sensitivity of PMR transmission to lineshape. Future TDL measurements should be preceded by accurate measurements of all the line parameters for the lines used, especially the line strength, self-broadening coefficient, and pressure shift coefficient. The values of these coefficients should be of an accuracy sufficient to reduce the contribution of the line data to the uncertainty in the pressure modulated transmission to approximately the same magnitude as that of the modelled cycle. This required accuracy will vary with the modulator characteristics, pressure, amplitude and frequency, but for the flight H_2O PMU will be about 1%. This accuracy should be attainable from analysis of TDL measurements of the transmission in cells of different lengths and at different pressures.

The model predictions should also be compared with measurements of the modulator emission, which is possibly the best diagnostic of the temperature cycle, and with broadband transmission measurements. Such measurements have been made by Irwin [1991] using the brassboard modulator, and agreed within the expected uncertainty with his analytical model. However, conversion of the cycles of both emission and transmission into temperature are highly dependent upon the quality of the line data, and, because of the number of lines active in the broad spectral bandpass required for such measurements, line data of an accuracy sufficient to show disagreement cannot be measured using a tunable diode laser spectrometer, and are unlikely to be available for some time.

8.4 Spectroscopic measurements

High-resolution spectroscopic measurements of the spectral lineshapes produced by pressure modulators containing H_2O and N_2O have been made. The agreement of the measured and calculated transmissions for water vapour indicates that the pressure modulation technique may be applied to polar molecules like H_2O, resolving a long-standing uncertainty. The implications are that both the PMIRR and (differences in the cell dimensions notwithstanding) the ISAMS H_2O pressure modulator channels should provide useful results, which was to some extent under question after the difficulties of interpretation experienced by the SAMS H_2O channel. My attempts to measure the isotopic CO_2 lineshape were unsuccessful; in retrospect, more information would have been obtained using a normal (non-isotopic) band for CO_2. However, shortage of time prevented a change of diode, since almost all of the time available for the measurements had been spent investigating water vapour, the primary objective of this study.

All the measured pressure modulator lines were asymmetric, with peaks of different heights in the wings. This asymmetry was not an artifact of the measurement technique, and can be explained in terms of pressure shifts, with a coefficient approximately one twentieth of the self-broadening coefficient in the case of water. There were obvious channel spectra in the CO_2 PMU, and to a lesser extent in the H_2O PMU. These arose from multiple reflections between the cell windows, and were significantly larger than the absorptions measured for lines in PMIRR's isotopic CO_2 channel which is used for low-altitude temperature sounding. As a result they may impair the understanding of the results of that channel. Of course, the TDL beam is highly collimated, and the effect of the channel spectra is reduced as the beam angle increases (appendix D), so that the effect will be significantly smaller in the instrument because the beam converges to a focus within the CO_2 pressure modulator cell, although it is almost parallel in the H_2O cell.

The absolute magnitude of the pressure modulated transmission is uncertain, because of the difficulties in measuring the 'full scale' pressure modulated transmission, *i.e.*, with the cell alternating between fully-transparent and fully-opaque. It should be possible to simulate this using the slow chopper technique described in chapter 6, provided that the chopper phase is much more tightly controlled. Efforts should also be made to reduce the noise in the measurements, and where possible to remove any suspicion of detector non-linearity by using a photovoltaic detector. These, together with accurate data for the lines used (above), will provide a closer tolerance on the calculated quality of fit, revealing deficiencies in the model.

More work is required to perfect the sweep integration technique as applied to pressure modulators. In particular, more averager memory locations are necessary to increase the spectral resolution and allow the simultaneous use of several lines, and measurements of the signal with the laser diode switched off should be made to determine the PMR emission signal. The zero and full scale signal levels should also be determined, preferably removing the effect of ambient light by some other means than AC coupling. The result should be an independent confirmation of the full-scale pressure modulator signal.

Several more sets of measurements of pressure-modulated lineshape should also be made. An ordinary (non-isotopic) band of CO_2 should be measured in order to verify the performance of channels 1, 2 and 3 of PMIRR (this has already been done in part at JPL). Measurement of the pressure modulated transmission of NO_2 could also resolve some of the reasons why the measurements of NO_2 made by the Balloon-borne Pressure Modulator Radiometer (BPMR) as part of the 1982/1983 Balloon Intercomparison Campaign (BIC)

were so much at variance with measurements made by other instruments at the same time [Roscoe *et al.* 1990]. It may also be worthwhile to measure a few lines for each of the channels of future PMR instruments as part of the calibration process. Future tunable diode laser measurements should be taken in conjunction with model runs to simulate the experiments, allowing immediate derivation of the transmissions and comparison with predictions.

Appendix A:
Calculation of Optical Transmission

The fraction of incident light transmitted through or absorbed by a medium is, in general, a function of the frequency or wavelength of the light, the amount of absorber, the state of the absorber (its pressure, temperature *etc.*), the quantum mechanics of the absorbing material, and the intensity of the incident light if it is very large. In the case of molecular absorbers, the quantum mechanics are complex, with many possible transitions between states, which can, however, be labelled with quantum indices related to physical changes of behaviour.

To calculate the absorption or transmission it is seldom necessary to consider the quantum mechanics in detail, as lists of spectral parameters of absorption lines (such as HITRAN) are available which contain the information in a more usable form.

The transmission τ at frequency ν through a medium which exhibits absorption at a number of spectral lines is given by the Beer-Lambert law:

$$\tau(\nu) = \exp - \left(\sum_{lines} \chi_{line}(\nu) \right) \qquad (A.1)$$

in which the sum in the exponent is the *optical path* or *optical depth*. The contribution of each individual line to the total optical path is given by

$$\chi_{line}(\nu) = \int_{path} S_{line} f(\nu - \nu_{line}) dm \qquad (A.2)$$

where S is the line strength, f the line shape, ν_{line} is the centre frequency of the line, and dm the amount of absorber in a small part of the path. The line strength and shape are discussed below.

In low- and medium-resolution spectroscopy, where the instrumental linewidth of the spectrometer is similar to or larger than that of the absorbing path, the calculated (infinite-resolution) spectrum must be degraded (convolved with the instrumental lineshape) so that calculated and measured spectra can be compared. For my experiments, however, this was not necessary, as the TDL linewidth was about 30MHz (0.001 cm^{-1}) and the width of the water lines was about 0.05 cm^{-1}.

A continuum of absorption is also frequently described in spectroscopic literature. It is often estimated by calculating contributions from spectral lines far away in wavelength from the region of interest. On this basis, I have compared calculations of transmission in which all the lines in the database are included with restricted-range calculations to choose a cut-off range beyond which the contribution from the remaining outlying lines is negligible. This was about 5 cm^{-1} for the water vapour lines measured in the TDL experiment.

A.1 Line strength

Infrared spectral lines are associated with changes in the vibrational and rotational quantum numbers of molecules. The strength of an individual line depends on the probability of a transition between two states of the molecule, (the so-called 'matrix element'), and the population of the two states. For absorption the lower state population is by far the more important, and in the infrared this can be a strong function of temperature, as collision processes redistribute the population between the various possible energy levels according to the Boltzmann factor $\exp(-\varepsilon_i/k_BT)$ for the energy ε_i of the ith state at a temperature T. This results in a dependence of line strength on temperature, which is of importance in both the atmosphere, because of the differences in temperature at different heights, and in the pressure modulator, because of the temperature cycle in the cell.

A.1.1 Variation of linestrength with temperature

Tables of line parameters, such as HITRAN [Rothman *et al.* 1987] give the values of line strength for each line at a reference temperature T_r, usually 296 K. The line strength is proportional to the population of the lower state. This is calculated from the ratio of the Boltzmann factor for that state to the sum of the Boltzmann factors for all states, which is the partition function Q.

The partition function for a vibrating rotor may be split into rotational and vibrational parts. The rotational partition function Q_r is proportional to the temperature T raised to a power m which is 1 for linear molecules (e.g. CO_2 and N_2O) and 1.5 for non-linear molecules (*e.g.*, H_2O and CH_4) in the classical limit. The vibrational partition function for equally spaced (at ν') non-degenerate vibrational energy levels is proportional to $1 - \exp(-hc\nu'/k_BT)$.

For the gases I have been using, simple but reasonably accurate approximations to the total partition functions exist:

$$Q_{CO_2}(T) = T(0.9700672 - T(0.7177665 \times 10^{-3} - 0.2766523 \times 10^{-5}T)) \qquad (A.3)$$

$$Q_{N_2O}(T) = 109.402 + T(0.470896 + 0.348261 \times 10^{-2}T) \qquad (A.4)$$

$$Q_{H_2O}(T) = T^{\frac{3}{2}} \qquad (A.5)$$

Equation A.3 is J.Eyre's cubic fit (reported by Muggeridge [1986]) to calculations by Gray and Young [1969] and equation A.4 is due to Jones [1983]. In equation A.5, the vibrational partition function is assumed to be unity.

The N_2O expression agrees excellently over a wide temperature range with values published by McClatchey *et al.* [1973]. The CO_2 values are not as satisfactory but the discrepancy of a few percent at 100K difference from the reference temperature is swamped by the uncertainty in the line strengths in the database, particularly for the isotopic CO_2 band used in channel 1 of PMIRR (see chapter 2).

On the basis of the above, the ratio of the line strength at a temperature T to that at a reference temperature T_r can be written:

$$\frac{S(T)}{S(T_r)} = \frac{Q(T_r)}{Q(T)} \frac{\exp(-E''/k_BT)}{\exp(-E''/k_BT_r)} \frac{1 - \exp(-hc\nu_0/k_BT)}{1 - \exp(-hc\nu_0/k_BT_r)} \qquad (A.6)$$

where ν_0 is the frequency of the line. The last term is to correct for the small amount of stimulated emission which reduces the effective absorption.

A.2 Lineshape

The lineshape of individual spectral lines of a molecule may be calculated according to a simple model involving collisional termination of emitted wave trains and Doppler shifting due to the velocity distribution of the molecules. In general it is a function of the difference (the 'detuning') between the frequency and that of the centre of the line, and has a maximum at line centre, with values tending to zero in the 'wings', at large detuning.

The theory of pressure-broadening using random-length wave trains interrupted by collisions leads to the prediction that the width of the spectral lines is proportional to the collision rate, and therefore to the pressure. The collisional processes predict a Lorentz function for the lineshape:

$$f_L(\nu - \nu_0) = \frac{\alpha_L}{\pi} \frac{1}{(\nu - \nu_0)^2 + \alpha_L^2} \tag{A.7}$$

where α_L is known as the *pressure-broadening coefficient* or *width*, proportional to the pressure, of order 0.1 $\mathrm{cm}^{-1}/\mathrm{atm}$, and to a power of the temperature which is usually near 0.5. The effect of the Doppler shifts, averaged over the Maxwell-Boltzmann velocity distribution, is described by a Gaussian lineshape:

$$f_D(\nu - \nu_0) = \frac{1}{\alpha_D \sqrt{\pi}} \exp\left(-\left(\frac{\nu - \nu_0}{\alpha_D}\right)^2\right) \tag{A.8}$$

where the *Doppler-broadening coefficient*,

$$\alpha_D \approx \beta \nu_0 \sqrt{\frac{T}{M}}, \tag{A.9}$$

in which

$$\beta \approx 4.3 \times 10^{-7} \mathrm{K}^{-\frac{1}{2}} \tag{A.10}$$

In the pressure régimes found in PMUs, both pressure and Doppler broadening are significant; the balance of the two varies from line to line. It is customary, and for most purposes adequate, to use the Voigt function, which results from the convolution of the Doppler and Lorentz profiles, to describe observed line shapes. The wings of the Doppler line profile tend to zero faster with increasing detuning than those of the Lorentz profile, so the wings of the Voigt lineshape are Lorentzian, while the centres are dominated by the Doppler component. I have used Reichel's algorithm [Chiarella and Reichel 1968] to evaluate the profile, because of its high speed.

Measured lineshapes frequently differ in the wings from those predicted with the Voigt profile, giving so-called sub- or super-Lorentzian behaviour, according as the absorption is less or greater than that predicted. However, as the dependence of the absorption on pressure is small in the wings, where $\nu - \nu_0 \gg \alpha_L$ this effect can be neglected for pressure-modulated lines, because of the exponential dependence of transmission on optical path.

A.2.1 Broadening coefficients

The magnitude of the pressure-broadening coefficient α_L in equation A.7 varies from line to line and has different values for different colliding molecules. It is also temperature-dependent, according to a power law. Since in the pressure modulator the gas is essentially pure, the self-broadening coefficient is most important. The temperature swings are small compared to the absolute temperature, so that the effect of the temperature dependence of the self-broadening coefficients is small.

For water lines in the databases for which only the air- or nitrogen-broadening coefficients are given (true for all lines in HITRAN 86), it is usual to assume that the self-broadening coefficient is four or five times the air-broadening coefficient. The ration of self- to air-broadening coefficient is smaller for other gases.

A.3 Pressure shifts

More advanced theory also predicts that the lines' centre frequencies are linearly dependent on the pressure, with a coefficient which depends upon the perturbing molecule, as is the case for pressure broadening. The line centre frequency may increase or decrease with increasing pressure. However, very few pressure-shift coefficients have been measured or published.

Pressure shifts for water vapour in the 720 nm wavelength region have been measured by Grossmann and Browell [1989a, 1989b]. The coefficients they obtained lie in the range of -0.079 to +0.053 cm^{-1}/atm. They conclude that while there is no apparent relationship between the pressure shift coefficient and the broadening coefficient or the vibrational state for pure water vapour, there is a correlation with both in the foreign-perturber case. The shift is temperature-dependent, with an exponent of 1.57. The shifts they measured are in marked disagreement with the calculations of Gamache and Davies [1983]. Liebe *et al.* [1969] and Bauer *et al.* [1987] have made measurements of pressure shifts for two lines in the microwave region, reporting shift coefficients also widely at variance with these calculations. My conclusion is that the theory of pressure shifting is not yet sufficiently developed for useful application to spectral calculations.

Appendix B:
Analytical model of the pressure modulator

To understand the behaviour of the PMU system, and in particular to design the drive circuit, I used a simple analytical mechanical model in which the system was represented by two masses connected by three springs to each other and to fixed ends. Such a model has been used before [Venters 1991] to derive the effect of spring strength mismatch on momentum imbalance and to predict the operating frequencies of balanced pressure modulators. The PMU–drive circuit system is treated as a second-order forced oscillator, neglecting effects such as the non-linearity of the 'spring' term arising both from non-linear behaviour of the springs (chapter 3) and compression of the gas. The frictional force is assumed to be proportional to the velocity, and only the differential mode is considered. The force provided by the drive coils is assumed to be proportional to the current through them.

B.1 Complementary function

The equation of motion of the PMU pistons when combined to give the differential mode is

$$m\frac{d^2\mathbf{x}}{dt^2} + \lambda\frac{d\mathbf{x}}{dt} + k\mathbf{x} = \mathbf{F_1}\exp i\omega t \qquad (B.1)$$

where $\mathbf{x}$ is the (complex) position of the piston, $\mathbf{F_1}$ is the (complex) magnitude of the driving force, m is the effective moving mass, λ is the first-order frictional coefficient, and k is the force constant of the effective restoring spring, which includes contributions from the suspension springs, coupling spring and gas. The complementary function of B.1 is

$$\mathbf{x} = \mathbf{x}_{CF}\exp\left(\frac{-\lambda t}{2m}\right)\ \exp i\left(\frac{\sqrt{4km-\lambda^2}}{2m}\ t\right) \qquad (B.2)$$

where $\mathbf{x}_{CF}$ is an amplitude at time $t = 0$. The complementary function frequency is

$$\omega_{CF} = \sqrt{\frac{4km-\lambda^2}{4m^2}} \qquad (B.3)$$

When there is no damping, this takes the familiar form:

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad (B.4)$$

B.2 Magnification Factor (Q)

The total energy E of the system is equal to the elastic energy when the piston is stationary. If x_{CF} is real, the kinetic energy at $t = 0$ is zero, as it is after one cycle, when $t = 2\pi/\omega_{CF}$, so that

$$E_{(t=0)} = \tfrac{1}{2}kx_{CF}^2 \tag{B.5}$$

$$E_{(t=\frac{2\pi}{\omega})} = \tfrac{1}{2}kx_{CF}^2 \exp\left(\frac{-2\pi\lambda}{\sqrt{4km-\lambda^2}}\right) \tag{B.6}$$

The energy lost per cycle at the complementary function frequency ω_{CF} is the total energy divided by the magnification factor Q:

$$\frac{E}{Q} = \Delta E = \tfrac{1}{2}kx_{CF}^2\left(1 - \exp\left(\frac{-2\pi\lambda}{\sqrt{4km-\lambda^2}}\right)\right) \tag{B.7}$$

from which it can be shown that

$$\left(1 - \frac{1}{Q}\right) = \exp\left(\frac{-2\pi\lambda}{\sqrt{4km-\lambda^2}}\right) \tag{B.8}$$

For convenience, I shall let

$$\beta = \frac{1}{2\pi}\ln\left(\frac{Q}{Q-1}\right) = \frac{\lambda}{\sqrt{4km-\lambda^2}} \tag{B.9}$$

β is physically significant in that the energy of the system drops by a factor $\exp(-\beta)$ in each radian. It thus relates the conceptually familiar concept of Q with the more physically useful but obscure damping factor λ, and spring constant k. From equations B.3 and B.9,

$$\lambda = 2\beta m\omega_{CF} \tag{B.10}$$

$$k = m\omega_{CF}^2(1+\beta^2) \tag{B.11}$$

$$\omega_{CF} = \frac{\beta}{2m\lambda} \tag{B.12}$$

and

$$\omega\lambda = 2\omega\omega_{CF}m\beta \tag{B.13}$$

B.3 Particular Integral

The particular integral of B.1 is $\mathbf{x} = \mathbf{x_1}\exp i\omega t$, and back-substitution implies

$$(-\omega^2 m + i\omega\lambda + k)\,\mathbf{x_1}\exp i\omega t = \mathbf{F_1}\exp i\omega t \tag{B.14}$$

Substitution from equations B.10–B.13 yields the following:

$$\frac{\mathbf{F_1}}{\mathbf{x_1}} = -\omega^2 m + i\omega\lambda + k = -\omega^2 m + 2i\omega\omega_{CF}m\beta + m\omega_{CF}^2(1+\beta^2) \tag{B.15}$$

In chapter 3 it was shown that the power consumption is proportional to the square of the driving force, so the force to amplitude ratio should be minimized in the drive circuit to provide the most power-efficient running conditions.

$$\left|\frac{\mathbf{F}_1}{m\mathbf{x}_1}\right| = \left|\omega_{CF}^2 - \omega^2 + \omega_{CF}^2\beta^2 + i2\omega\omega_{CF}\beta\right| \tag{B.16}$$

$$= \sqrt{\left(\left(\omega_{CF}^2(1+\beta^2) - \omega^2\right)^2 + 4\omega^2\omega_{CF}^2\beta^2\right)} \tag{B.17}$$

$$= \sqrt{\left(\omega_{CF}^4(1+\beta^2)^2 + \omega^4 + 2\omega^2\omega_{CF}^2(\beta^2-1)\right)} \tag{B.18}$$

At any given frequency ω, the tangent ε of the phase difference ϕ is (from B.16)

$$\varepsilon = \frac{2\omega\omega_{CF}\beta}{\omega_{CF}^2(1+\beta^2) - \omega^2} \tag{B.19}$$

B.4 Power minimization

The most efficient frequency is that for which the force for a given amplitude is a minimum. Differentiating its square of equation B.18 with respect to ω^2:

$$\frac{d}{d\omega^2}\left|\frac{\mathbf{F}_1}{m\mathbf{x}_1}\right|^2 = 2\omega^2 + 2\omega_{CF}^2\left(\beta^2 - 1\right) \tag{B.20}$$

so the most efficient frequency

$$\omega_{ME} = \omega_{CF}\sqrt{(1-\beta^2)} \tag{B.21}$$

For large Q, when β is very small, therefore, $\omega_{ME} \approx \omega_{CF} \approx \omega_0$ for the most efficient operation. The frequency ω_{ME} is achieved by servoing the phase lag between the PPO position ($\mathbf{x}$) and the drive current ($\mathbf{F}$). The phase shift ε_{ME} at the most efficient frequency ω_{ME}, substituting from B.21 and B.19, is

$$\varepsilon_{ME} = \frac{2\omega_{CF}^2\beta\sqrt{1-\beta^2}}{\omega_{CF}^2(1+\beta^2) - \omega_{CF}^2(1-\beta^2)} \tag{B.22}$$

$$= \frac{\sqrt{1-\beta^2}}{\beta} \tag{B.23}$$

which is independent of ω_{CF} and ω_{ME}. From consideration of a triangle with perpendicular sides of lengths 1 and β, it is evident that

$$\beta = \cos\phi_{ME} \tag{B.24}$$

and so for large Q (and small β), $\phi_{ME} = 90°$, as expected.

B.5 Relationship of frequency and power to phase

The frequency of the system for a given phase shift may be obtained from equation B.19 by solving the quadratic equation for ω:

$$\frac{\omega}{\omega_{CF}} = \frac{-\beta \pm \sqrt{\beta^2 + \varepsilon^2(1+\beta^2)^2}}{\varepsilon} \qquad (B.25)$$

From B.11 and B.4,

$$\frac{\omega_{CF}}{\omega_0} = \frac{1}{\sqrt{1+\beta^2}} \qquad (B.26)$$

These equations give ω in terms of the damping-independent ω_0. The sign of the addition, giving the two values of ω in B.25, depends on whether the phase is less than or greater than 90°, to give a positive frequency. The power consumption is calculated from substitution of B.25 and B.26 into B.18. At the most efficient frequency, substituting B.21 into B.18 gives the power consumption

$$\left| \frac{\mathbf{F}_1}{m\mathbf{x}_1} \right|^2 = 4\beta^2 \qquad (B.27)$$

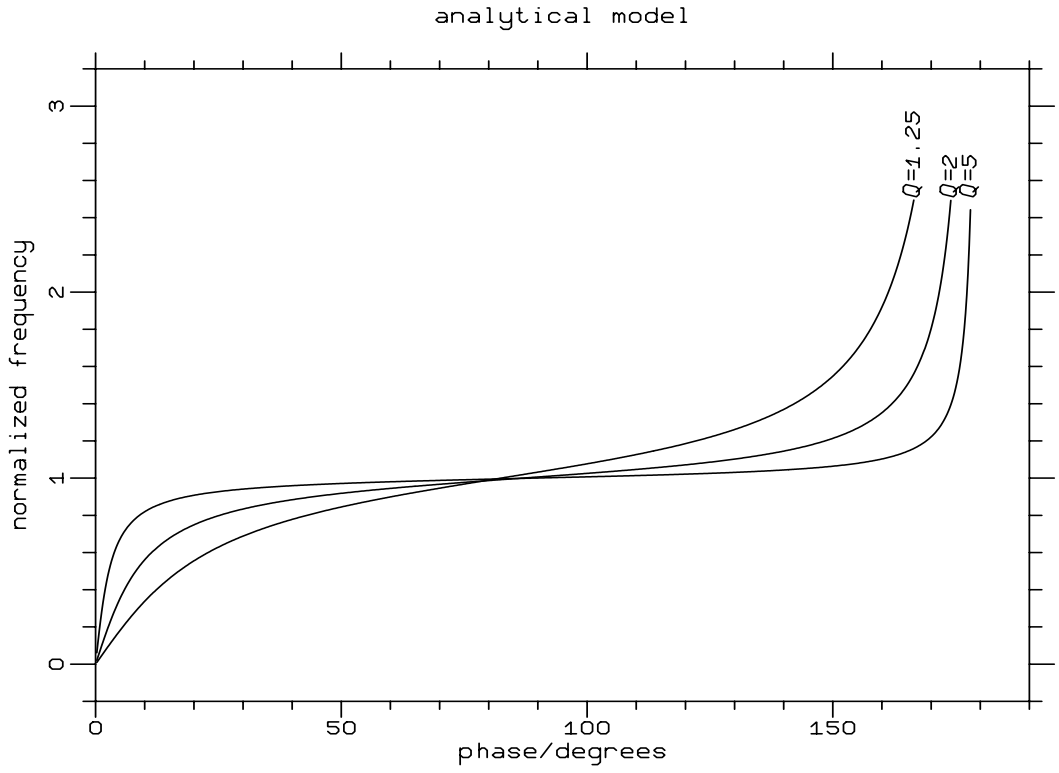

Figure B.1: Frequency / phase relationship calculated for three values of the magnification factor Q. The frequency is in units of ω_0, the undamped resonant frequency. As the damping increases, the dependence of frequency on phase increases.

Figure B.1 shows the frequency/phase relationship for values of Q of 1.25, 2 and 5, and figure B.2 shows the corresponding power consumption. For likely values of Q (greater than 1.1) the frequency/phase relationship is quite flat around 90° phase shift, so that the system frequency is to a certain extent independent of the phase setting around that

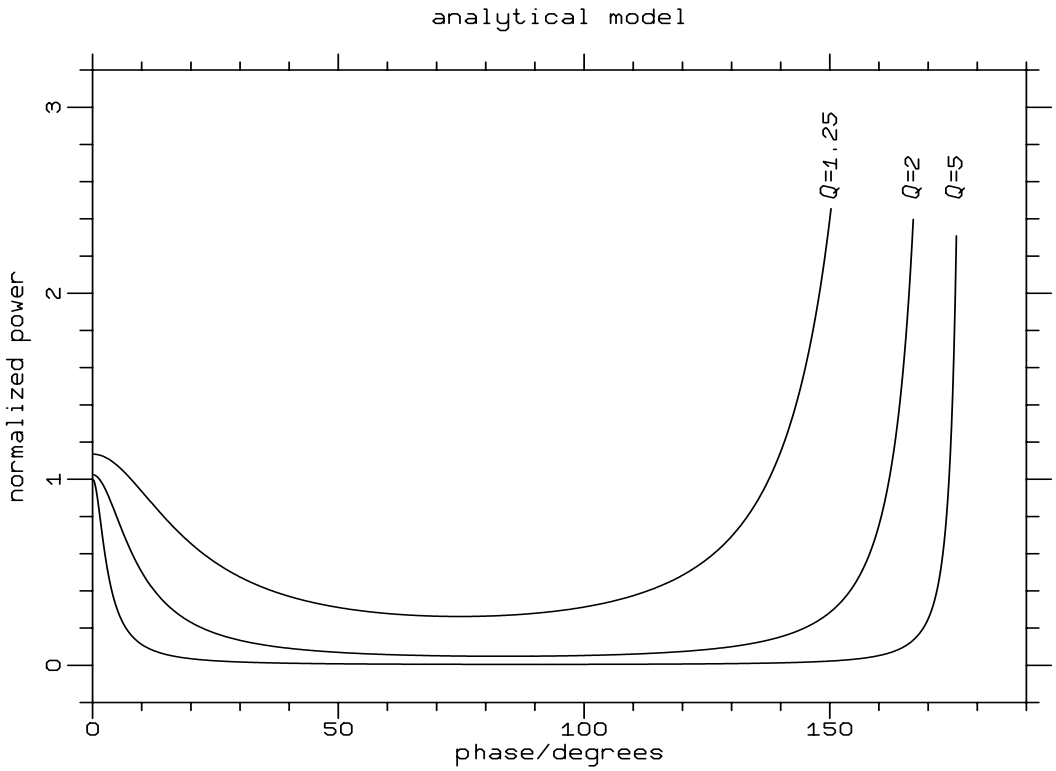

Figure B.2: Relative power consumption as a function of phase, for 3 values of Q, showing the importance of the phase setting when the damping is large.

value. Substitution in equation B.18 demonstrates that the power/phase curve is quite flat-bottomed around that point, but with steep sides for large phase errors. Thus for all practical purposes the phase shift may safely be set to 90° according to this model. This was confirmed by measurements with the brassboard modulator (see chapter 4). However, the phase error arising from the inductance of the coils in a voltage-driven system (such as the ISAMS drive circuit) is far more significant; for the PMIRR CO_2 coils at a frequency of 40 Hz, the current lags the voltage by 58°, so that the effective PPO/current phase difference is 32° for a 90° demand. In a voltage-driven system, with a Q of 1.25 implied from the measured phase for minimum power with the brassboard modulator, the power would be 1.71 times its minimum value.

This page intentionally left blank.

Appendix C:
Temperature structure in cylinders

C.1 Physical assumptions

The temperature distribution in the cell or the internal volumes of the pressure modulator is a complex problem, but an important one, because it defines the heat input into the gas, which affects the thermodynamic cycling and thus the optical transmission cycle within the cell. There are many processes in operation: conduction through the gas to the walls, convection, and turbulence are the most significant.

To produce simple terms for the conducted heat which can be included in the numerical model of the PMU, I assume that all the volumes are cylindrical in shape, that thermal conduction is the major mechanism, and that the temperature distribution is axially symmetrical. These assumptions produce simple expressions for the time constants, bulk temperature and conducted heat, and yield some insight into the physics of the heat transfer in the modulator. Most of the volumes are not truly cylindrical, although there is cylindrical symmetry; the internal volumes are strangely shaped as a result of the mechanical constraints on their design, and the PMIRR CO_2 cell is slightly conical in section. I have assumed that gas enters the volumes in a uniformly-distributed way. This also is questionable, but the extra accuracy gained by considering the connecting geometry is not justified as the errors introduced by the approximations are small compared with other uncertainties, even in the cell volume.

C.2 Mathematics

The heat conduction and heat capacity equations are:

$$Q = -\sigma A \frac{dT}{dx} \tag{C.1}$$

and

$$\delta Q = m c_v \, \delta T \tag{C.2}$$

In these expressions, σ is the thermal conductivity, A is a cross-sectional area for heat transfer, m is the mass of gas in the volume, c_v is the specific heat capacity at constant volume, T is the temperature, x is a length dimension, and Q is an amount of heat. Equations C.1 and C.2 may be used to derive the basis functions of the temperature distribution. In the following discussion, the cylinder has length L and radius ξ. The axial coordinate is z, running from 0 to L. The radial coordinate is r. For simplicity the temperature T refers to the difference between the gas temperature at a point and the temperature of the walls.

An elemental annulus of thickness dr and length dz has four surfaces that conduct heat. The amounts of heat entering from the inside and leaving from the outside are

$$dQ_{inside} = -\sigma\, 2\pi r\, dz \frac{\partial T}{\partial r}_{\text{at } z,r} \tag{C.3}$$

and

$$dQ_{outside} = -\sigma\, 2\pi (r + dr)\, dz \frac{\partial T}{\partial r}_{\text{at } z,r+dr} \tag{C.4}$$

The amounts of heat entering from the 'bottom' (nearer $z = 0$) and leaving from the 'top' are

$$dQ_{bottom} = -\sigma\, 2\pi r \frac{\partial T}{\partial z}_{\text{at } z,r} \tag{C.5}$$

and

$$dQ_{top} = -\sigma\, 2\pi r \frac{\partial T}{\partial z}_{\text{at } z+dz,r} \tag{C.6}$$

The net heat input power is matched by a temperature rise:

$$\frac{dQ}{dt}_{net} = \frac{\partial T}{\partial t}_{\text{at } z,r} c_v \rho\, 2\pi r\, dr\, dz = 2\pi\sigma \left(dz\, dr \frac{\partial}{\partial r}\left(r \frac{\partial T}{\partial r} \right) + r\, dr\, dz \frac{\partial^2 T}{\partial z^2} \right) \tag{C.7}$$

Division by $2\pi\, dz\, r\, dr$ yields

$$\frac{1}{dr} \frac{\partial}{\partial r}\left(r \frac{\partial T}{\partial r} \right) + \frac{\partial^2 T}{\partial z^2} = \frac{c_v \rho}{\sigma} \frac{\partial T}{\partial t} \tag{C.8}$$

which is suitable for separation of variables:

$$T = R(r)\, Z(z)\, \Theta(t) \tag{C.9}$$

Substitution and division by T gives

$$\frac{1}{rR} \frac{d}{dr}\left(r \frac{dR}{dr} \right) + \frac{1}{Z} \frac{d^2 Z}{dz^2} = \frac{c_v \rho}{\sigma} \frac{1}{\Theta} \frac{d\Theta}{dt} \tag{C.10}$$

As each term is a function of an independent variable, it must have a constant value; let the radial one be $-A^2$, and the axial one $-B^2$.

The radial equation from C.10 when expanded is Bessel's equation in zeroth order:

$$r \frac{d^2 R}{dr^2} + \frac{dR}{dr} + A^2 R r = 0 \tag{C.11}$$

whose solutions are $J_0(Ar)$. Other Bessel functions (Y_n type) are not appropriate here because the value at $r = 0$ must be bounded. The value of A can be chosen so that the temperature is zero at the boundary $r = \xi$; thus it may take any value p_n/ξ, where p_n designates the nth root (from $n = 0$) of $J_0(x) = 0$, *i.e.*, 2.405, 5.520... This produces a set of radial functions with different radial scales. Thus

$$R = \sum_{n=0}^{\infty} a_n J_0\left(\frac{p_n}{\xi} r \right) \tag{C.12}$$

where p_n is one of the values discussed above, and a_n is the weight of the nth radial mode. $A = p_n/\xi$.

The axial dependence in C.10 is sinusoidal, with antinodes at the ends, and the constant B is equal to $m\pi/L$, so the axial dependence is given by

$$Z = \sum_{m=1}^{\infty} b_m \sin\left(\frac{m\pi}{L}z\right) \tag{C.13}$$

The time-dependence in C.10 is a simple exponential decay:

$$\Theta = \exp\left(-\frac{(A^2 + B^2)\sigma}{c_v\rho}t\right) \tag{C.14}$$

The spatial components C.12 and C.13 may be combined:

$$S = \sum_{n=0}^{\infty}\sum_{m=1}^{\infty} a_n b_m J_0\left(\frac{p_n}{\xi}r\right)\sin\left(\frac{m\pi}{L}\right) \tag{C.15}$$

For convenience, I shall introduce a 'shape factor', $s = L/\xi$ in the rest of the discussion.

C.3 Rate and time constants

The characteristic rate constant in the temporal dependence is

$$\lambda_{n,m} = \frac{1}{\tau_{n,m}} = \frac{(A^2 + B^2)\sigma}{c_v\rho} \tag{C.16}$$

$$= \frac{\sigma}{c_v\rho}\left(\left(\frac{p_n}{\xi}\right)^2 + \left(\frac{m\pi}{L}\right)^2\right) = \frac{\sigma}{c_v\rho}\left(\frac{p_n}{\xi}\right)^2\left(1 + \left(\frac{m\pi}{p_n s}\right)^2\right) \tag{C.17}$$

which increases with both n and m, showing that the more complex modes decay faster than the simpler ones. Physically this is expected, as the more complex the mode, the steeper the gradient at the side or end wall of the boundary, so that more heat is conducted to or from the wall.

Any given initial temperature distribution may therefore be conceived as decaying towards the simplest modes. In the long (prolate) cell case, the radial modes decay fastest, so that there is a progression first to the simplest radial mode with a monotonic radial decrease in temperature difference from the centre to the wall, but with the axial distribution remaining largely constant, and thence via axial mode decay to the simplest mode in both directions, the last axial mode being that with a monotonic axial decrease in both directions from the centre of the cell to the ends. For a short (oblate) cell the order is reversed. In reality both processes occur simultaneously. For a cylinder in which the length is much greater than the radius, (such as the optical cell) the conduction to the sides dominates, and the axial rates form a sort of 'fine structure' on the radial rates. For a short cylinder (such as the space between the pistons) the converse is true.

C.4 Bulk mean temperature

Since it would be impractical to consider all the modes in any given volume as part of a computer model of the PMU, it is necessary to evaluate the volume- or mass- weighted mean temperature, thereby producing a single coordinate. The temperature variations within the cell are small compared with the absolute temperature, so the two weightings are largely equivalent. The volume of the elemental annulus used to derive the temperature distribution is

$$dV = 2\pi r\, dr\, dz \tag{C.18}$$

and thus, from C.15, the contribution to the mean temperature from the (n, m)th mode is

$$\pi\xi^2 L\overline{T} = 2\pi a_n b_m \int_0^{\xi} r J_0\left(\frac{p_n}{\xi}r\right) dr \int_0^{L} \sin\left(\frac{m\pi}{L}z\right) dz \tag{C.19}$$

It is convenient to substitute $x = p_n r/\xi$, and so

$$\int_0^{\xi} r J_0\left(\frac{p_n}{\xi}r\right) dr = \left(\frac{\xi}{p_n}\right)^2 \int_0^{p_n} x J_0(x)\, dx = \left(\frac{\xi}{p_n}\right)^2 p_n J_1(p_n) \tag{C.20}$$

Thus for even m the mean temperature contribution is zero; for odd m it is

$$\pi\xi^2 L\, \overline{T}_{n,m} = a_n b_m \frac{2L}{m\pi}\left(\frac{\xi}{p_n}\right)^2 p_n J_1(p_n) \tag{C.21}$$

$$\Rightarrow \overline{T}_{n,m} = a_n b_m \frac{2}{m\pi^2}\frac{J_1(p_n)}{p_n} \tag{C.22}$$

The fact that m has to be odd has the physical reason that each half of the length of the volume contributes oppositely to the mean temperature. As the even m axial functions are antisymmetric about the centre of the cell, they make no net contribution (This does not apply for conical volumes.)

C.5 Heat conducted

The heat conducted into the gas is calculated from the temperature gradient at the wall, integrated over all the boundaries. Again considering the temperature structure of the (n, m)th mode from C.15, and the contributions from the walls (integrating over z) and the ends (integrating over r):

$$H_{n,m} =$$

$$2\pi\sigma a_n b_m \left(\xi\frac{p_n}{\xi}J_1(p_n)\int_0^{L}\sin\left(\frac{m\pi z}{L}\right)dz + \frac{m\pi}{L}\cos\left(\frac{m\pi z}{L}\right)\int_0^{\xi} J_0\left(\frac{p_n r}{\xi}\right) r\, dr\right) \tag{C.23}$$

which may be simplified to

$$H_{n,m} = 2\pi\sigma a_n b_m\, p_n J_1(p_n)\left(\frac{2L}{m\pi} + \frac{2m\pi}{L}\left(\frac{\xi}{p_n}\right)^2\right) \tag{C.24}$$

$$= \sigma a_n b_m \, 2\pi p_n J_1(p_n) \frac{2L}{m\pi} \left(1 + \left(\frac{m\pi}{p_n s} \right)^2 \right) \tag{C.25}$$

where m is odd; $H_{n,m}$ is zero for even m. The physical explanation of this is that antisymmetric axial modes contribute oppositely to the heat conducted between the gas and the wall in the two axial halves of the cell.

Several terms ($a_n b_m \, (2L/m\pi) \, p_n J_1(p_n)$) in this expression cancel terms in the expression for the temperature contribution of the mode, so that for one mode

$$\frac{H_{n,m}}{\overline{T}_{n,m}} = 2\pi^2 \sigma \, p_n^2 L \left(1 + \left(\frac{m\pi}{p_n s} \right)^2 \right) \tag{C.26}$$

This expression is proportional to the time constant for the mode, (equation C.17):

$$\frac{H_{n,m}/\overline{T}_{n,m}}{\lambda_{n,m}} = 2\pi^2 \, c_v \rho \, \xi^2 L \tag{C.27}$$

C.6 Mode weightings

To find a useful and reasonable set of mode weights a_n and b_m, I considered two extreme cases, firstly a uniform temperature distribution, and secondly the single lowest mode, for which $a_n = 0, n > 0$ and $b_m = 0, m > 1$. To be useful for the purposes of the thermodynamic model the ratio of the heat conducted to the temperature difference should be well-defined. The uniform temperature distribution could be achieved by weighting the axial modes as $1/m$ for odd m and zero for even m, analogously to the weights of harmonics in a square wave. The radial weights I found by trial and error to vary as

$$\frac{(-1)^n}{\sqrt{p_n}} \tag{C.28}$$

The value of $H_{n,m}/\overline{T}_{n,m}$ for such a distribution is an unbounded sum, because the temperature gradient at the wall for such a distribution is infinite by nature. Although such a temperature distribution could be set up, the higher modes would decay quickly, leaving the low n and m ones. For these reasons I abandoned the uniform temperature distribution.

The true temperature distribution in each volume cell will vary with time, but because of the simplicity of considering the single-mode case, and the long time constant associated with the $n = 0, m = 1$ mode, I have assumed it to be the only one present in the calculations, using the expression:

$$\frac{H}{\overline{T}} = 2\pi^2 \sigma \, p_0^2 L \left(1 + \left(\frac{\pi}{p_0 s} \right)^2 \right) \tag{C.29}$$

This ratio depends upon the thermal conductivity of the gas, and the length and shape of the volume. For a composite volume which consists of several cylinders of the same shape factor stacked end-to-end, the total heat conducted is given by using the expression C.29 with s as the shape of each partition and L the total length of the volume. The volume between the pistons in the pressure modulator was considered as a stack of two cylinders

separated by the coupling spring; the volume behind the pistons was considered as a stack of 3 cylinders, with boundaries at the suspension springs.

The thermal relaxation rate of this mode is

$$\lambda_{0,1} = \frac{\sigma}{c_v \rho}\left(\frac{p_0}{\xi}\right)^2\left(1+\left(\frac{\pi}{p_0 s}\right)^2\right) \tag{C.30}$$

By comparison with C.29, this may be written as

$$\lambda_{0,1} = \frac{H/\overline{T}}{2\pi C_v} \tag{C.31}$$

where C_v is the (total) heat capacity of the gas in the volume. For both PMIRR PMUs the relaxation rate in the optical cell had a value similar in magnitude to the PMU frequency, so that thermal conduction was an important consideration in the modelling.

C.7 Range of point temperatures

The temperature at a point for the single (n, m)th mode case C.15 may be described in terms of the mean temperature C.22 as follows:

$$\frac{T(r,z)}{\overline{T}_{n,m}} = \frac{\pi^2}{2}\, p_n \frac{J_0\left(\frac{p_n r}{a}\right)}{J_1(p_n)}\, m \sin\left(\frac{m\pi z}{L}\right) \tag{C.32}$$

The 'maximum' temperature, *i.e.*, the greatest difference in temperature between the gas and the wall, occurs in the centre of the cell, when $r = 0$ and $z = L/2$; the value of C.32, for the (0,1)th mode at that point is $\pi^2 p_0/(2J_1(p_0))$, approximately 22.8. This large value represents a small volume of gas because the radius is small. The standard deviation may be evaluated as

$$\frac{\sigma_T^2}{2\overline{T}_{0,1}^2} = \left(1 - 2\pi + \frac{\pi^3}{4} p_0^2\right) \tag{C.33}$$

which is approximately 40; thus the ratio of the standard deviation of temperature σ_T to the mean temperature is about 9. This is indicative of a large variation in temperature throughout the volume.

The mean temperature near the centre of the cell (in the middle third of the length and the central area radially, to simplify the integration) is

$$T_{centre} = 6\overline{T}_{0,1}\frac{1}{2\sqrt{2}}\frac{J_1(p_0/\sqrt{2})}{J_1(p_0)} = 2.35\,\overline{T} \tag{C.34}$$

The ranges seem counter-intuitively large; they imply that a (relatively modest) 10K variation in mean temperature is related to a 200K variation in centre temperature and a spatial distribution of temperature with a 100K standard deviation, both of which are unrealistic. The conclusion therefore is that the temperature distribution must be more uniform than the single-lowest-mode case admits, and that more modes must be present. Greater uniformity in either direction yields a smaller spread in temperature.

As it is not possible by reasoning to deduce the relative mode weights and because the 'fine structure' modes (which are axial in the case of a long cylinder) have similar ratios of heat conduction to mean temperature, it is reasonable to approximate the conducted heat by that predicted for the lowest mode alone. However, as the spread of temperatures it produces is unrealistic, it should not be used to infer the distribution of temperature in each volume.

Appendix D:
Multiple reflections from parallel surfaces

Since the windows of the pressure modulator cells are plane and parallel, multiple reflections can occur both within each window and between the two inner faces. Fringes with a spectral separation characteristic of the latter were observed in the tunable diode laser measurements for both PMIRR flight modulators, but fringes from reflections inside the window material were not, since the free spectral range was larger than the frequency range of the measurements. Accordingly this appendix has been written with window-to-window reflections particularly in mind.

Hecht and Zajac [1974] give an expression for the ratio of the transmitted intensity I_t to the incident intensity I_i for fringes of equal thickness arising from a plane parallel beam of light of wavelength λ, incident at an angle θ to the normal:

$$\frac{I_t}{I_i} = A(\theta) = \frac{1}{1 + F\sin^2(\delta/2)} \qquad (D.1)$$

$A(\theta)$ is known as the *Airy function*, and F is the *coefficient of finesse*:

$$F = \left(\frac{2r}{1 - r^2}\right)^2 \qquad (D.2)$$

where r is the reflectance from one surface. δ is the phase difference between the interfering beams, and since the refractive index of the gas is close to unity,

$$\delta = 4\pi\nu d\cos\theta \pm \pi \qquad (D.3)$$

where ν is the reciprocal of wavelength, and d is the separation of the windows. If the $\sin^2$ term in equation D.1 is changed to $\cos^2$, the $\pm\pi$ at the end of equation D.3 can be ignored.

The free spectral range (wavelength interval between fringes) is

$$(\Delta\nu)_{fsr} = \frac{1}{2d\cos\theta} \qquad (D.4)$$

When the coefficient of finesse is small, the depth (maximum transmission deficit) of the fringes predicted by equation D.1 is, by inspection, $1 - 1/(1 + F)$.

The depth of the fringes measured in the wideband transmission of the PMIRR flight model H_2O modulator was about 2.5%, and the beam was well collimated, so the coefficient of finesse was approximately 0.026, which corresponds to a reflectance of 0.08. The free spectral range (FSR) expected for a cell 10 cm long, is 0.05 cm^{-1}, from equation D.4, as observed in the TDL measurements (chapters 6 and 7). For the CO_2 cell, which is 4 cm long, the FSR is 0.125 cm^{-1}.

When the beam is not perfectly collimated, the expression D.1 must be averaged over the solid angle it subtends. Contributions at different angles have different free spectral ranges, so that the fringes are merged into each other, and for large beam angles the depth of the fringes is reduced. Assuming that the radiation distribution is isotropic, the transmitted intensity at each angle θ from the normal is weighted by the solid angle element of a conical shell of thickness $d\theta$ at θ,

$$d\Omega = 2\pi \sin\theta d\theta = -2\pi d(\cos\theta) \qquad (D.5)$$

For small values of F, the approximation

$$A(\theta) \approx 1 - \frac{F}{2} + \frac{F}{2}\cos\delta \qquad (D.6)$$

may be used; it demonstrates variation about a mean transmission of $1 - F/2$. Ignoring the $\pm\pi$ in the expression for δ, since it is a phase constant which does not affect the general form but only the position of the fringes, and integrating up to the beam half-cone angle θ_1, the weighted transmission becomes

$$\frac{I_t}{I_i} \approx 1 - \frac{F}{2} + \frac{F}{8\pi\nu d}\frac{\sin 4\pi\nu d - \sin 4\pi\nu d\cos\theta_1}{1 - \cos\theta_1} \qquad (D.7)$$

This expression, which cannot easily be simplified because νd is so large and therefore small-angle and derivative approximations cannot apply, shows that the effect of the multiple reflections is still to produce a mean transmission of $1 - F/2$, but that there will be two overlapping sets of fringes, at spacings characteristic of the central ray and the extreme ray. These beat together, modulating the transmission envelope.

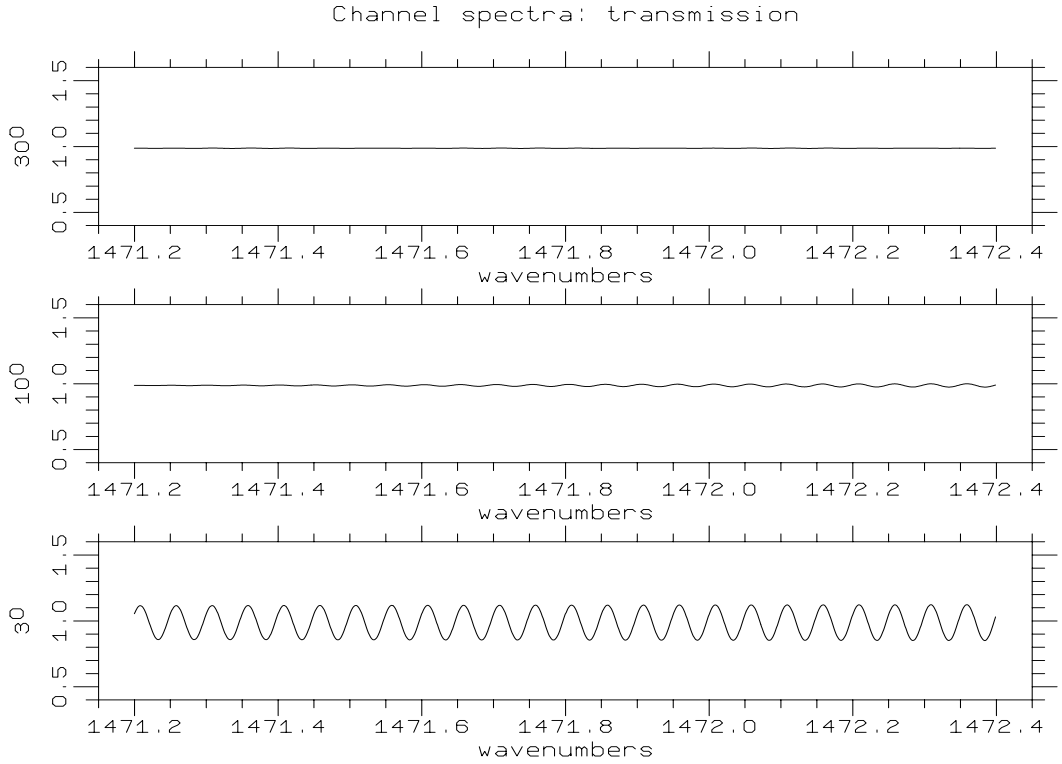

Figure D.1: Calculated transmission of the empty PMIRR H_2O pressure modulator cell over the frequency interval covered by the TDL measurements, for three maximum beam angles (3, 10 and 30 degrees), and a coefficient of finesse of 0.026.

Figure D.1 shows the resulting fringe transmission or channel spectrum effect. Absorption by the gas itself has been ignored here; the multiple reflections will, of course, enhance the absorption as occurs in a White cell, although for such a small coefficient of finesse the effect is small. It is clear from this figure that the magnitude of the fringe effects decreases as the beam angle increases, the implication being that use of a converging or diverging beam is advantageous. When the beam is collimated, the windows must be wedged and/or angled with respect to the normal to prevent these channel spectra.

This page intentionally left blank.

Appendix E:

References

Bauer, A, Godon, M, Kheddar, M, Hartmann, J H, Bonamy, J and Robert, D: Temperature and Perturber Dependences of Water Vapor 380 GHz Line Broadening *JQSRT* **37** 531, 1987

Brunauer, S, Emmett, P H and Teller, E: Adsorption of Gases in Multimolecular Layers *J. Am. Chem. Soc.* **60**309, 1938

Chiarrella, C and Reichel, A: *Mathematics of Computation* **22** 137, 1968

Conrath, B J: Planetary-scale Wave Structure in the Martian Atmosphere *Icarus* **48** 246, 1981

CRC: Handbook of Chemistry and Physics (64th edition, Weast R C) *Chemical Rubber Publishing Co., Boca Raton, Florida*, 1984

Curtis, P D: Measurement of temperatures in the upper atmosphere using satellite techniques *D. Phil thesis, University of Oxford*, 1974

Davis, G R: Remote sensing of atmospheric water vapour by pressure modulation radiometry *D. Phil thesis, University of Oxford*, 1987

Delaye, C, Hartmann, J-M, and Taine, J: Calculated tabulations of H_2O line broadening by H_2O, N_2, O_2, and CO_2 at high temperature *Applied Optics* **28** 5080, 1989

Eisenberg, D and Kauzmann, W: The Structure and Properties of Water *O.U.P.*, 1969

Elsey, R J: Outgassing of Vacuum Materials - I *Vacuum* **25** 299, 1976

Farmer, C B and LaPorte, D D: The detection and mapping of water vapour in the Martian Atmosphere *Icarus* **16** 34–46, 1972

Farmer, C B, Davies, D W, Holland, A L, Laporte, D D and Doms, P E: Mars: Water

Vapor Observations from the Viking Orbiters *J G R* **82** 28 p4225, 1977

Gamache, R R and Davies, R W: Theoretical Calculations of N_2 broadened halfwidths of H_2O using Quantum Fourier Theory *Applied Optics* **22** 4013, 1983

Goody, R and Belton, M J S: Radiative relaxation times for Mars: a discussion of Martian atmospheric dynamics *Planet. Spac.* **15** 247, 1967

Gray, L D and Young, A T: Relative intensity calculations for carbon dioxide: IV Calculation of the partition function for the isotopes of CO_2 *JQSRT* **9** 569, 1969

Grossmann, B E and Browell, E V: Spectroscopy of Water Vapor in the 720-nm Wavelength Region: Line Strengths, Self-Induced Pressure Broadenings and Shifts, and Temperature Dependence of Linewidths and Shifts *J of Molecular Spectroscopy* **136** 264, 1989a

Grossmann, B E and Browell, E V: Water Vapor Line Broadening and Shifting by Air, Nitrogen, Oxygen and Argon in the 720-nm Wavelength Region *J of Molecular Spectroscopy* **138** 562, 1989b

Hecht, E, and Zajac, A: Optics *Addison-Wesley*, 1974

Houghton, J T, Taylor, F W and Rodgers, C D: Remote Sounding of Atmospheres *Cambridge University Press*, 1984

Irwin, P G J: Remote Sounding of the Martian Atmosphere *D.Phil thesis, University of Oxford*, 1991

Jones, R L: Measurement of the Atmospheric Distribution of CH_4 and N_2O from Satellites *D.Phil thesis, University of Oxford*, 1983

Kaye, G W C and Laby, T H: Tables of Physical and Chemical Constants (14th edition) *Longman and Co, London*, 1973

Keyes, F G: The Consistency of the Thermodynamic Data for Water Substance Vapor Phase to 550 °C, Part VII *J. Chem. Phys.* **17** 923, 1949

Leovy, C B: Exploration of the Atmosphere and Climate System of Mars *in Szegö, K (ed.)* **The Environmental Model of Mars** Cospar Colloquia Series, Volume 2 (Pergamon Press), 1990

Liebe, H J, Thompson, M C and Dillon, T A: Dispersion Studies of the 22 GHz Water Vapor Line Shape *JQSRT* **9** 31, 1969

McClatchey, R A, Benedict, W S, Clough, S A, Burch, D E, Calfee, R F, Fox, K, Rothman, L S and Garing, J S: AFCRL Atmospheric Absorption Line Parameters Compilation *Air Force Cambridge Research Laboratory report TR-73-0096 Environment Research Papers 434*, 1973

McCleese, D J, Schofield, J T, Zurek, R W, Martonchik, J V, Haskins, R D, Paige, D A, West, R A, Diner, D J, Locke, J R, Chrisp, M P, Willis, W, Leovy, C B and Taylor, F W: Remote sensing of the atmosphere of Mars using infrared pressure modulation and filter radiometry *Applied Optics* **25** 4232, 1986

May, R D, McCleese, D J, Rider, D J, Schofield, J T and Webster, C R: Tunable diode laser spectral diagnostic studies of a pressure modulator radiometer *Applied Optics* **27** 3591, 1988

Mooradian, A: High Resolution Tunable Infrared Lasers *Chapter 4 in* **Smith, R A (ed)** Very High Resolution Spectroscopy, 1976

Muggeridge, A H: The retrieval of atmospheric pressure and temperature from high resolution absorption spectra *D. Phil thesis, University of Oxford*, 1986

Munro, R: Middle Atmosphere Water Vapour *D. Phil thesis, University of Oxford*, 1991

Mutlow, C T: Studies of atmospheric water vapour using satellite data *D. Phil thesis, University of Oxford*, 1984

NASA: On Mars: Exploration of the Red Planet 1958–1978 *NASA, Washington*, 1984

Newman, F H and Searle, V H L: The General Properties of Matter (fourth edition) *Edward Arnold, London*, 1948

Owen, T and Mason, H P: Mars: Water Vapour in its atmosphere *Science* **165** 893, 1969

Pollack, J B, Colburn, D S, Flasar, F M, Kahn, R, Carlston, C E and Pidek, D: Properties and Effects of Dust Particles Suspended in the Martian Atmosphere *J G R* **84** 2929, 1979

Press, W H, Flannery, B P, Teukolsky, S A and Vetterling, W T: Numerical Recipes in C – the Art of Scientific Computing *Cambridge University Press, New York*, 1988

Remedios, J J: Spectroscopy for Remote Sounding in the Atmosphere *D. Phil thesis, University of Oxford*, 1990

Roberts, J K and Miller, A R: Heat and Thermodynamics *Blackie and Son, London*, 1956

Roscoe, H K: Intercomparison aspects of balloon-borne measurements of NO_2 *Proc. Quad. Int. Ozone Symposium*, 1988

Roscoe, H K and Wells, R J: The variation of pressure, temperature and transmission within a pressure modulator: measurements with a high-compression modulator *JQSRT* **41** 259, 1989

Roscoe, H K, Kerridge, B J, Pollitt, S, Louisnard, N, Flaud, J-M, Camy-Peyret, C, Alamichel, C, Pommereau, J-P, Ogawa, T, Iwagmi, N, Coffey, M T, Mankin, W, Evans, W F J, McElroy, C T and Kerr, J: Intercomparison of remote measurements of stratospheric NO and NO_2 *J. Atmos. Chem* **10** 111, 1990

Rothman, L S, Gamache, R R, Goldman, A, Brown, L R, Toth, R A, Pickett, H M, Poynter, R L, Flaud, J-M, Camy-Peyret, C, Barbe, A, Husson, N, Rinsland, C P, and Smith, M A H: The HITRAN database – 1986 edition *Applied Optics* **26** 4058, 1987

Schofield, J T: The remote sounding of the Venus atmosphere *D. Phil Thesis, University of Oxford*, 1980

Schorn, R A; Farmer, C B and Little, S J: High-dispersion spectroscopic studies of Mars. III: Preliminary results of 1968–69 water vapour studies *Icarus* **11** 283, 1969

Taylor, F W: Pressure Modulation Radiometry *in* **Vanasse, G A** Spectrometric Techniques III, chapter 3, 1984

Toth, R A: Frequencies of N_2O in the 1100 to 1440 cm^{-1} region *Journal of the Optical Society of America B* **3** 1263, 1986

de Vaucouleurs, G: Physics of the planet Mars *Faber and Faber, London*, 1954

Venters, P: The Pressure Modulation System in the Improved Stratospheric and Mesospheric Sounder *D. Phil Thesis, University of Oxford*, 1991

Webster, C R, Menzies, R T and Hinkley, E D: Infrared Laser Absorption: theory and applications *in* **Laser remote chemical analysis** ed. Measures, R D, chapter 3 (Wiley and

Sons), 1988

Wells, R A: Geophysics of Mars *Elsevier Scientific Publishing Co., New York*, 1979

Wells, J S, Schneider, M and Maki, A G: Calibration Tables Covering the 1460 to 1550 cm^{-1} Region from Heterodyne Frequency Measurements on the ν_3 Bands of $^{12}CS_2$ and $^{13}CS_2$ *Journal of Molecular Spectroscopy* **132** 422, 1988

Zemansky, M W, and Dittman, R H: Heat and Thermodynamics *McGraw-Hill*, 1981

This page intentionally left blank.