UMBC_LBL vers 7: An Algorithm to Compute Line-by-Line Spectra

Sergio De Souza-Machado, L. Larrabee Strow,
David Tobin, Howard Motteler and Scott Hannon

Physics Department
University of Maryland Baltimore County
Baltimore, MD 21250 USA

v6 March 4, 2002

Sergio De Souza-Machado: sergio@umbc.edu
L. Larrabee Strow: strow@umbc.edu

Contents

List of Tables

List of Figures

ABSTRACT

We have developed a line-by-line code to compute the spectral lineshapes of gases under varying pressure and temperatures. This code can use the spectral line parameters of a variety of databases such as HITRAN or GEISA. The code ensures that the same lines are used in all layers of an atmosphere, so that the resulting line profiles are smooth. The latest models for CO2 line mixing and in water vapor lineshapes are incorporated in the code. The code is can be used to compute absorption spectra at user specified resolutions, using one of various lineshapes. One of our uses for the code is to generate new versions of the spectroscopic database for kCARTA, which is the reference for the AIRS forward model.

This document is very much a work in progress. Some major omissions include references, significant examples of output, and comparisons of output to GENLN2. These omissions will be rectified in the future. Please give us your feedback on both the code and the documentation!

1 Introduction

This documentation describes a line-by-line code that can be used to compute optical depths for various gases (that appear in the HITRAN database). The code has been written in Matlab, with many of the routines written as FORTRAN-MEX files in the interests of speed. The lineshape database that is used in the computations, is currently the HITRAN98 database. Other databases, such as the GEISA or HITRAN92 databases, can be merged in and used. With this, we can incorporate the latest lineshape studies and parameters into our computations.

This code will be used to generate a new database for kCARTA, our radiative transfer algorithm that is the reference for the AIRS forward model. AIRS is a high resolution infrared instrument due to be launched by NASA in the year 2000. The instrument will be used to make measurements of the gas content and temperatures of the Earth's atmosphere, as well as study the global climate and weather.

For all gases except carbon dioxide and water vapor, the lineshapes of the individual lines are simply added together to give the overall absorption spectrum. To speed things up, the code uses the binning methods of GENLN2, where the lines are divided into near lines, medium-far lines and far lines. The general code to compute the lineshape of this majority of gases is called $run6.m$ Note that we also have code to incorporate the computation of the lineshapes of the cross section gases as well.

For the case of carbon dioxide, we have utilised Dave Tobin's work on linemixing and duration of collisions effects in the 4 μm band, and applied it to the 15 μm band as well. The code to compute the lineshape of carbon dioxide is called $run6co2.m$

For the case of water vapor, we have written two source codes. $run6water.m$ can compute the spectral lineshapes with or without the basement term removed, and add on the user specified continuum (CKD 0,2.1,2.3 or 2.4). In addition, this code also includes Tobin's work on water vapor chi functions. We also have a specialised continuum code $run6watercontinuum.m$ that can output one of the following (a) the completely computed water continuum (b) the self continuum coefficients (c) the foreign continuum coefficients or (d) combinations of the above three.

We begin the documentation by briefly describing how our code computes an optical depths, given the necessary input parameters. This section is divided into three; the first can be used for all gases, while the second and third apply to water vapor and carbon dioxide respectively.

The next sections describe the $UMBC\_LBL$ Version 6 code implemetation. A section on the basic code for most gases, $run6.m$ is then followed by section on $run6water(continuum).m$ and $run6co2.m$. We have improved upon Version 6 by defaulting most parameters to a set of values that are usually used; if the user wants to change these settings, he/she can send in the new value(s) by using a Matlab structure. This will be the *UMBC_LBL* Version 7 implementation. All the basic physics is the same as version 6.

Theis is followed by sections which describe spectral lineshapes in detail. First is a section on general spectral lineshape theory, where the meaning of terms such as line widths and line broadening is explained. Next are two sections, one on computing water vapor spectra and the other computing carbon dixide spectra. As mentioned in the previous paragraphs, to accurately compute the absorption spectra of these two gases, one needs to use more involved lineshape theories than those for other gases. This is particularly important as these two gases are actively important in the infrared portion of the radiation emitted by the Earth, and so a thorough understanding of their spectra would be beneficial to the remote sensing community.

Finally, a short set of appendices briefly describe the general code to compute partition functions, line widths and line strengths.

2 General algorithm to compute optical depth

Assume the user wants the optical depth to be computed for a certain gas, using known parameters. These parameters are the total pressure P , partial (self) pressure PS, temperature T and gas amount (column density) U. The gas amount U is related to self pressure, temperature, path cell length (or atmospheric layer thickness) L by the equation

$$U = \frac{CF \times L \times \times PS}{R \times T}$$

where R is the molar gas constant (8.31 J/mol/K), PS is the partial pressure in atmospheres, T is the temperature in Kelvin, and L is thepath length in cm. CF is a units conversion factor of 1.01325×10^{-4} to give us U in units of $kiloMoles/cm^2$. This equation uses the Ideal Gas Law to calculate molecular density for a unit volume, and then multiplies the density by the path length L to convert density to column density (gas amount).

Knowing for what gas the spectrum is to be computed, and within which frequency interval, the lineshape parameters for that gas need to be read in. These parameters can be obtained from a database such as HITRAN98, and should include the following :

TYPE	DESCRIPTION	USE
gasid	input HITRAN GasID	1=water,2=CO2, etc
iso	list of isotopes	1=most abundant,2=next ...
linct	number of lines in wavenumber interval	
wnum	wavenumbers of the line centers $[cm^{-1}]$	$\nu_0(1), \nu_0(2), \ldots$
tsp	line center shift due to pressure P [cm-1/atm]	$\nu_0(j) \to \nu_0(j) + P \times tsp(j)$
stren	line strengths (See Eqn 26)	$S(T) \simeq S(296)Z(296)/Z(T) \times \rho$
	$[cm^{-1}/(molecules cm^{-2})]$	Z are partition functions (Eqn 21)
		$\rho \simeq$ upper/lower state populations
abroad	air broadened half widths HWHM $[cm^{-1}/atm]$	$brd_{air} = (P - PS) \times abroad$
sbroad	self broadened half widths HWHM $[cm^{-1}/atm]$	$brd_{self} = (PS) \times sbroad$
abcoef	temperature dependence of air	$brd = brd_{air} + brd_{self}$
	broadened half width	$brd \to brd \times (296/T)^{abcoef}$
tprob	transition probabilility $[debyes^2]$	
els	lower state energy $[cm^{-1}]$	computing S(T)
usgq	Upper state global quanta index	identifying
lsgq	Lower state global quanta index	P,Q,R branches
uslq	Upper State local quanta	local quantum numbers
bslq	Lower State local quanta	local quantum numbers

ai	accuracy indices
ref	lookup for references

Note that the $stren$ parameter is used in units of $cm^{-1}/(moleculescm^{-2})$ while the gas amounts U are in units of $kiloMoles/cm^2$. To get absorption coefficients and optical depths that are in the correct units, the program eventually multiplies $stren$ by 6.023×10^{26}, which is the number of molecules per kilomole (ie Avagadro's Number times 1000).

Having read in these parameters, the code is almost ready to proceed with the computations. Before this, it needs to read in the mass of each of the isotopes of the current gas. In addition, it needs to compute the partition function Z, and thus it needs four parameters a, b, c, d for each isotope (refer to Eqn 21). For the current temperature T , $Z(T)$ is then readily computed as :

$$Z(T) = a + bT + cT^2 + dT^3$$

The partition function is used when computing the line strength, as that term is proportional to the relative populations of the lower and upper levels of the transition.

Using parameter tsp, any shift of the line centers due to the total pressure is then computed, giving the adjusted line centers

$$\nu_0(P) = \nu_0 + P \times tsp(j)$$

The broadening of each line, due to the self component and the foreign component, is then computed.

$$brd_{air} = (P - PS) \times abroad$$

$$brd_{self} = (PS) \times sbroad$$

$$brd = brd_{air} + brd_{self}$$

$$brd \to brd \times (296/T)^{abcoef}$$

With the above information for each line having been determined, the absorption coefficients for the individual lines can now be computed, using the required lineshape (Lorentz, Doppler, Voigt, etc). Looking at the simplest lineshape equation, the Lorentz lineshape is given by

$$k_L(\nu) = \frac{1}{\pi} \left(\frac{\gamma_L}{(\nu - \nu_0)^2 + \gamma_L^2} \right)$$

where γ_L is the linewidth and ν, ν_0 are the wavenumber and line center frequency respectively. This means the units that result from a lineshape computation is cm

The final term required to compute the absorption spectrum is the line strength associated with each line. In the equation below, $S_i(T_{ref})$ is the line strength at the HITRAN reference temperature of $T_{ref} = 296K$, E_i is the lower state energy (given by els) and ν_i is the line center wavenumber :

$$S_i(T) = S_i(T_{ref}) \frac{Z(T_{ref})}{Z(T)} \frac{\exp(-hcE_i/kT)}{\exp(-hcE_i/kT_{ref})} \frac{[1 - \exp(-hc\nu_i/kT)]}{[1 - \exp(-hc\nu_i/kT_{ref})]}$$

With S_i being in units of $cm^{-1}/(moleculescm^{-2})$, when this mutliplies the lineshape, the final abosrption coeficient units are $1/(moleculescm^{-2})$

By summing over the individual lines, the absorption spectra can be computed. However, for purposes of radiative transfer in the atmosphere, it is more convenient to think in terms of optical depths (and transmittances) instead of absorption spectra. With this in mind, and to get the units right, the adjusted line strength is simply multiplied by the gas amount times number of molecules per kilomole : $U(kiloMoles/cm^2) \times 6.023 \times 10^{26}(molecules/kiloMoles)$, from which the optical depth of the line (in dimensionless units) can be computed. By summing over all the individual optical depths, the total optical depth can be computed :

$$k_i^{opticaldepth}(\nu, \nu_0) = S_i(T) \times U \times 6.023E^{26} \times LineShape(P, PS, T, \nu_0(i), \gamma(i), mass(i))$$

$$k(\nu)^{opticaldepth} = \sum_i k_i^{opticaldepth}(\nu, \nu_0)$$

The actual steps described above are summarized in Fig. 1. In the case of water vapor and carbon dioxide, the spectral lineshape that is used for each line can be more complicated than a simple lorentz or voigt shape. This will be discussed in more detail later. For now, let us obtain some order of magnitude estimates for the optical depths of some of the gases in the atmosphere. The gas amounts U come from the AIRS layers. Assume the temperature T is 296 K, so all temperature effects are unimportant. Here we use units $\eta = molecules/cm^2$ for gas amounts U. Following are typical line paramter values obtained from the HITRAN database, in the regions where the optical depths for the individual gases peak in the atmosphere :

gas	wavenumber cm^{-1}	abroad cm^{-1}/atm	sbroad cm^{-1}/atm	S cm^{-1}/η	U η	mix ratio %
water	1500	0.08	0.4	2E-19	5.6E+21	0.77
CO2	2400	0.08	0.1	4E-18	2.2E+20	0.0036
N2	2400	0.05	0.05	4E-28	4.8E+23	78.1
CH4	1320	0.06	0.09	1E-19	1.0E+18	0.00017

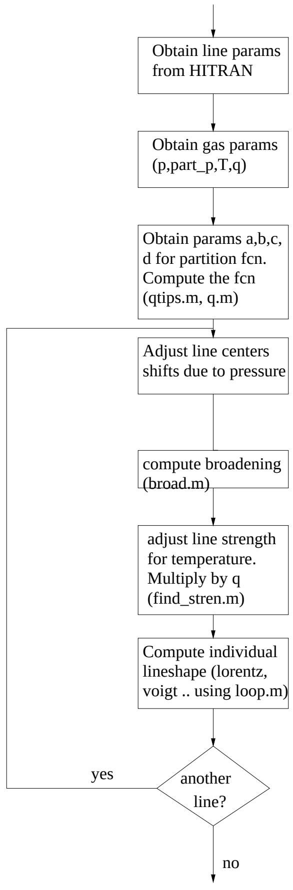

Figure 1: Flow diagram to compute $k(\nu)$

The amount of water vapor in the atmosphere varies greatly both in time and space. The mixing ratio of water near the surface might be anywhere between 0.01 and 4a few percent, while CH4 may vary by a factor of two. Note that due to the size and range of variation of the water mixing ratio, as well as the lage difference between air- and self- broadened widths, that the total line widths can vary by upwards of 15amount of water in the atmosphere. That is, the widths of water lines in a dry atmosphere will be noticeably narrower than the widths in a wet atmosphere. This is not true for any of the other important atmospheric gases; their widths are nearly constant. This is because either their mixing ratio is relatively constant, very tiny, or their air- and self-broadened widths are very similar.

The peak absorption coefficent in the lower atmosphere is approximately given by $S/(\pi \times \gamma)$ where γ is the linewidth computed from the self and foreign broadening. Using the values in the previous table, we get the following estimates for the optical depths in the lowest 200 meters of the atmosphere :

gas	total P atm	self P atm	γ $[cm^{-1}]$	peak abs $[1/\eta]$	peak optical depth
water	1.0	9.8E-3	8.8E-2	7.2E-19	4000
CO2	1.0	3.4E-7	8.5E-2	1.5E-17	3000
N2	1.0	8.4E-1	5.5E-2	2.3E-27	0.001
CH4	1.0	1.8E-6	6.4E-2	4.9E-19	0.5

One immediately sees that methane and nitrogen have much smaller peak optical depths than water or carbon dioxide. Most of the other gases in the atmosphere also have maximum optical depths of less than one. Remember that when the spectra are actually computed, there is a sum over the individual lines, so the peak values in the table above are slight underestimates. In addition, the above table is only for the lowest 200m meters of the atmosphere; the total surface-to-space optical depths are typically 50 times larger.

Using typical values from the table above (which use parameters for the lowest layer in the atmosphere), the absorption due to an individual line has decayed by $\frac{\gamma}{25^2+\gamma^2}/\frac{\gamma}{\gamma^2} \simeq 10^{-5}$ by the time it is $25cm^{-1}$ away from line center. For most gases except H2O and CO2, this $25cm{-}1$ distance may be used a safe cutoff limit beyond which the contribution of an individual line to the overall spectrum can be considered negligible. The large optical depth of the strong CO2 and water lines means we must go a couple hundred wavenumbers into the wing before the optical depth drops off to a negligible level!

To speed up the code while maintaining accuracy, our algorithm uses a variant of the GENLN2 method of binning the lines while computing the spectra. To compute the absorption spectrum in the chosen interval, the code divides up the interval into bins of width $fstep \simeq 1cm^{-1}$. It then loops through these bins, computing the overall spectra in each bin using three stages. To reinforce the ideas, actual example numbers will be used. Suppose the wavenumber interval being considered is 1005 - 1030 cm^{-1}. There are 25 bins of width $ffin = 1cm^{-1}$ in this interval. Suppose we are considering the $4th$ bin, which is the bin spanning 1008 to 1009 cm^{-1}.

- (a) *fine mesh stage* : lines that are $\pm xnear \simeq 1cm^{-1}$ on either side of the edges of this bin, are used in computing the overall lineshape at this stage, at a very fine resolution of $ffin \simeq 0.0005cm^{-1}$. The results of this computation are then boxcar averaged to the output resolution $nbox \times ffin \simeq$

$0.0025cm^{-1}$
Thus the lines that are used in this stage have their centers spanning 1007 to 1010 cm^{-1}.

- (b) *medium mesh stage* : lines that are an additional $xmed - xnear \simeq (2-1) = 1cm^{-1}$ on either side of the edges of this bin, are now used in computing the overall lineshape, at a medium resolution of $fmed \simeq 0.1cm^{-1}$. The results of this computation are then splined onto the output resolution and added onto the running sum from above. Thus the lines that are used in this stage have their centers spanning (1006,1007) and (1010,1011) cm^{-1}.

- (b) *coarse mesh stage* : lines that are an additional $xfar - xmed \simeq (25-2) = 23cm^{-1}$ on either side of the edges of the medium meshes, are now used in computing the overall lineshape, at a coarse resolution of $fcor \simeq 0.5cm^{-1}$. The results of this computation are then splined onto the output resolution and added onto the running sum from above. Thus the lines that are used in this stage have their centers spanning (983,1006) and (1011,1034) cm^{-1}.

The speed up in the code is gained by computing contributions at the medium and coarse resoutions as much as possible, instead of using the fine resolution grid all the time.

As was pointed out above, the strengths of some of the lines sometimes sometimes make it necessary to use lines that could be upto 200 cm^{-1} away (in the coarse part of the computation). Instead of doing this, a common practice is to remain with the 25 cm^{-1} width of the coarse meshes, but include the effects of lines outside these coarse mesh by adding on an extra *continuum*. This is true for oxygen and nitrogen, and in particular water vapor (see below). Carbon dioxide also requires this; however we choose not to use the continuum in this case, but simply use the entire 200 cm^{-1} wide coarse meshes.

Carbon dioxide, like water vapor, is radiatively active in the Earth's atmosphere. Just like water, it is important to accurately compute the absorption spectra using the correct lineshape, as optical depths of these gases vary greately in the atmosphere, allowing one to probe several layers of the atmosphere within a small spectral region. A thorough understanding of the absorption spectra of these two gases is therefore very important for the remote sensing community.

2.1 Computing water vapor absorption coefficients

Almost all that has been said above remains valid when computing the lineshape of water vapor. However, the lineshape far away from line center is sub-Lorentz ($k/k_{lor} \le 1$), while the lineshape close to line center is super Lorentz ($k/k_{lor} \ge 1$).

To account for this behavior, the above algorithm has to be slightly modified. This leads to the water vapor lineshape algorithm to include the following three different modifications, put together :

- Instead of using a lorentz lineshape, a *local lineshape* is used. Upto 25 cm^{-1} away from line cenetr, this is defined as lorentz less the lorentz value $25cm^{-1}$, and zero everywhere else.
- To include the super Lorentz behavior close to the line center, the local lineshape is then multiplied by a chi (χ) function
- To include the sub Lorentz behavior far from the line center, a continuum function is then added on. This is done after the effects of all lines have been used. Another way of thinking of this continuum is to say that the water lines are quite strong, and thus the computational algorithm should not restricted to using only lines that are at most 25 cm^{-1} wavenumbers away from the spectral region under consideration. But instead of individually using lines that could be upto 200 cm^{-1} away (and modelling their sub Lorentz far wing behavior), the far wing effects of these far lines are all lumped into the continuum.

To summarize, when computing a water vapor optical depth, the code proceeds as in the general case described above, except that it uses a *local* lineshape for each line j, multiplied by a χ function. After all the necessary lines have been included, a *continuum* absorption coefficient is also added on :

$$k(\nu) = k_{continuum}(\nu) + \sum_j k_{local}(\nu, \nu_j)\chi(\nu)$$

which can be rewritten, for the individual lines j

$$k_{local}(\nu, \nu_j) = \begin{cases} (k_{lorentz}(\nu, \nu_j) - k_{lorentz}(\nu, 25 + \nu_j))\chi(\nu) & \text{if } |\Delta\nu| \le 25\text{cm}^{-1} \\ 0 & \text{if } |\Delta\nu| > 25\text{cm}^{-1} \end{cases}$$

where appropriate factors of $\nu \tanh(\beta\nu/2)$ multiply the above coefficients.

The computation for k_{local} proceeds as described in the previous section, *viz.* using fine, medium and coarse meshes.

2.2 Computing carbon dioxide absorption coefficients

Computing the spectral lineshape of carbon dioxide can be quite complicated. There are many bands within which there are lines that are very closely spaced. Collisions have the effect of mixing these lines together, transferring intensity from the line wings to the line centers. Furthermore, the collisions are not instantaneous, but have a finite duration. This also makes the lineshape deviate from Lorentz, especially far from the line centers. Additionally, some of the bands are very strong and have an effect on the absorption spectrum, at quite large distances from their band (line) centers. This third effect can be accounted for by either using a continuum, or by allowing the inclusion of effects of lines that are upto 200 cm^{-1} away from the region of interest.

2.2.1 Line Mixing : probability of mixing (parameter β)

Deviations from the Lorentz lineshape in regions of overlapping spectral lines have been observed in many cases. In particular, large deviations are found in infrared Q-branches, where the spectral lines are very closely spaced. Theories that treat the collisions between molecules as instantaneous will be accurate only in spectral regions close to the line centers.

Within an ensemble of colliding molecules, the Hamiltonian of a single molecule is not a conservative system. The total Hamiltonian of the molecule is given by $H(t) = H_0(t) + H_1(t)$. In this expression, $H_0(t)$ is the Hamiltonian of the molecule without any interaction with its perturbers. $H_0(t)$ therefore has real eigenvalues and its eigenvectors are the stationary states of the molecule. $H_1(t)$ is the Hamiltonian representing the interaction of the molecule with its perturbers, such as the effects of inelastic collisions. In the context of line mixing of a band, the eigenvalues of $H_0(t)$ give the energies (line centers) of the individual transitions within a band.

For a given band, if the off-diagonal elements of H_1 are zero, the lineshape that results is a sum of individual Lorentzians : the model obeys the impact approximation. The role of the off-diagonal elements of the interaction potential is the interaction between spectral transitions. If these off-diagonal elements of H_1 are non-zero, intensity can be transferred from one line to another, with the “amount” of line-mixing determined by the magnitude of the corresponding off-diagonal element of H_1. These elements are proprtional to the **probability** β that spectral intensity is transferred from one line to another. If $\beta = 0$, there is no line mixing and the lineshape that results is Lorentzian; if $\beta \neq 0$, there is line mixing, and the lineshape deviates from a sum of Lorentzians.

Calculating an absorption coefficient for many transitions over a large spectral range is computationally most efficiently performed using

$$k_{mix}(\nu) = \frac{N}{\pi}\mathbf{IM}\left(\mathbf{d} \cdot \mathbf{G}(\nu)^{-1} \cdot \rho \cdot \mathbf{d}\right)$$

where $\mathbf{G} = \nu - \mathbf{H}$ and $\mathbf{H} = \nu_\mathbf{0} + \imath P\mathbf{W}$, where ν_0 are the eigenvalues of H_0 and $H_1 = \imath P\mathbf{W}$ where P is the pressure and $\imath\mathbf{W}$ is the interaction matrix.

$\mathbf{H}$ is diagonalized with a complex matrix $\mathbf{A}$ to get the diagonal matrix $\mathbf{L} = \mathbf{A}^{-1} \cdot \mathbf{H} \cdot \mathbf{A}$. $\mathbf{G}$ is also diagonalized by $\mathbf{A}$ and $k_{mix}(\nu)$ is written as

$$k_{mix}(\nu) = \frac{N}{\pi}\mathbf{IM}\left(\sum_i \frac{(\mathbf{d}\cdot\mathbf{A})_i(\mathbf{A}^{-1}\cdot\rho\cdot\mathbf{d})_i}{\nu - l_i}\right)$$

where l_i are the diagonal elements of $\mathbf{L}$.

Using time-independent perturbation theory, Rosenkranz found the first-order approximation for $k_{mix}(\nu)$ to be

$$k_{1st}(\nu) = \frac{N}{\pi}\sum_j S_j\left(\frac{P\gamma_j + (\nu-\nu_j)PY_j}{(\nu-\nu_j)^2 + (P\gamma_j)^2}\right) \text{ with } Y_j = 2\sum_{k\neq j}\frac{d_k}{d_j}\frac{W_{kj}}{\nu_j - \nu_k}$$

where Y_j are first-order mixing coefficients. For a single transition, this lineshape is the sum of a Lorentzian and an asymmetric term. Far from the line centers, the asymmetric terms become proportional to ν^{-1}. In order for $k_{1st}(\nu)$ to go to zero in these regions, the sum of the coefficients must vanish. That is, detailed balance must be obeyed. In this context, Strow and Reuter[**?**] showed that detailed balance is obeyed if

$$\sum_j S_j Y_j = 0.$$

They used this result to show that, in the far-wing limit, the ratio of mixing and Lorentz absorption coefficients is a constant[**?**]

$$\Omega = \frac{k_{1st}}{k_L} = 1 + \frac{\sum_j S_j Y_j \nu_j}{\sum_j S_j \gamma_j}$$

This is a useful result because it allows the mixing lineshape to be calculated by simply multiplying the Lorentz lineshape by a constant in the far wing. Note that the constant Ω depends on the band, as well as on the pressure and temperature. However, as the interaction matrix $W_{jj'}$ has to be computed in each case, it is easy to obtain the mixing coefficients Y_j and transition amplitudes d_j, from which Ω can be computed.

The only point left to be addressed is the determination of the off-diagonal, or mixing, terms of $\mathbf{W}$. This will be described in more detail below.

2.2.2 Determining the Interaction Matrix W

Empirical scaling laws based on energy changes caused by inelastic collisions are often used to model the interactions. The PEG law models the energetically upward state-to-state inelastic collisional rates as a function of the rotational energy difference, $\Delta E_{j'j}$. An upward rate going from the state j to state j' is modeled as

$$K_{j'j} = a_1 \left(\frac{\mid \Delta E_{j'j} \mid}{B_0} \right)^{-a_2} \exp \left(\frac{-a_3 \mid \Delta E_{j'j} \mid}{kT} \right)$$

where B_0 is the rotational constant and a_1, a_2, and a_3 are adjustable parameters which are discussed below.

Detailed balance is obeyed if

$$K_{jj'}(2j'+1)e^{-\frac{E_{j'}}{kT}} = K_{j'j}(2j+1)e^{-\frac{E_j}{kT}}$$

This relation essentially ensures that energy is conserved and gives the downward rates, $K_{jj'}$:

$$K_{jj'} = K_{j'j}\frac{2j+1}{2j'+1}e^{\frac{\Delta E}{kT}}$$

a_1, a_2, and a_3 above are determined by equating the width of a spectral line to the sum of all of the rates which limit the lifetime of that transition via a least-square fit to the known linewidths, as any rates which shorten the molecule's lifetime in a specific energy state broadens the spectral line. These rates include all of those which occur in either the lower or upper state of the transition. Since vibrational energies are much greater than rotational energies, only collisions between states within the same vibrational level are considered. Therefore, using line widths extracted from experimental data, a_1, a_2, and a_3 are determined by requiring

$$\gamma_j \equiv \mathbf{W_{jj}} = \sum_{j' \neq j} \text{All } K_{j'j} \text{ which limit the transition lifetime.}$$

The details of the exact expression used in calculations depend on the type of transition which is occurring, such as Q branch mixing for a $\Sigma - \Pi$ transition.

The off-diagonal elements of the matrix **W** are then taken to be proportional to the corresponding collisional rates of the matrix **K**. *These are the mixing terms.* For two rotational levels which are energetically close, the collisional rate between them is relatively large and they experience mixing. On the other hand, if two levels are energetically far from each other, the corresponding **K** rates are negligible and no mixing occurs.

Summarizing the calculational procedures, the relaxation rates, $K_{jj'}$, are first determined by adjusting a_1, a_2, and a_3 so that the sum of all relaxation rates which limit the lifetime of a transition equals the known line width. The off-diagonal elements of **W** are then taken to be proportional to the corresponding off-diagonal elements of **K**. The details of this step depend on the symmetry of the band.

The diagonal elements of **W** are equated to the line widths. The absorption coefficients are then calculated using full or first order mixing.

2.2.3 Duration of collisions : Birnbaum chi function (τ parameter)

The Lorentz model ignores the effects of finite durations of collisions, and so has a larger intensity far from line center than that actually measured. To account for the finite duration of collisions, Birnbaum developed a theory which resulted the lineshape being described by a chi function mutiplying the Lorentz lineshape (see Eqn. 68

$$k_B(\nu) = k_L(\nu)\chi_B(\nu) = k_L(\nu)A_m z K_1(z)\exp(\tau\gamma + \tau_0\Delta\nu)$$

with

$$z = \sqrt{(\gamma^2 + \Delta\nu^2)(\tau_0^2 + \tau^2)} \text{ and } \Delta\nu = \nu - \nu_j$$

where $K_1(z)$ is a modified Bessel function of the second kind, $\tau_0 = \frac{0.72}{T}$ [1], and A$_m$ is an adjustible constant (set to 1.0 in the code). T is the temperature, γ the linewidth and and τ is the duration of collision parameter.

This "corrective" factor, $\chi_B(\nu, \gamma, \tau_0(T), \tau)$, removes much of the far-wing absorption of the impact approximation for the Lorentz line shape. For the parameters used in the code (Earth atmosphere), χ_B is only very weakly dependent on the linewidth γ. When computing the Birnbaum function, our code uses a look up table to do an interpolation in temperature/duration-of-collision.

2.2.4 Combined Line-Mixing and Duration-of-Collision Lineshape

Both line-mixing and the duration-of-collision effect have shown to reduce the amount of absorption in the far-wing limit. Since the line-mixing theory is valid only under the impact approximation, the combined effects of line-mixing and duration-of-collision are approximated as if each effect were independent. In order to include the effects of line-mixing over the entire frequency range, we have

$$k(\nu) = \sum_i k_{1st}(\nu_i, \nu)\chi_B(\nu_i, \nu)$$

Note that this lineshape is parameterised by only two parameters : β which is the probability that a collision will transfer intensity from one transition to another, and τ which is the duration of collisions parameter.

if the first-order approximation is too inaccurate (which is seldom true for atmospheric applications), the full line-mixing lineshape $k_{mix}(\nu)$ can also be implemented by using

$$k(\nu) = \frac{k_{mix}(\nu)}{\sum_i k_{Lor}(\nu_i, \nu)} \sum_i k_{Lor}(\nu_i, \nu)\chi_B(\nu_i, \nu)$$

[1] τ_0 and τ have been converted to units of cm by multiplication by 2πc.

However, all of our data fits were performed using the previous expression, which we have chosen to remain with. The two expressions are not interchangeable ie parameters obtained from a fit using the first expression cannot be used in the second expression.

2.3 Cousin versus Birnbaum chi functions

To model the CO2 spectrum in the 4 μm region, Cousin developed a set of empirical chi functions, with many parameters to account for varying temperatures, broadeners and so on. Multiplying the Lorentz lineshape with the suitable Cousin chi function can model the experimental spectrum, epsecially far away from band center and for low optical depths. The Cousin chi functions combine the physics of line mixing and finite duration of collisions. However, these Cousin functions are also used in the 15 μm region of CO2 as well. As mentioned above, the β factor in this region is 0.5, about half that in the 4 μm region. Thus it is incorrect to use the Cousin functions in the 15 μm region. There will be too much intensity transferred to line centers, which means the absorption in between lines is too weak, leading to lower optical depths and higher transmissions. As remote sensing algorithms frequently use the information "in between" lines to retrieve atmospheric properties, this could certainly be a source of errors in the retrievals.

If the user asks for "birnbaum" chi function to be used, our code "blends" in the Cousin and Birnbaum chi functions. Close to the lines (typically within 25 cm^{-1} for the strong bands, and within 1-2 cm^{-1} for the weaker bands), the Birnbaum chi function is used. Outside of this region, the Cousin function is turned on gradually, till it is fully on a distance 1 cm^{-1} away from this edge. This birnbaum chi function can be used whether or not line mixing is turned on or not.

If the user asks for "cousin" chi function to be used, our code just turns on the Cousin chi function only. Because the Cousin function includes the effects of line mixing, this feature can only be used if line mixing is turned off.

2.4 Computing the line-mixing lineshape

We have identified many of the strong P,Q,R bands in the 4 and 15 μm regions, where line mixing computations are performed. For all these bands, the code used to compute the line mixing lineshape is almost the same and is therefore common. As mentioned earlier, the main difference comes in the setting up of the interaction matrices **W**, whiere the details of the symmetry peculiar to the band under consideration have to be explicitly implemeted.

In this section, we outline the steps in the computation of the $\Sigma - \Pi$ **Q** branch lineshape (eg **Q 667**). Almost all the steps will be applicable to he other bands.

Given the gas amount U in $kiloMoles/cm^2$, the self pressure and temperature, the code first finds the path length L.

2.4.1 Loading in line parameters, storing β, τ

It then calls subroutine $loader$ which performs a number of initializations. For instance, it computes the line mixing strength

$$\beta = \frac{\beta_{self} p_{self} + \beta_{for} p_{for}}{p_{total}}$$

where $p_{total} = p_{self} + p_{for}$. Similarly it computes the duration of collision parameter (usually not needed for **Q** branches, as the lines are all very close together)

$$\tau = \frac{\tau_{self} p_{self} + \tau_{for} p_{for}}{p_{total}}$$

The code then initialises other constants, and loads in the relevant HITRAN parameters for the band in question, such as line strengths and widths. It then adjusts the line widths for the pressures and temperatures in question, as well as initialising the partition fiunctions and computing the temperature adjusted line strengths. For the given band, the rotational quantum indices and associated parameters (widths, strengths etc) are all ordered in terms of increasing J, using routine $orderer$.

2.4.2 Computing lower state rotational energies E_{lower}

Routine $efitter$ is then called, to compute the rotational energies of any missing levels. When interaction matrix **W** is built up, a complete knowledge of lower and upper state rotational energies is required, so that state to state collision rates can be computed. However, this may require the code to know the rotational energies of some missing levels. To circumvent this, a three parameter least squares fit of the lower state energies $elower$ versus rotational number j is performed :

$$E_{lower} = B \times j(j+1) - D \times (j(j+1))^2 + E_{vib}$$

For the **Q667** branch, the even j's from 2 to 102 are stored on the HITRAN tape, as are the associated lower state energies E_{lower} The least squares fit thus gives an estimate for parameters B, D, E_{vib}. Having obtained these parameters, using the above equation the code can easilty compute the energies of missing levels if necessary – in this case, the odd j's and $j = 0$. The final step is to save in memory the rotational energies $B \times j(j+1) - D \times (j(j+1))^2$ for the $j = 0, 1, 2, 3, ...$

2.4.3 Computing relaxation matrix K using PEG law

Now comes the set of routines which actually use the above information to set up the interaction matrix **W**. As mentioned earlier, the physics of the band in question is used in setting up this matrix. Having computed the rotational energies of the lower states as described above, a set of routines now fit the foreign broadening widths $\gamma_{for}(j)$ to the three parameters $a_1^{for}, a_2^{for}, a_3^{for}$ of the Power Exponential Gap (PEG) equations, and then similarly the self broadening widths $\gamma_{self}(j)$ to the corresponding three parameters $a_1^{self}, a_2^{self}, a_3^{self}$. The procedure is described below.

As mentioned earlier, the PEG law models the energetically upward state-to-state inelastic collisional rates as a function of the rotational energy difference, $\Delta E_{j'j} = E_{j'} - Ej$. First set up matrix $\mathbf{\Delta E}$ which governs the energy transitions from lower state j to upper state j' :

$$\mathbf{\Delta E} = E_{j'j} = \begin{pmatrix} 0 & E_{j2} - E_{j1} & E_{j3} - E_{j1} & E_{j4} - E_{j1} \\ E_{j1} - E_{j2} & 0 & E_{j3} - E_{j2} & E_{j4} - E_{j2} \\ E_{j1} - E_{j3} & E_{j2} - E_{j3} & 0 & E_{j4} - E_{j3} \\ E_{j1} - E_{j4} & E_{j2} - E_{j4} & E_{j3} - E_{j4} & 0 \end{pmatrix}$$

Note that if one takes the lower triangular diagonal elements of the above matrix, one has the lower to upper state transitions $j \to j'$, while the upper triangular diagonal elements of the above matrix yield the upper to lower state transitions $j' \to j$.

Using matrix $\mathbf{\Delta E}$, the matrix containing the upward relaxation rates going from lower state j to upper state j' is then

$$K_{j'j} = a_1 \left(\frac{\mid \Delta E_{j'j} \mid}{B_0} \right)^{-a_2} \exp \left(\frac{-a_3 \mid \Delta E_{j'j} \mid}{kT} \right)$$

where B_0 is the rotational constant and a_1, a_2, and a_3 are adjustable parameters.

Detailed balance is obeyed when the upward transition rate is equal to the downward transition rate, giving the downward rates in another matrix :

$$K_{jj'} = K_{j'j} \frac{2j+1}{2j'+1} e^{\frac{\Delta E}{kT}}$$

Having constructed matrix $K_{j'j}$, *we can now construct the relaxation rate matrix* **K** to fit for parameters $a_l, l = 1, 2, 3$. We can put all this information into one matrix, where the upper triangular part contains the downward rates while the lower triangular partion contains the upward transition rates. This can be more conveniently expressed by using the lower triangular part of $K_{j'j}$ as is (for the upward transition rates), and using detailed balance to rewrite the the upper triangular part of $K_{j'j}$ for the downward transition rates

$$K^{rate}_{j'j} = K^{lowertriang}_{j'j} + K^{lowertriang}_{j'j} \frac{2j+1}{2j'+1} e^{\frac{\Delta E(j'j)}{kT}}$$

The upper diagonals have the downward relaxation rates, while the upper relaxation rates are in the lower diagonals. This matrix now has all the elements that can interact with one another, as well as extra elements which may not be needed (remember $efitter$ precomputed the rotational energies of any missing j levels, in case they are required in what follows).

Recall that the relaxation rates $\mathbf{K_{jj'}}$ are determined by adjusting a_1, a_2, a_3, so that the sum of the relaxation rates that limit the lifetime of a transition equal the linewidth. Upto this point, the discussion above has been quite general and applicable to all the bands. Now however, the particular details of the band become important. For the **Q667** branch, the line widths are given by

$$W_{jj} \equiv \gamma_j = -\frac{1}{2}\left(\sum_{j' \neq j}^{even} K^{\Sigma(e\leftarrow e)}_{j'j}\right) - \frac{1}{2}\left(\sum_{j' \neq j}^{even} \beta K^{\Pi(f\leftarrow f)}_{j'j} + \sum_{j' \neq j}^{odd} (1-\beta) K^{\Pi(e\leftarrow f)}_{j'j}\right)$$

where the β probability factor, introduced by Strow and Edwards [**?**] differentiates between the collisional relaxation rates which connect states of similar and opposite rotational symmetry. As mentioned earlier, β is a measure of the probability that a collision transfer intensity from one transition to another. For each of the bands for which line mixing is done, we have experimentally determined β. For the **Q667** band, the superscript on $K_{j'j}$ denotes the symmetry of the levels involved. In the Π vibrational level, the relaxation rates have been divided into two groups. Those which connect rotational levels of similar symmetry are multiplied by β and those which connect levels of opposite symmetry are multiplied by $(1 - \beta)$. The β factor is not needed in the Σ state simply because the odd rotational levels do not exist there. For the **Q667** $\Sigma - \Pi$ band, a cartoon of the allowed transitions, and the lines to which they can mix, is shown in Fig. 2.

The user is referred to Tobin's thesis [**?**] for implementation details of the other bands.

2.4.4 Computing the interaction matrix W

By using a least squares fitting procedure, parameters a_1, a_2, a_3 can be adjusted so that the sum of the relaxation rates on the RHS of the previous equation are almost equal to the actual line width $W_{jj} \equiv \gamma_{Qj}$. Using these fitted parameter values, the

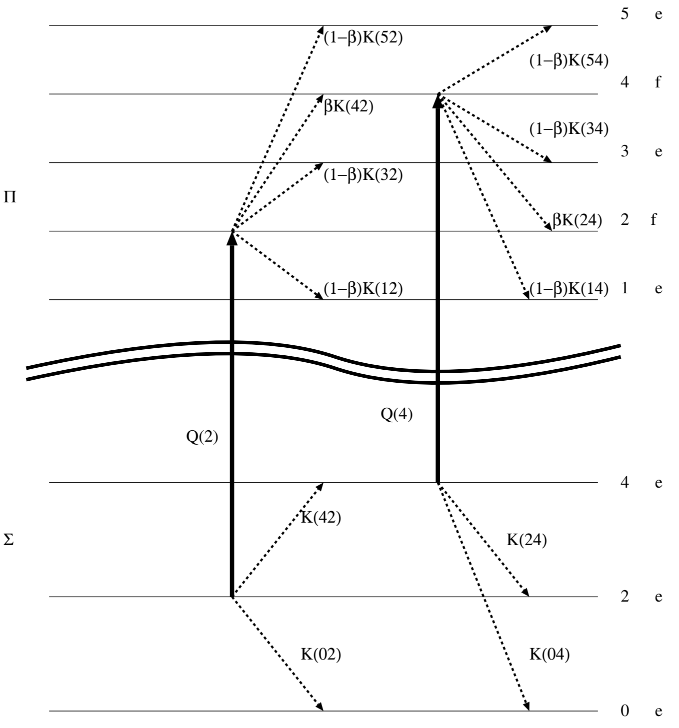

Figure 2: Rotational relaxation rates, $K_{j'j}$, for several Q-branch lines of a $\Sigma \leftarrow \Pi$ transition. In our line-mixing model, Π rates connecting rotational levels of similar parity are multiplied by β and rates connecting levels of opposite parity are multiplied by $(1-\beta)$.

off diagonal elements of $K_{j'j}$ are recomputed using the PEG scaling law

$$K_{j'j} = a_1 \left(\frac{\mid \Delta E_{j'j} \mid}{B_0} \right)^{-a_2} \exp\left(\frac{-a_3 \mid \Delta E_{j'j} \mid}{kT} \right)$$

while the diagonal elements are equated to the linewidths $K_{jj} = \gamma_j$. This is done separately for the self and foreign broadened widths. For the pressures under consideration, the complete relaxation matrix is then a weighted sum of the two relaxation matrices

$$\mathbf{K} = \frac{\mathbf{K_{for}} p_{for} + \mathbf{K_{self}} p_{self}}{p_{for} + p_{self}}$$

The diagonal elements of interaction Hamilitonian are the linewidths of **K.** If there is no line mixing, the mechanics of the algorithm will just give a Lorentzian lineshape.

The off diagonal elements of the perturbation or interaction matrix are then simply proportional to the off diagonal elements of the relaxation matrix; if there is no line mixing ($\beta = 0$), these elements are zero.

Combining the above two Hamiltonians, the interaction matrix $\mathbf{W} \simeq H_1$ is

$$\mathbf{W} = diag(\mathbf{K}) + \beta \times offdiag(\mathbf{K})$$

2.4.5 Computing the transition population amplitudes

In order to compute the full line mixing lineshape, a knowledge of the transition population amplitudes is necessary. These amplitudes are computed in $trans\_pop$

The transition amplitude is computed using

$$d = \sqrt{\frac{S}{\rho}}$$

where d is the transition amplitude (which is the square root of the transition probability), S is the line strength and ρ is the population.

For the isotope in questions, the partition function coefficients are read in a, b, c, d (obtained from qtips.f). This enables the computation of the partition functions both at the reference and actual temperatures, $Z(296), Z(T)$. For a line of rotational quantum number j, centered at $\nu_0(j)$ with lower rotational energy $E_{lower}(j)$ and strength S_j, the reference temperature transition amplitude d is

$$\alpha(296) = \frac{8\pi^3}{3} \frac{\nu_0(j)(1 - e^{-c_2 \nu_0(j)/296})}{hc}$$

$$\beta(296) = (2j + 1) \times e^{-c_2 E_{lower}(j)/296} \times 1e^{-36} / Z(296)$$

$$d = \sqrt{\frac{S_j}{\alpha(296) \times \beta(296) \times 10^{-7}}}$$

The population ρ at the desired temperature can similarly be computed using temperature T instead of $T_{ref} = 296$ in the expressions for α, β above to get

$$\rho = \alpha(T) \times \beta(T) \times 10^{-7}$$

The powers of 10 in the above expressions are 10^{-36} for a $debye^2$ unit in terms of $ergscm^3$, and a factor of 10^{-7} to change Planck's constant from $J.s$ to $erg.s$

2.4.6 Computing the full mixing lineshape

The complete Hamiltonian is written as

$$H = H_0 + H_1 = diag(\nu_0) + \imath W$$

from which the eigenvalues l_j and eigenvectors $\mathbf{A_j}$ of H are found. Having done this, and having already computed the transition amplitudes d_j (dipole moment matrix emlements) and density ρ_j (matrix elemtns that represent population difference between lower and upper levels), it is now straightforward to loop over the individual lines and compute the total lineshape. Note that if necessary, a birnbaum χ function would have been mutiplied into each contribution :

$$k^{optdepth}_{mix(j)}(\nu) = \frac{N}{\pi}\mathbf{IM}\left(\frac{(\mathbf{d}\cdot\mathbf{A})_j(\mathbf{A^{-1}}\cdot\rho\cdot\mathbf{d})_j}{\nu - l_j}\right)\times\chi_{birnbaum}(\gamma_j, T, p, ps)$$

$$k^{optdepth}_{mix}(\nu) = \sum_j k^{optdepth}_{mix(j)}(\nu)$$

where $N = \frac{ps}{p_{ref}}\frac{T_{ref}}{T}\frac{L\mu}{\pi}$, L being the path length and μ the molecular density.

Recall that if there is no linemixing, the lineshape should just be a sum of Lorentzians. Thus if H_1 has no off diagonal elements ie its elements consists entirely of the linewisths on the diagonal, then the eigenvalues of H would be the line centers $\nu_0(j)$, and the spectrum would end up being a sum of Lorentzians, with line center $\nu_0(j)$ and line width γ_j. If there *is* linemixing, then the linecenters would still be close to $\nu_0(j)$, but the lineshape would deviate from a sum of lorentzians.

2.4.7 Computing the first order mixing lineshape

The first-order approximation for $k_{mix}(\nu)$ is

$$k_{1st}(\nu) = \frac{N}{\pi}\sum_j S_j\left(\frac{\gamma_j + (\nu-\nu_j)Y_j}{(\nu-\nu_j)^2 + (\gamma_j)^2}\right) \text{ with } Y_j = 2\sum_{k\neq j}\frac{d_k}{d_j}\frac{W_{kj}}{\nu_j - \nu_k}$$

where, as above, $N = \frac{ps}{p_{ref}}\frac{T_{ref}}{T}\frac{L\mu}{\pi}$, S_j is the line strength, and the mixing coefficients Y_j are obtained from the interaction matrix $\mathbf{W}$ and transition amplitudes d_j.

2.5 Implementing the line mixing/duration of collision algorithm

Assume the code is about to compute the contribution of band B to the absorption spectrum. For any input wavenumber interval, this could be done by following the prescription above. However, the code could be sensitive to the numerics of the matrix diagonalization and inversion, especially far from the band. As mentioned above, far from the band center, the ratio of line mixing and Lorentz absorption coefficients is a constant Ω, that depends on the band, temperature and partial and total pressures. In this far away region, it is numerically more appropriate to use this ratio, than to expect the code to behave correctly.

In addition, when fitting experimental data to the line mixing/ Birnbaum chi functions, the emphasis was on the region close to the 4.3 μm bandhead, and in between the lines. On the other hand, experimentally, far from band center, we see that the Cousin lineshape describes the observed experimental spectra quite well. Combining the above two, to compute the lineshape far from line centre, it is apropriate to just use the Cousin chi function to multiply a Lorentz or Voigt lineshape.

For these reasons, when computing the contribution of the current band B to the absoption coefficient, the code divides the input wavenumber region into 5 intervals. Assume the wavenumbers of the band edges are b_1, b_2. In the region f_1, f_2 where $f_1 = b_1 - \theta$, $f_2 = b_2 + \theta$, full mixing is done. θ varies according to the band, and is a variable that tells the code where full mixing should be computed. For the 4.3 μm band, we have the following values :

Band	θ (cm^{-1})	Description
2350	40	Strong $\Sigma - \Sigma$ band
2310	0.025	Strong $\Pi - \Pi$ band
2320	0.025	Strong $\Delta - \Delta$ band
2352	0.025	Weak $\Sigma - \Sigma$ band
2353	0.025	Weak $\Sigma - \Sigma$ band
2354	0.025	Weak $\Sigma - \Sigma$ band
others	15	Rest of the 4 μm bands

while for the 15 μm band, we have the following values :

Band	θ (cm^{-1})	Description
667	5	Strong $\Sigma - \Pi$ band
others	15	Rest of the 15 μm bands

One wavenumber away from f_1, f_2, the linemixing/Birnbaum lineshape is turned off, and Cousin is turned on (ie $f_1 - f_3 = f_4 - f_2 = 1cm^{-1}$, where in the region to the left of f_3 and to the right of f_4, the lineshape is $k_{lor} \times \chi_{Cousin}$

The wavenumber regions are illustrated below

```
  -------|------------|--------|==========|--------|-------------|--
--------
         |            |       b1         b2       |             |
        f3            f1                          f2            f4
         |            |                           |             |
  cousin | full/cousin|          full             | full/cousin |  cousin
         |            |                           |             |
   I     |     IV     |          III              |     V       |    II
```

- : Regions I,II : these regions are far away from the band edges. If the Birnbaum function was being used as the chi function, it is turned off and the Cousin chi function is turned on.
- : Regions III : this region is at the band center itself. Full or first order line mixing as described above, is computed. In addition, the Birnbaum chi function is used, if the user has specified it is to be used.
- : Regions IV,V : these are intermediate distance regions. To smoothly blend on the complete line mixing/birnbaum computation onto the Cousin chi function, a linear mix of the above two is done

Note that the code can perform full linemixing in regions III,IV,V above, or just use first order linemixing. The $k = k_{lor} \times \theta$ algorithm is implememted in regions IV,V, if first order linemixing is specified.

In addition, the full mixing lineshape theory is developed using Lorentz lines. The lineshape in the lower atmosphere ($p > 0.158 atm$) is Lorentz. However, the lineshape in the upper atmosphere ($p < 0.021 atm$) is turning towards a Doppler profile. Thus it is incorrect to use the full line mixing lineshape at lower pressures. A linear blending of the lineshape, from full to first order, is performed by the code, illustrated by the cartoon below.

```
     -------------------------- TOA

          first order

     -------------------------- 2.065 e-2 atm
```

```
          blend

---------------------------- 1.584 e-1 atm

          full mix

---------------------------- Ground
```

Thus a combination of the above two blends are done - in terms of wavenumbers away from the band center, and total pressure. Note for the 15 μm band, PR linemixing is done only for the 720 $\Sigma - \Pi$ and 741 $\Delta - \Pi$ bands. For the other bands, no PR linemixing is done.

3 run6

The Matlab code has four main driver files : $run6.m$, $run6co2.m$ and $run6water.m$ and $run6watercontiunuum.m$.

$run6.m$ is a general code that will work for all gases; however one should run the specialized codes for water and CO2, so as to to utilise the above physics in the computed lineshapes. We describe the $run6.m$ program parameters and algorithm in detail below; in the next two sections, we will discuss the corresponding similarities and differences for the water and carbon dioxide codes.

The input argument list to all codes also includes the name of a profile file which specifies the layer number, total pressure, gas partial pressures (both in atm), gas temperature (in Kelvin) and gas amount (in $kilomoleculescm^{-2}$). The profile should be in a 5 column format, and should be a text file.

3.1 Mex files and HITRAN database

(this is for run6)
If the user wants to change the name of the line database file that is used, he/she will have to go into the run6* files and change the name of the file in the line beginning with the word $fnamePRE$, which is currently set to :

```
fnamePRE='/salsify/scratch4/h98.by.gas/g';
```

(this is for run7)
The default is to use HITRAN2000; if the user wants to change this, all that has to be done is supply an additional input parameters, topt.HITTRAN = xxxx (see below for details!!!)

To speed the code up, a number of loops have been written as fortran MEX files. All these files are in subdirectory FORTRANFILES, and assume input arrays/matrices that are smaller than certain limits. If the user wants to change these limits, he/she will have to edit the file $max.inc$ and recompile the Mex files.

```
c this is max length of arrays that can be used in the Mex Files
c this number came out of
c   200000 = max number of elements in mesh
c               eg (755-655)/0.0005 = 160000
c        4 = number tacked on to arrays so boxint(y,5) can be done
      integer MaxLen
      parameter(MaxLen=200010)

c assume max number of any of P,Q,R lines = 300
```

```
      integer MaxPQR
      parameter(MaxPQR=300)

c assume max number of any of layers = 100
      integer kMaxLayer
      parameter(kMaxLayer=100)
```

To compile the Mex files, the user has to type $makemex1$ at the UNIX prompt (if only $run6/7.m$ is being used), or type $makemex$ (if $run6/7co2.m$ will be used). This compiles all the Mex files, and creates symbolic links to these files from the necessary subdirectories.

If the user is going to use $run6/7co2.m$, he/she will also need to go to the C02_COMMON subdirectory, and type $link.sc$ so that symbolic links from the CO2 subdirectories to the common files are created.

3.2 Water, nitrogen, oxygen continuum

Through parameter CKD (see below), the user can toggle the continuum calculation on/off for three gases : water, oxygen and nitrogen (gasIDs 1,7,22 respectively).

- Water : CKD can be set to -1 (no continuum), or 0,21,23,24 for CKD versions 0, 2.1, 2.3, 2.4. Note when the first 3 versions of CKD are included, the computation proceeds by using the code which does not require a "local" lineshape. However, CKD 2.4 does require a "local" lineshape.
- Oxygen : CKD can be set to -1 (no continuum), or +1 (continuum)
- Nitrogen : CKD can be set to -1 (no continuum), or +1 (continuum)
- For all other gases, the value of CKD is irrelevant

3.3 run6.m input parameters

A typical call to run6 would involve sending in the following :

$[outwave, outarray] = run6(gasID, fmin, fmax, ffin, fmed, fcor, fstep$ $xnear, xmed, xfar, nbox, strfar, strnear, LVG, CKD, profile)$

where the right hand side variables would be

TYPE	VAR	DESCRIPTION	TYPICAL VALUE

integer	gasID	HITRAN gas ID	3
integer	fmin	minimum freq (cm-1)	605
integer	fmax	maximum freq (cm-1)	630
real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0
real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	25.0
integer	nbox	boxcar sum size (odd integer)	1,5
real	strfar	min line strength for far wing lines	
real	strnear	min line strength for near wing lines	
char	LVG	(L)orentz,Voi(G)t,(V)anHuber	'V'
integer	CKD	continuum no = -1 (most gases) yes for water (0,21,23,24) yes for N2, O2 (gases 7,22)	-1
matrix	profname	Nx5 matrix (gasID,pressure,pp,temp,amt) pressure, partial pressure in atm, T in K gas amount in kilomolecules/cm2	

The output arguments from the function call are the output wavevector, outwave, and the computed line spectra in outarray. The vector outwave (and thus the output array outarray) spans the wavenumber range from $fmin$ to $fmax - (ffin \times nbox)$, at a resolution of $ffin \times nbox$.

3.4 run7.m input parameters

The same parameters as above are used. However, many of them are now defaulted to preset values, and can be reset by using optional input argument $topts$, where $topts$ is a structure.

A typical call to run7 would involve sending in the following :
$[outwave, outarray] = run7(gasID, fmin, fmax, profile, \{topts\})$;
where the required right hand side variables would be

TYPE	REQUIRED	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	3
integer	fmin	minimum freq (cm-1)	605
integer	fmax	maximum freq (cm-1)	630

matrix	profname	Nx5 matrix (gasID,pressure,pp,temp,amt) pressure, partial pressure in atm, T in K gas amount in kilomolecules/cm2

and the optional right hand side arguments would be sent in structure $topts$; the default values are as shown :

TYPE	OPTIONAL	DESCRIPTION	DEFAULT VALUE
real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0
real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	25.0
integer	nbox	boxcar sum size (odd integer)	5
real	strfar	min line strength for far wing lines	0.0
real	strnear	min line strength for near wing lines	0.0
char	LVG	lineshape : (L)orentz,Voi(G)t,(V)anHuber	'V'
string	HITRAN	path to HITRAN database	/asl/data/hitran/h2k.by.gas
integer	CKD	continuum no = -1 (most gases) yes for water (0,21,23,24) yes for N2, O2 (gases 7,22)	-1

The output arguments from the function call are the output wavevector, outwave, and the computed line spectra in outarray. The vector outwave (and thus the output array outarray) spans the wavenumber range from $fmin$ to $fmax - (ffin \times nbox)$, at a resolution of $ffin \times nbox$.

3.5 Detailed description of the input parameters

Two of the input parameters are self-describing. The first parameter, $gasID$ is an integer value specifying which gas you want to compute the line spectra for. This integer value is the same as that used for the HITRAN database; for example gasID=3 corresponds to ozone. LVG is a character parameter that tells the code which lineshape to use for *all* the lines. Of the lineshapes described previously, our code can compute one of the following three - **L**orentz, Voi**G**t or **V**anVleck-Huber. The VanVleck-Huber is computed with a Voigt lineshape, and is the one we recommend; to use this lineshape, LVG is set to 'V'.

When the code starts running, it loads in all lines whose centers lie between $fmin - xfar$ and $fmax + xfar$, and whose database line strength is greater than min(strfar,strnear), as the user assumes these are the lines which will have a discernible effect on the overall spectra. Using these lines, and their associated parameters, the computations are performed on a fine mesh resolution $ffin$ and then boxcar averaged to an output resolution $ffin \times nbox$. The results of the computations are output for a wavector that spans $fmin$ to $fmax - ffin \times nbox$. Internally, the computations are essentially performed on a fine mesh that spans $fmin - (nbox - 1)/2$ to $fmax - ffin \times nbox + (nbox - 1)/2$. In this way, the boxcar averaging can be done on the endpoints.

If this direct method were used, depending in the gasID and wavenumber region chosen, the code could be agonizingly slow. In the interests of speed (and mantaining the accuracy), the code therefore requires some more parameters to be sent in.

With these additional parameters, the output wavevector $fmin$ to $fmax - ffin \times nbox$ is divided into equal sized "wide meshes" of size $fstep\ cm^{-1}$. Thus there are N wide meshes, where

$$N = \frac{fmax - fmin}{fstep} \tag{1}$$

Suppose we are considering the ith widemesh, and we denote the start frequency of this widemesh by $f1$, and the stop frequency by $f2$. These two numbers are related to each other and to the other numbers by

$$f1 = fmin + (i - 1) \times fstep - ffin \times (nbox - 1)/2$$

$$f2 = fmin + ii \times fstep - (ffin \times nbox) + ffin \times (nbox - 1)/2 \tag{2}$$

This "finemesh" thus spans $(f1, f2)$ at the fine resolution of $ffin\ cm^{-1}$

Associated with this finemesh is a medium resolution mesh, that spans $(f3, f4)$ at a coarser resolution $fmed$ where

$$f3 = fmin + (i - 1) \times fstep$$

$$f4 = fmin + ii \times fstep \tag{3}$$

In addition there is a coarse resolution mesh, that spans $(f3, f4)$ at a coarsest resolution $fcor$ where $f3, f5$ are the same, as are $f4, f6$:

$$f5 = fmin + (i - 1) \times fstep$$

$$f6 = fmin + ii \times fstep \tag{4}$$

The spectral region of the output wavevector that corresponds to these three meshes is essentially $f3, f4$, adjusted for the last point. In other words, the ith output region $fout(i)$ spans $f3$ to $f4 - nbox \times ffin$, at a resolution of $nbox \times ffin$

For any of the N widemeshes, lines are grouped into three categories, depending where they fall within the three categories defined below :
(a) near lines are those whose line centers lie in the wavenumber interval $(f3 - xnear, f4+xnear) = (w1, w2)$. All computations using these lines are performed on the fine mesh (spanning $f1, f2$) of point spacing ffin, and then boxcar averaged to the output wavevector $fout(i)$
(b) medium lines are those whose line centers lie in the wavenumber interval $(f3 - xnear - xmed, f3 - xnear) \cup (f4 + xnear, f4 + xnear + xmed) = (w3, w1) \cup (w2, w4)$. All computations using these lines are performed on the medium mesh (spanning $f3, f4$) of point spacing fmed, and then splined to the output wavevector $fout(i)$
(c) far lines are those whose line centers lie in the wavenumber interval $(f3 - xmed, f3 - xfar) \cup (f4 + xmed, f4 + xfar) = (w5, w3) \cup (w4, w6)$. All computations using these lines are performed on the coarse mesh (spanning $f3, f4$) of point spacing fcor, and then splined to the output wavevector $fout(i)$
The cartoon in Figure **??** summarizes the above relationships.

With the above description, the following restrictions on the parameters are now self explanatory :
(1) xnear $\leq$ xmed $\leq$ xfar
(2) xnear $\geq$ fstep
(3) xmed/fmed xnear/fmin fstep/fmed fstep/ffin are integers
(4)fstep/(nbox*ffin) fcor/ffin are integers
(5)(fmax-fmin)/fstep (fmax-fmin)/fcor are integers

A useful rule of thumb is that ffin,fmed,fcor are chosen so that they are all equal to $(1/2)/10^n$ $n \leq 5$, with n chosen as necesary for the three parameters. For example, $n = 3, 1, 0$ gives $ffin = 0.0005, fmed = 0.05, fcor = 0.5cm^{-1}$.

The above algorithm used is almost the same as that used by GENLN2, except that GENLN2 does not currently have the medium resoltion mesh ie the overall lineshape is a sum of boxcar averaged fine mesh contribution and a spline computed coarse mesh contribution.

Other differences found between this LBL code and GENLN2 is that all computations here are in real*8, while GENLN2 mixes bewteen real*8 and real*4. The discrepancies between these two representations is noticeable in computations of eg the partition fucntions. In addition, we believe that the contribution of a line whose center is in the "far line" regime, is incorrectly splined at the last interval

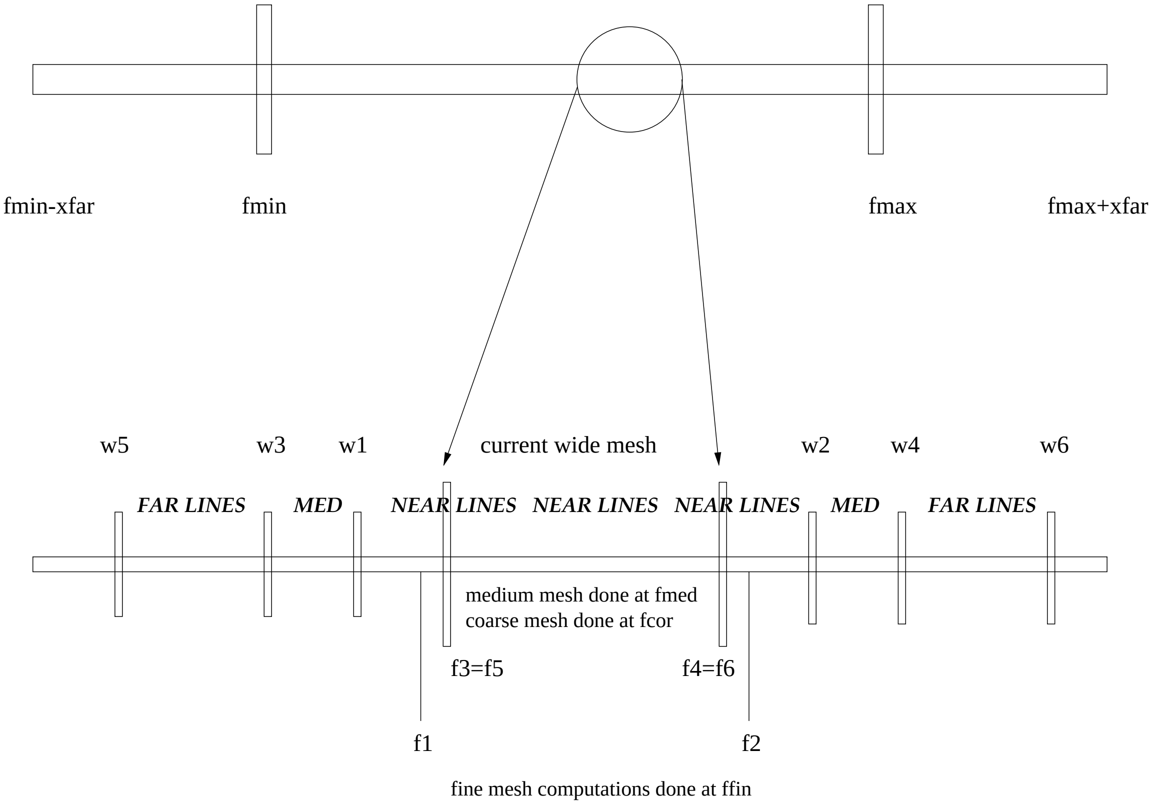

Figure 3:

$|x_{center} - x| \sim xfar$

3.6 Detailed description of the output parameters

Two parameters are passed out after running the code : a 1d array $outwave$ that contains the output wavevector, and a 2d matrix $outarray$ that contains the computed lineshapes at the user set atmospheric levels.

3.7 Outline of the algorithm

The code starts out by checking to ensure that the input parameters make sense and that they are self consistent. For example, parameter LVG must be set to one of the allowable line shapes. In addition, the parameters should all be self consistent in that they have to staisfy the restrictions given at the end of the previous subsection.

Having ascertained the self consistency of the parameters sent in by the user, the program loads in the required mass isotopes for the chosen gas. For example ozone has 5 isotopes.

The program then loads in the user specified profile for the gas. Having done this, the program then uses $fmin, fmax, ffin, nbox$ to define the output wave vector. After this, the gas initializes the $qtipts$ coefficients that are used to compute the partition functions. This is essentially a GENLN2 subroutine, similar to the program "tips" by R.R.Gamache.

The program is now ready to load in the gas line parameters from the HITRAN database. As described above, it loads in all lines whose centers lie between $fmin - xfar$ and $fmax + xfar$, and whose database line strength is greater than min(strfar,strnear).

At this point, the program is almost ready to start running in earnest. Before doing that, it computes the number of wide meshes N and the number of points in each wide mesh that will be mapped to the output wavevector. If the user loaded in a profile that has more than one layer in it, the program calls subroutine $doUnion2$, that computes the optical depth of each linecenter for the chosen profile conditions; if a line is strong enough to be used in any *one* of the levels, it will be used at *all* levels. This will ensure that the optical depth profiles are smooth. An example of this is the case of ozone, where lines could "turn-on" high in the atmosphere, but have almost no optical depth lower in the atmosphere. The importance of this is when the output from the code is used to generate the kCARTA database using Singular Value Decomposition; the SVD algorithm would work more efficiently with smoothly varying data (achieve better compression).

Figure 4 outlines the above initialisation stages of the algorithm.

The program is now ready to loop over the far,medium and near lines. For each of the wide meshes, the program first defines the fine, medium and coarse meshes (frequencies and indices), as described in the previous section. It then sorts all the lines it has loaded into three bins; near, medium and far, also as described in the previous section.

It then enters a loop over layers. For the current layer, the program uses the gas profile to determine the gas amount, temperature, total and self pressures. For each layer, it first computes the contribution due to the near lines, then the medium lines and finally the far lines. The near line spectrum is computed on the fine mesh, and the results are boxcar averaged and added onto the output array. The medium line spectrum is computed on the medium mesh, spline interpolated onto the output wavevector and added on to the output array.The far line spectrum is computed on the coarse mesh, spline interpolated onto the output wavevector and added on to the output array.

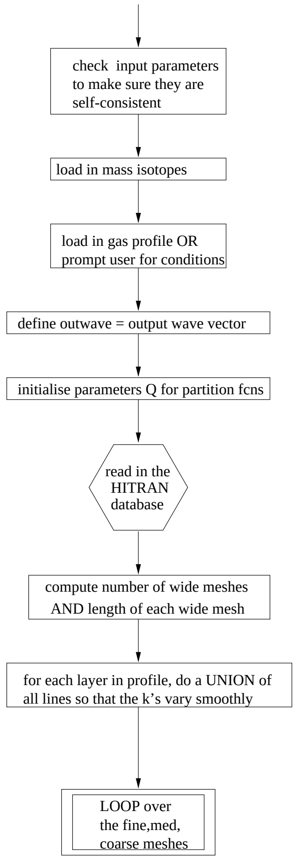

Figure 4:

For each of the fine,medium and coarse computations, the code computes the following line parameters, for each of the lines
(a) the partition function, using $qfcn = q(A, B, C, D, G, lines, tempr)$
(b) the line center frequency, taking the pressure of the current layer into account $freq = lines.ZWNUM + press(jj) * lines.ZTSP$
(c) the overall broadening of the line, using the self and foreign broadening contributions $brd = broad(p, ps, 1.0, forbrd, selfbrd, pwr, tempr, gasID)$
(d) the line center line strength, using the necessary layer parameters such as temperature, gas amount and necessary line parameters
$strength = find_stren(qfcn, freq, tempr, energy, s0, GasAmt(jj))$
The above computations are essentially GENLN2 routines.

Figure 5 outlines the loop stage of the algorithm.

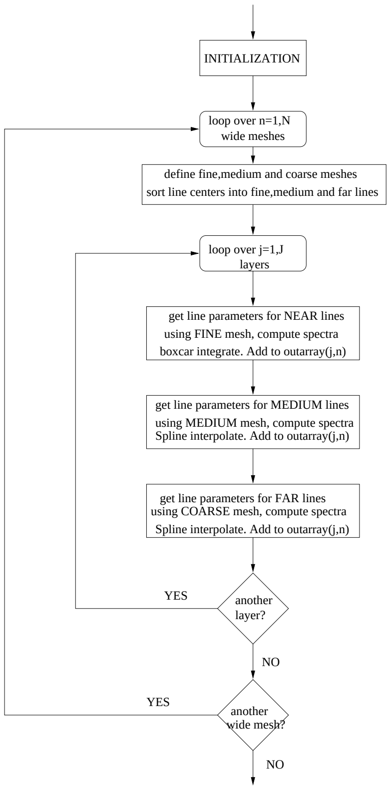

Figure 5:

4 run6water

run6water.m is a specialised code for H2O, so as to to utilise the above physics, namely local lineshape and the CKD continuum effects in the computed lineshapes. If the user simply wants to do a Lorentz or Voigt computation, then it would behoove him/her to use $run6.m$ instead of this special code.

4.1 run6water.m input parameters

A typical call to run6water would involve sending in the following :

$[outwave, outarray] = run6water(gasID, fmin, fmax, ffin, fmed, fcor, fstep, xnear, xmed, xfar, nbox, strfar, strnear, LVF, CKD, selfmult, formult, usetoth, local, profname);$

where the right hand side variables are the same as those for run6 described above; there are 5 new variables on the right side.

TYPE	VAR	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	1
integer	fmin	minimum freq (cm-1)	705
integer	fmax	maximum freq (cm-1)	730
real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0
real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	150.0
integer	nbox	boxcar sum size (odd integer)	1,5
real	strfar	min line strength for far wing lines	
real	strnear	min line strength for near wing lines	
char	LVG	(L)orentz,Voi(G)t,(V)anHuber	'V'
integer	CKD	continumm no (-1)	-1
		yes water : (0,21,23,24)	
real	selfmult	multiplier for self part of contiuum	0¡x¡1
	formult	multiplier for for part of contiuum	0¡x¡1
integer	usetoth	use Toth or HITRAN	+1 to use Toth
			-1 to use HITRAN
integer	local	use local lineshape	+1 to use local*chi defn
			0 to use local defn
			-1 to use run6 defn

The output arguments from the function call are once again the output wavevector, outwave, and the computed line spectra in outarray. The vector outwave (and thus the output array outarray) spans the wavenumber range from $fmin$ to $fmax - (ffin \times nbox)$, at a resolution of $ffin \times nbox$.

4.2 run7water.m input parameters

The same parameters as above are used. However, many of them are now defaulted to preset values, and can be reset by using optional input argument $topts$, where $topts$ is a structure.

A typical call to run7 would involve sending in the following :

$[outwave, outarray] = run7water(gasID, fmin, fmax, profile, \{topts\});$

where the required right hand side variables would be

TYPE	REQUIRED	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	1
integer	fmin	minimum freq (cm-1)	605
integer	fmax	maximum freq (cm-1)	630
matrix	profname	Nx5 matrix (gasID,pressure,pp,temp,amt) pressure, partial pressure in atm, T in K gas amount in kilomolecules/cm2	

and the optional right hand side arguments would be sent in structure *topts*; the default values are as shown (notice that *CKD* is defaulted to -1 (OFF) and the lineshape is augmented to “local”) :

TYPE	OPTIONAL	DESCRIPTION	DEFAULT VALUE
real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0
real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	25.0
integer	nbox	boxcar sum size (odd integer)	5
real	strfar	min line strength for far wing lines	0.0
real	strnear	min line strength for near wing lines	0.0
char	LVG	lineshape : (L)orentz,Voi(G)t,(V)anHuber	’V’
string	HITRAN	path to HITRAN database	/asl/data/hitran/h2k.by.gas
integer	CKD	continuum no = -1 (most gases) yes for water (0,21,23,24) yes for N2, O2 (gases 7,22)	-1
real	selfmult	multiplier for self contiuum 0¡ x¡ 1	1
real	formult	multiplier for foreign continuum 0¡ x¡ 1	1
integer	local	modification to LVG lineshape +1 to use local*chi lineshape defn 0 to use local lineshape defn -1 to use standard LVG lineshape defn	0

4.3 Detailed description of the input parameters

As mentioned above, most of the input parameters are the same as for $run6$ and a description is not repeated here. However, five of the last six parameters are new, and so will be explained below.

CKD is a integer parameter that tells the code which continuum to use. Note that based on whether or not the "local" lineshape was used, the appropriate CKD lookup tables are used. For CKD 0,21,23 the code can compute the continuum whether or not the local lineshape was used; for CKD24, only the local lineshape can be used.

$selfmult$ is a real parameter between 0 and 1, that is used to scale the "self" contribution to the continuum.

$formult$ is a real parameter between 0 and 1, that is used to scale the "foreign" contribution to the continuum.

$usetoth$ is a integer parameter that tells the code whether or not to use the Toth database.

$local$ is a integer parameter that tells the code whether or not to compute the local lineshape (must be set to "0" or "1" to use CKD2.4)

5 run6watercontinuum

run6watercontinuum.m is a specialised code for H2O, that only computes the CKD continuum. This code is to be used in conjunction with run6water.m

5.1 run6watercontinuum.m input parameters

A typical call to run6watercontinuum would involve sending in the following (note that the 'LVG" parameter has been replaced by "divide") :

$[outwave, outarray] = run6water(gasID, fmin, fmax, ffin, fmed, fcor, fstep, xnear, xmed, xfar, nbox, strfar, strnear, divide, CKD, selfmult, formult, usetoth, local, profname);$

where the right hand side variables are the same as those for run6 described above; there are 5 new variables on the right side.

TYPE	VAR	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	1
integer	fmin	minimum freq (cm-1)	705
integer	fmax	maximum freq (cm-1)	730

real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0
real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	150.0
integer	nbox	boxcar sum size (odd integer)	1,5
real	strfar	min line strength for far wing lines	
real	strnear	min line strength for near wing lines	
integer	divide	what is overall computation	-1
integer	CKD	continumm no (-1)	24
		yes water : (0,21,23,24)	
real	selfmult	mult for self part of contiuum	0¡ x¡ 1
	formult	mult for for part of contiuum	0¡ x¡ 1
integer	usetoth	use Toth or HITRAN	+1 to use Toth
			-1 to use HITRAN
integer	local	use local lineshape	+1 to use local*chi defn
			0 to use local defn
			-1 to use run6 defn

The output arguments from the function call are once again the output wavevector, outwave, and the computed line spectra in outarray. The vector outwave (and thus the output array outarray) spans the wavenumber range from $fmin$ to $fmax-(ffin \times nbox)$, at a resolution of $ffin \times nbox$.

5.2 run7watercontinuum.m input parameters

The same parameters as above are used. However, many of them are now defaulted to preset values, and can be reset by using optional input argument $topts$, where $topts$ is a structure.

A typical call to run7 would involve sending in the following :

$[outwave, outarray] = run7watercontinuum(gasID, fmin, fmax, profile, \{topts\})$;

where the required right hand side variables would be

TYPE	REQUIRED	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	1
integer	fmin	minimum freq (cm-1)	605
integer	fmax	maximum freq (cm-1)	630

matrix	profname	Nx5 matrix (gasID,pressure,pp,temp,amt) pressure, partial pressure in atm, T in K gas amount in kilomolecules/cm2

and the optional right hand side arguments would be sent in structure *topts*; the default values are as shown (as is readily appreciated, most of the run7 input arguments are unnecessary and have been REMOVED; also sice the continuum is smooth, we do not really need 5 point averaging and so nbox is defaulted to +1 instead of +5) :

TYPE	OPTIONAL	DESCRIPTION	DEFAULT VALUE
real	ffin	fine point spacing (cm-1)	0.0025
integer	nbox	boxcar sum size (odd integer)	1
integer	CKD	continuum no = -1 (most gases) yes for water (0,21,23,24)	-1
integer	divide	what is overall computation	-1
real	selfmult	mult for self contiuum 0 ¡ x ¡ 1	1
real	formult	mult for foreign continuum 0 ¡ x ¡ 1	1
integer	local	modification to LVG lineshape +1 to use local*chi lineshape defn 0 to use local lineshape defn -1 to use standard LVG lineshape defn	0

5.3 Detailed description of the input parameters

$ffin \times nbox$ determines the output spacing

CKD is a integer parameter that tells the code which continuum to use. Note that based on whether or not the "local" lineshape was used, the appropriate CKD lookup tables are used. For CKD 0,21,23 the code can compute the continuum whether or not the local lineshape was used; for CKD24, only the local lineshape can be used.

selfmult is a real parameter between 0 and 1, that is used to scale the "self" contribution to the continuum.

formult is a real parameter between 0 and 1, that is used to scale the "foreign" contribution to the continuum.

local is a integer parameter that tells the code whether or not to compute the local lineshape (must be set to "0" or "1" to use CKD2.4)

divide is an IMPORTANT parameter that enables the user to ouput the correct total continuum, or along with $selfmult$, $formult$, just parts of it (such as self or foreign or combination). The effects are described as below

DIVIDE	SELFMULT	FORMULT	DIVIDES BY	RESULT
-1	xx	xx	1.0	q v tanh(c2 v/2T) (296/T) × (ps CS + pf CF)
+1	1.0	0.0	q v tanh(c2 v/2T) (296/T) * ps	CS
+1	0.0	1.0	q v tanh(c2 v/2T) (296/T) * (p-ps)	CF
+1	xx	xx	q v tanh(c2 v/2T) (296/T)	(ps CS + pf CF)

6 run6co2

run6co2.m is a specialised code for CO2, so as to to utilise the above physics, namely line mixing and duration of collision effects in the computed lineshapes. If the user simply wants to do a Lorentz or Voigt computation, then it would behoove him/her to use $run6.m$ instead of this special code.

6.1 run6co2.m input parameters

A typical call to run6co2 would involve sending in the following :

$[outwave, outarray] = run6co2(gasID, fmin, fmax, ffin, fmed, fcor, fstep, xnear, xmed, xfar, nbox, strfar, strnear, LVF, IO, birn, profile)$

where the right hand side variables are the same as those for run6 described above; there are three new variables on the right side.

TYPE	VAR	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	2
integer	fmin	minimum freq (cm-1)	705
integer	fmax	maximum freq (cm-1)	730
real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0
real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	150.0
integer	nbox	boxcar sum size (odd integer)	1,5

real	strfar	min line strength for far wing lines	
real	strnear	min line strength for near wing lines	
char	LVF	(L)orentz,(V)anHuber,(F)ullMixing	’F’
char	IO	’0’ for no mixing, ’1’ for mixing	’1’
char	birn	(n)o chi , (c)ousin, (b)irnbaum	’b’

The output arguments from the function call are once again the output wavevector, outwave, and the computed line spectra in outarray. The vector outwave (and thus the output array outarray) spans the wavenumber range from $fmin$ to $fmax - (ffin \times nbox)$, at a resolution of $ffin \times nbox$.

6.2 run7co2.m input parameters

The same parameters as above are used. However, many of them are now defaulted to preset values, and can be reset by using optional input argument $topts$, where $topts$ is a structure.

A typical call to run7co2 would involve sending in the following :

$[outwave, outarray] = run7co2(gasID, fmin, fmax, profile, \{topts\});$

where the required right hand side variables would be

TYPE	REQUIRED	DESCRIPTION	TYPICAL VALUE
integer	gasID	HITRAN gas ID	2
integer	fmin	minimum freq (cm-1)	605
integer	fmax	maximum freq (cm-1)	630
matrix	profname	Nx5 matrix (gasID,pressure,pp,temp,amt) pressure, partial pressure in atm, T in K gas amount in kilomolecules/cm2	

and the optional right hand side arguments would be sent in structure $topts$; the default values are as shown (notice that $xfar$ is defaulted to 250 cm^{-1}, lineshape is augmented to “full linemixing” with ‘birnbaum” turned on; also, do linemixing on all major bands and PQR branches) :

TYPE	OPTIONAL	DESCRIPTION	DEFAULT VALUE
real	ffin	fine point spacing (cm-1)	0.0005
real	fmed	medium point spacing (cm-1)	0.1
real	fcor	coarse point spacing (cm-1)	0.5
real	fstep	wide mesh width size (cm-1)	1.0
real	xnear	near wing distance(cm-1)	1.0

real	xmed	med wing distance(cm-1)	2.0
real	xfar	far wing distance(cm-1)	250.0
integer	nbox	boxcar sum size (odd integer)	5
real	strfar	min line strength for far wing lines	0.0
real	strnear	min line strength for near wing lines	0.0
char	LVG	lineshape : (L)orentz,(V)oigt/anHuber,(F)ull	'F'
char	IO	upper atmosphere line mixing computations	'0'
		'1' for first order line mix, '0' for none	
char	birn	duration of collisions effects	'b'
		'b','B' for Birnbaum	
		'c','C' for Cousin	
		'n','N' for none	
string	HITRAN	path to HITRAN database	/asl/data/hitran/h2k.by.gas
array	band	which bands to use in line mixing (all = -1)	-1
		choose from the below combinations	
	CO2q_sigpi	[618 648 662 667 720 791 1932 2080 2129];	
	CO2q_delpi	[668 740 2093];	
	CO2pr_sigsig	[2350 2351 2352 2353 2354];	
	CO2pr_deltdelt	[2310 2311];	
	CO2pr_pipi	[2320 2321 2322];	
	CO2pr_sigpi	[618 648 662 667 720 791];	
	CO2pr_delpi	[668 740]; (note k/klor for PR_delpi=0.5)	
array	PQRallowed	which branches to use in line mixing (all = -1)	-1
		choose from the below combinations	
		01 = Q delt pi	
		02 = Q sig pi	
		-14 = P sig pi	
		-15 = P delt pi	
		+14 = R sig pi	
		+15 = R delt pi	
		-11 = P sig sig	
		-12 = P delt delt	
		-13 = P pi pi	
		+11 = R sig sig	
		+12 = R delt delt	
		+13 = R pi pi	

6.3 Detailed description of the input parameters

As mentioned above, most of the input parameters are the same as for $run6$ and a description is not repeated here. Howvere, the last three parameters are either new or slightly different than before, and so will be expounded upon below.

LVF is a character parameter that tells the code which lineshape to use for the lines. Of the lineshapes described previously, this code can compute one of the following three - **L**orentz, **V**anVleck-Huber or **F**ull mixing. The VanVleck-Huber is computed with a Voigt lineshape, while the full mixing is essentially a Lorentz computation and so should only be used at higher pressures.

IO is a character parameter that works if LVF is not 'F' or 'f'. It tells the code whether or not to do *no* line mixing (IO='0') or to do first order line mixing (IO='1').

$birn$ is a character parameter that determines the chi function to be used. There are three choices : (n)o, (c)ousin and (b)irnbaum. If $birn$='C','c' then the cousin lineshape cannot be used with mixing turned on; hence the user cannot have LVF='F','f' or IO='1', with cousin chi functions on.

6.4 Outline of the algorithm

The initialisation part of the code is the same as that of $run6$. This code is designed to compute the line mixing in various P,Q,R bands using the latest knowledge. As such, after the code loads in the line parameters from the HITRAN database, it then goes through a list of bands/branches that it has specific code for, and removes *all* the lines that fall within this list, saving the parameters in assorted *.mat files. All the lines that remain are considered background lines.

Figure 6 outlines the above initialisation stages of the algorithm for run6co2.

The computation of the spectrum then proceeds in two stages. The first stage involves using only the background lines, and the spectrum associated with these lines is computed in the exact same fashion as described above for $run6$. The slight difference is that only one of two lineshapes is used : if the user specified LVF = 'L','l', then a Lorentz/vanhuber line shape is used, else a voigt/vanhuber line shape is used.

The second stage is entirely peculiar to run6co2. For the bands that lie within the HITRAN lines read in, the program proceeds as follows. First the code creates a "high resolution" wavevector that spans $f_{high} = (fmin, fmax)$, at the fine (highest) resolution. The code is then ready to loop. The outermost loop is over the bands/branches, while the inner loop is over atmosphere layers. For each layer, the program stores current layer profile data : total and self pressure, gas amount and temperature. Depending on the total pressure of the layer, it then follows one of the following four choices :

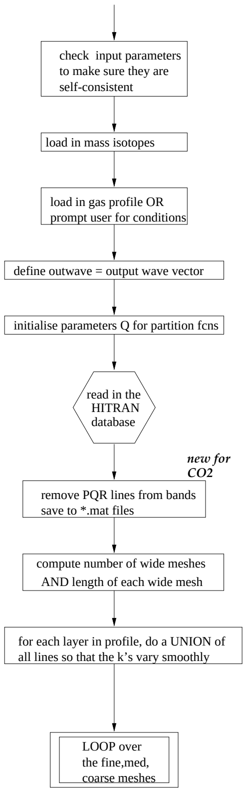

Figure 6:

(a) if LVF = 'F','f' and total pressure is above 0.158 atm, do full line mixing computation, as described below. This is because pressures are high enough that a Lorentz line shape is deemed good enough OR
(b) if LVF = 'F','f' and total pressure is below 0.021 atm, do first order mixing computation, as described below. This is because pressures are low enough that a Doppler line shape is required OR
(c) if LVF = 'F','f' and total pressure is between 0.021 and 0.158 atm, do a weighted average of first order and full mixing computation, as described below. This is because pressures are high enough that first order approximations are not good enough, while pressures are low enough that a Doppler line shape is required OR
(d) if LVF = 'L','V', do the user input lorentz/voigt with linemixing turned on or off (depending on the setting of IO).

Figure 8 outlines the second stage of the algorithm for run6co2.

For each of the bands, and within each of the above 4 choices, the program pro-

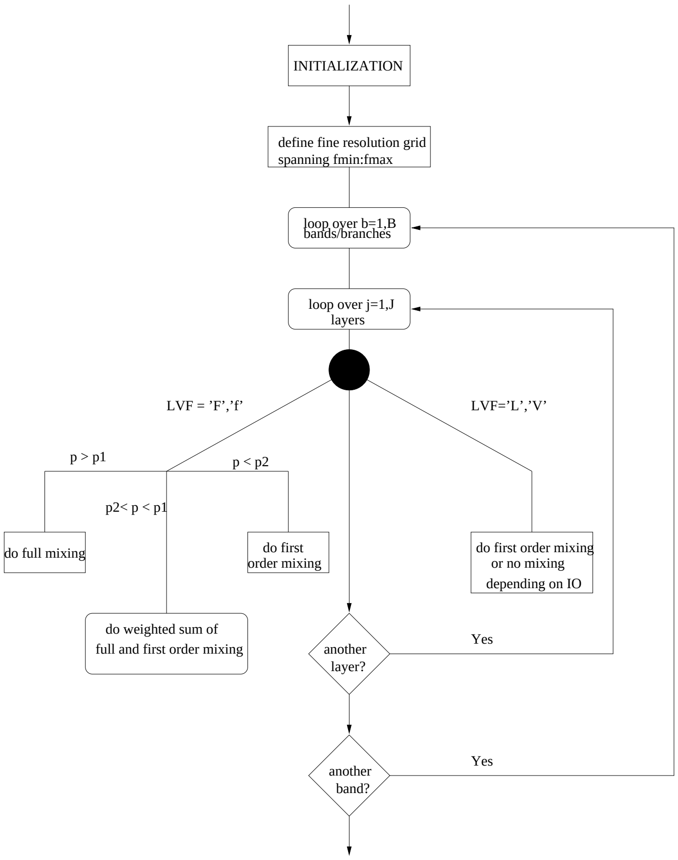

Figure 7:

ceeds as follows. First the HITRAN line parameters for the particular band/branch combination are read in from the hit*.mat saved earlier. The code then computes parameters such as foreign and self broadening, line strength and so on. This is done in $loader.m$

Now the code makes some important speed-affecting decisions based on the current band/branch, and the input high resolution wavevector. Suppose the band lines span the wavevector range (A, B). A preset lookup table has a set of pairs of numbers (a, b) outside of which the effects of linemixing can be simply computed from the lorentz line shape i.e. $k = k_{lorentz} \times MixRatio$. For the current profile conditions, this ratio is computed on the fly using the sum rules derived by Strow et al. If this ratio is computed to be negative, it is reset to a very small number (practically zero). The input wavevector f_{high} then can be divided into three regions

: one that has wavenumbers that are less than a, another that has wavenumbers that are greater than b and a third region whose wavenumbers lie in the interval (a, b).

To compute the ratio, the code has to go ahead and compute the matrices/mixing coefficients associated with the current band/branch combination. The matlab routines are typically $efitter$, $wfunco2er$, $wfun1co2$, $wgradco2$, $trans_pop$

The code now loops over the three wavenumber regions defined above :
(a) for the first two regions, the code simply does a lorentz computation, and multiplies it by the required ratio. Note that in order to speed things up, the computation is done at the $output$ wavevector resolution, so that no boxcar integrations, or splines, have to be done. If a birnbaum or cousin factor is required, it is computed for each line and multiplied in. Thus the overall lineshape in this region is

$$k(\nu) = ratio \times \sum_{i=1}^{i=F} k_{lorentz}(\nu, \nu_i, brd_i) \times \chi(\nu) \tag{5}$$

where the sum is over all the lines in the band, ν_i, brd_i being the line centers and broadening respectively, and χ is either the birnbaum factor or 1.00, and ν is the portion of the $output$ wavevector. The rationale behind the output wavevector being used for the computation is that the first two wavenumber regions are so far away from the lines in the band that the optical depth k is both very small and varying very slowly.
For the third region, the code does an entire mixing computation :

(b) If a full mixing computation is required, the code sets up the required matrix and finds the eigenvalues. The spectra is then built up, at the $fine$ resolution, using the eigenvalues as described earlier on in this document. If a birnbaum factor is required, it is computed as follows. For each line in the band, a lorentz line shape is computed, as is the corresponding birnbaum factor. A vector containing the sum of the products of the lorentz, birnbaum factors, as well as another vector containing the sum of the lorentz lines, is kept. The overall full mixing lineshape is then multiplied by the ratio of thses two sums. Thus the overall lineshape in this region is

$$k(\nu) = k_{full}(\nu, brd, strengths) \times \chi(\nu) \tag{6}$$

where χ is either 1.00 or computed as follows

$$\chi(\nu) = \frac{\sum_{i=1}^{i=F} k_{lorentz}(\nu, \nu_i, brd_i) \times birn(\nu)}{\sum_{i=1}^{i=F} k_{lorentz}(\nu, \nu_i, brd_i)} \tag{7}$$

Having done this the code then boxcar averages the result to the output resolution, and adds on the result to the required part of $outarray$

(c) If a first order mixing computation is required, the code sets up the required first order coeffs. The spectra is then built up, at the $fine$ resolution, using the mix coeffs. If a birnbaum factor is required, it is computed for each line, and then multiplied in. Thus for LVF = 'L', the overall lineshape in this region is

$$k(\nu) = \sum_{i=1}^{i=F} (k_{lorentz} + Y_i \times f(\nu - \nu_i)) \times \chi(\nu) \tag{8}$$

while for LVF = 'V', the overall lineshape in this region is

$$k(\nu) = \sum_{i=1}^{i=F} (k_{voigt} + Y_i \times f(\nu - \nu_i)) \times \chi(\nu) \tag{9}$$

where the sum is over all the lines in the band, ν_i, brd_i being the line centers and broadening respectively, χ is the birnbaum factor or 1.00, and ν is the portion of the $highresolution$ wavevector. Y_i are the first order mixing coefficients, which multiply a term that essentially depends on the distance away from the line center, $\nu - \nu_i$
Having done this the code then boxcar averages the result to the output resolution, and adds on the result to the required part of $outarray$

(d) If a cousin computation is required, the code is also broken up into the three regions above. However, the code does *not* have to do any of the mixing computations. The spectra is then built up, at the $fine$ resolution, for region 3, or at the lower output resolution, for regions 1 and 2. Thus for $birn$ = 'C', LVF = 'L', the overall lineshape is

$$k(\nu) = \sum_{i=1}^{i=F} k_{lorentz} \times \chi(\nu) \tag{10}$$

while for for $birn$ = 'C', LVF = 'L', the overall lineshape is

$$k(\nu) = \sum_{i=1}^{i=F} k_{voigt} \times \chi(\nu) \tag{11}$$

where the sum is over all the lines in the band, χ is the cousin factor, and ν is the relevant portion of the $highresolution$ or $outputresolution$ wavevector. Having done this the code then boxcar either averages the result to the output resolution, and adds on the result to the required part of $outarray$, or directly adds on the result to the required part of $outarray$, as appropriate.

Both the birnbaum and cousin computations are done as lookup tables.

Figure **??** outlines the above stages of the core workings of the PQR band/branches algorithm for run6co2, assuming the user wants to do either a line mixing or a cousin computation.

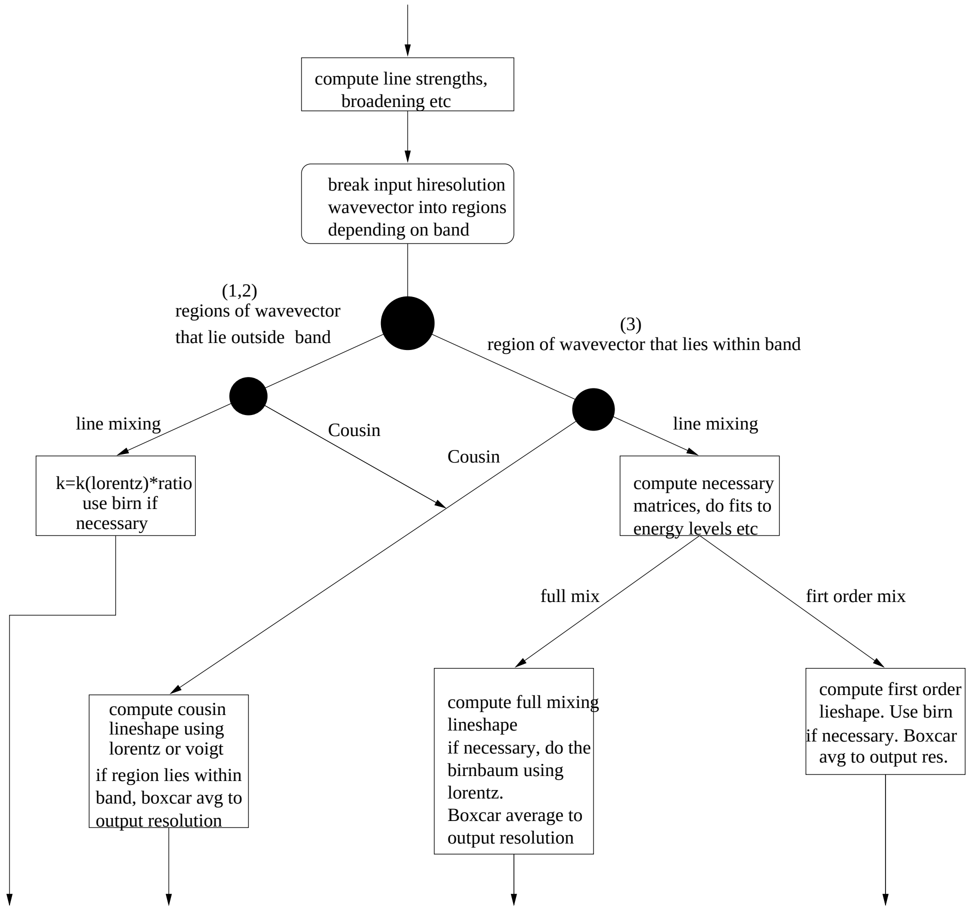

Figure 8:

6.5 Band details

Each of the bands used for the special linemixing computations in $runCO2$ are associated with a unique identifier, which is approximately equal to the band center. This identifier enables the code to pull out the lines from the HITRAN database that have similar quantum numbers. For example, 720 would pull out all HITRAN lines having the lower vibrational quanta index = 2, upper vibrational quanta index = 5, isotope = 1 (main isotope). This is the Q720 sigpi branch.

The bands are divided into subsets associated with the lower and upper level angular momenta : sigmapi, deltapi, sigmasigma, pipi, deltadelta. To see which HITRAN quantum numbers/isotopes are associated with each subset, the user can look at file *makeDAVEhitlin.m*. Only the strongest bands have been included in *run6co2*.

Since some of the bands in the subsets are associated with isotopes, the code used to compure the required mixing matrices for some bands within the same isotope, may be quite different from that for the other bands in the isotope. In particular, computing the $a1, a2, a3$ coefficients for the power-energy scaling gap can vary significantly.

Below, copied from *makeDAVEhitlin*, is a summary of the bands currently used for line mixing:

```
%find all PQR lines for isotope 1 (or 2 or 3) : these are PQR_sigpi
if (band==618)
  v_l=2;v_u=3;
elseif (band==648)
  v_l=1;v_u=2;isotope=2;
elseif (band==662)
  v_l=1;v_u=2;isotope=3;
elseif (band == 667)
  v_l=1;  v_u=2;
elseif (band == 720)
  v_l=2;  v_u=5;
elseif (band==791)
  v_l=3;v_u=8;
elseif (band==2080)
  v_l=1;v_u=8;

%find all PQR lines for isotope 1 : these are PQR_deltpi
elseif (band==668)
  v_l=2;v_u=4;
elseif (band==740)
  v_l=4;v_u=8;
elseif (band==2093)
  v_l=2;v_u=14;

%find all PQR lines for isotope 1 : these are PQR_sigsig
elseif (band==2350)
  v_l=1;v_u=9;
elseif (band==2351)
  v_l=1;v_u=9; isotope=2;
elseif (band==2352)
```

```
  v_l=1;v_u=9; isotope=3;
elseif (band==2353)
  v_l=3;v_u=23;
elseif (band==2354)
  v_l=5;v_u=25;

%find all PQR lines for isotope 1 : these are PQR_pipi
elseif (band==2320)
  v_l=2;v_u=16;
elseif (band==2321)
  v_l=2;v_u=16; isotope=2;
elseif (band==2322)
  v_l=2;v_u=16; isotope=3;

%find all PQR lines for isotope 1 : these are PQR_deltdelt
elseif (band==2310)
  v_l=4;v_u=24;
elseif (band==2311)
  v_l=4;v_u=24; isotope=2;
```

6.6 PQR_sigpi

For the Q branches, the code is quite straightforward, except for the 662 band. Since this is the 3rd isotope (O16 C12 O18) the symmetry is broken and now we can have many more allowed lines.

For the PR branches, only the 720 branch has the line mixing matrices being computed ... the rest of the bands use $k/klor = 0.5$

6.7 PQR_deltpi

For the Q branches, the code is quite straightforward. For the PR branches, all the bands use $k/klor = 0.5$

6.8 PR_sigsig

No Q branch mixing done. The strongest lines in the 4 μm region are from this subset. For the PR branches, full line mixing done. Band 2352 is special because it is a symmetry breaking isotope.

6.9 PR_pipi

No Q branch mixing done. For the PR branches, full line mixing done. Band 2322 is special because it is a symmetry breaking isotope. In addition, band 2321 needed a special $efitter$ initiliazation when fitting for missing energy levels.

6.10 PR_deltdelt

No Q branch mixing done. For the PR branches, full line mixing done.

7 General Spectral Lineshape Theory

This section examines the basic lineshape parameters, including line centers, shifts, widths, and strengths. The standard lineshapes for natural and Doppler broadening are then presented, followed by a review of collisional lineshapes. Most of this information is from Dave Tobin's PhD dissertation [**?**].

7.1 Molecular Absorption and Beer's Law

Molecular absorption occurs when a molecule absorbs light and simultaneously makes a transition to a higher level of internal energy. The absorption of incident photons decreases the outgoing radiation, and a spectral line is produced. The shape or frequency dependence of this absorption is often called the lineshape. Because of various factors, the absorption occurs not only at the resonant frequency of the transition (determined by the difference between the upper and lower energy levels), but over a spread of frequencies. These broadening factors lead to a finite width of the spectral line. While the resonant frequency and the intensity of the absorption are determined primarily by the structure of the molecule, the lineshape is determined by the molecules' environment.

The frequency dependence of the absorption coefficient, $k(\nu)$, determines the shape of a spectral line. Beer's law relates the absorption of radiation through a gaseous medium linearly to the incident radiation (see Figure 9):

$$-dI = k(\nu) I_0 P dl \tag{12}$$

with the absorption coefficient, $k(\nu)$ being the constant of proportionality. $-dI$ is the decrease in radiation flux over a path length of dl through a gas of constant and uniform pressure P. Integrating this equation over a homogeneous path length of L yields the integrated form of Beer's law:

$$T(\nu) = \frac{I_f(\nu)}{I_0(\nu)} = \exp\left(-k(\nu)PL\right). \tag{13}$$

$T(\nu)$ is is the transmission at frequency ν. $I_0(\nu)$ and $I_f(\nu)$ are the initial and final radiation intensities. Thus, the absorption coefficient is related to the observed transmission by

$$k(\nu) = -\frac{1}{PL}\ln\left(T(\nu)\right). \tag{14}$$

It should be noted that deviations from the linear form of Beer's law are only observed at extremely high photon densities. Under atmospheric conditions, however, the linear dependence of the extinction on the amount of absorbing material and incident radiation is valid.

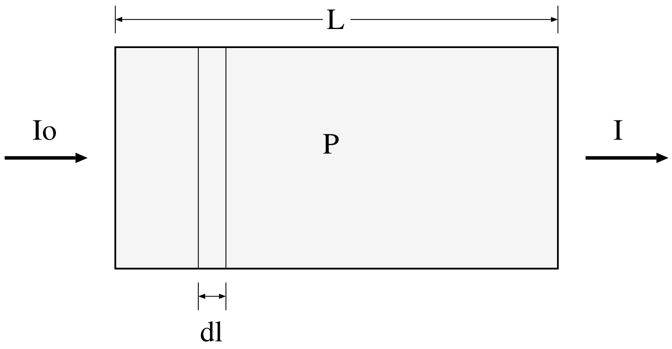

Figure 9: A gas cell of pressure P and length L with incident radiation, I_0, from the left. The amount of absorption at frequency ν is determined by the magnitude of the absorption coefficient, $k(\nu)$, the gas pressure, and the path length according to Beer's law.

7.2 Line Parameters

The lineshape of a single (non-interacting) transition is commonly characterized by several parameters including the line *center* (ν_0), line *strength* (S), and line *width* (γ). These are illustrated in Figure 10.

Several spectral line databases are available which provide a compilation of the line positions, strengths, and widths as well as several other important parameters such as the lower state energy, pressure induced line center shifts, isotopic abundances, rotational and vibrational quantum indexing, width-temperature exponents, and transition probabilities. HITRAN[**?**, **?**] (the high resolution transmission molecular absorption database) is one such database used in this work which is maintained by the Phillips Laboratory Geophysics Directorate . It currently lists the parameters of over 700,000 rotation and vibration-rotation spectral lines for 31 molecules of atmospheric importance from 0–23,000 cm^{-1}. This database represents the most accurate compilation of line parameters. However, due to its size, it is only updated every two to four years and thus recent state-of-the-art measurements and calculations are not always in the database and must be obtained elsewhere.

7.2.1 Line Centers

The line *center*, or position, of a spectral line is determined by the molecular structure, just as the allowable vibrational-rotational energy levels of the molecule are

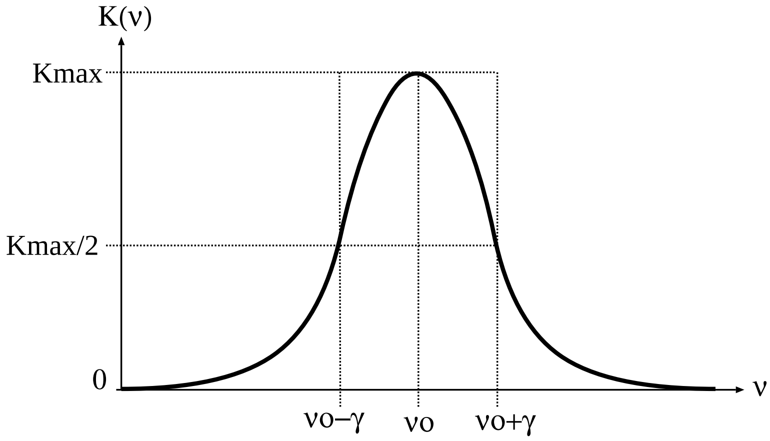

Figure 10: A spectral line depicting the line center, ν_0, and half-width, γ. The line strength, S, is the absorption coefficient, $k(\nu)$, integrated over all wavenumbers ν.

determined by its structure. Planck's relation:

$$\nu_0 = \frac{\Delta E}{h} \tag{15}$$

relates the transition frequency, ν_0 (cm^{-1}), to the change in internal energy, ΔE, where h is Planck's constant. The line centers are thus determined by the structure and allowed energy levels of the molecule and by transition selection rules.

Since the line centers do not depend critically upon the interactions with other molecules, or upon the population of various states, they do not vary significantly with temperature or pressure. One common exception to this are the very small shifts in line center with increasing pressure. Just as molecular collisions can disturb optical transitions leading to increased line widths (discussed later), these disturbances can also lead to an apparent change in the resonant frequency of the molecule's wave-train. Computationally, the shifted position is given by

$$\nu_0(P) = \nu_0 + P \cdot \delta_\nu \tag{16}$$

where P is the total pressure and δ_ν is the pressure induced frequency shift and is determined either theoretically or experimentally. A more theoretical explanation of line shifts which are due to distant collisions was first given by Lenz and more recently by Breene[**?**].

7.2.2 Line Strengths

The line *strength* or line intensity is a direct measure of the ability of a molecule to absorb photons corresponding to a given transition. It depends upon both the properties of the single molecule and the relative number of molecules in the upper and lower states. The strength, S, is defined as

$$S = \int k(\nu_0, \nu) d\nu \tag{17}$$

where the integral is over all ν. Experimentally, S can be determined using Equation 17 if $k(\nu_0, \nu)$ is measured. Alternatively, if the functional form of $k(\nu_0, \nu)$ is known, regression techniques can be used to determine S.

By far the strongest interaction between matter and an incident field of electromagnetic radiation involves the molecule's electric dipole moment. The intensity of a dipole transition is proportional to the square of the matrix element of the dipole moment operator M:

$$R_{i,j} = \int \Psi_i^* M \Psi_j dV \tag{18}$$

where dV is a volume element in configuration space and the integral is over all space. Ψ_i and Ψ_j are the wavefunctions of the lower and upper levels of the transition. The wavefunctions are orthogonal and therefore, if M is unchanged during the transition, $R = 0$. Consequently, for a dipole transition to occur, the electric dipole moment must change between the initial and final energy levels of a transition. Otherwise, the molecule is not linked to the incident radiation and no absorption occurs. Weaker transitions can occur, however, for quadrupole transitions even if there is no change in the dipole moment, although these are not considered in this work.

The line strength is also proportional to the relative populations of the upper and lower transition levels. In thermodynamic equilibrium the probability of a molecule being in a specific energy level is given by

$$\frac{N_i}{N} = g_i \exp(-\frac{hc}{kT} E_i)/Z(T) \tag{19}$$

where h is Planck's constant, c is the speed of light, k is Boltzmann's constant, T is the temperature, N is the total number of molecules, N_i is the number of molecules in energy level E_i, g_i is the statistical weight of the level, and $Z(T)$ is the partition function given by:

$$Z(T) = \sum_i g_i \exp(-\frac{hc}{kT} E_i) \tag{20}$$

In this work, the partition functions are computed using Gamache's [**?**] convenient parameterization:

$$Z(T) = a + bT + cT^2 + dT^3 \tag{21}$$

where a, b, c, and d have been tabulated for most molecules found in the lower atmosphere. Combining Equations 19 and 20, the relative population of the upper and lower energy levels is given by:

$$\frac{N_j - N_i}{N} = \frac{g_j \exp(-\frac{hc}{kT}E_j) - g_i \exp(-\frac{hc}{kT}E_i)}{Z(T)} \tag{22}$$

The line strength is then expressed as

$$S_{i,j} = \sigma_{i,j}\frac{N_j - N_i}{N} \tag{23}$$

where $\sigma_{i,j}$ are integrated absorption cross sections and are given by

$$\sigma_{i,j} = \frac{8\pi^3}{3h}\nu_{i,j}|R_{i,j}|^2. \tag{24}$$

Combining these results yields the full expression for the line strength:

$$S_{i,j} = \frac{8\pi^3}{3h}\nu_{i,j}|R_{i,j}|^2\frac{g_i \exp(-\frac{hc}{kT}E_i)}{Z(T)}[1 - \exp(-\frac{hc}{kT}\nu_{i,j})] \tag{25}$$

Using line strengths determined either theoretically or experimentally at some reference temperature T_{ref}, the strength can be converted to other temperatures using

$$S_i(T) = S_i(T_{ref})\frac{Z(T_{ref})}{Z(T)}\frac{\exp(-hcE_i/kT)}{\exp(-hcE_i/kT_{ref})}\frac{[1 - \exp(-hc\nu_i/kT)]}{[1 - \exp(-hc\nu_i/kT_{ref})]} \tag{26}$$

7.2.3 Line Widths

The line *width*, or halfwidth, is defined as half the frequency interval between ν_0 and the frequency at which $k(\nu)$ has fallen to one half of its maximum value. Values of line widths in the Earth's atmosphere can range from 0.0002 cm^{-1} for conditions where the molecules are isolated to 0.5 cm^{-1} for conditions of extreme pressure broadening. Under pressure broadening conditions, the resulting lineshape near line center is well approximated as a Lorentzian with a nominal line width of

$$\gamma = \frac{2r^2}{m}\frac{P}{R_m T}\left(\frac{3kT}{m}\right)^{\frac{1}{2}}, \tag{27}$$

which has been derived from classical Kinetic theory using the Equipartition theorem and the Ideal gas law. r s the effective radius of the molecule, m is the molecule's mass, P is the total pressure, R_m is the gas constant, and T is the temperature. A typical time between collisions for an atmospheric gas at room temperature and pressure is $\sim 10^{-10}s$, which leads to a Lorentz width of ~0.05 cm^{-1}. If the line width, γ_0, is determined at a given pressure, temperature combination (P_0, T_0), the line width at other conditions is

$$\gamma = \gamma_0 \left(\frac{T_0}{T}\right)^{\frac{1}{2}} \left(\frac{P}{P_0}\right) \tag{28}$$

Thus, the line width increases linearly with pressure and decreases with temperature. Although this kinetic theory does not result in accurate values of γ_0, the pressure dependence is observed in most cases. More commonly, γ_0 is determined either experimentally or calculated with more realistic theories when accurate measurements are not available. Furthermore, the temperature exponent, $\frac{1}{2}$, is generally replaced with a parameter n, which is also determined experimentally. The accuracy of n was investigated by Lui Zheng and Strow [**?**] for both $CO_2 \leftrightarrow CO_2$ and $CO_2 \leftrightarrow N_2$ collisions: $n \simeq 0.69$ for $CO_2 \leftrightarrow CO_2$ collisions and $n \simeq 0.75$ for $CO_2 \leftrightarrow N_2$ collisions. In general, n can vary with transition for the same molecule. For example, accepted values of n for H_2O range from 0.5 to 1. When n is unknown, default values of 0.64 and 0.68 are generally used.

For mixtures of gases, the total line width is the sum of the individual partial widths:

$$\gamma_{TOT} = \sum_i \gamma_{0,i} P_i \tag{29}$$

From quantum Fourier transform theory calculations, the line width for the $f \leftarrow i$ transition is calculated using [**?**]

$$\gamma_i = \left(\frac{n\nu}{2\pi c}\right) \sum_{J_2} \rho(J_2) \sigma_{if,J_2} \tag{30}$$

where n is the perturber density, c is the speed of light, ν is the mean relative thermal velocity ($\nu = \sqrt{8k_B T/\pi\mu}$), μ is the reduced mass of the perturber/absorber system, and $\rho(J_2)$ is the density of the perturber state J_2. σ_{if,J_2} are the absorption cross sections and are dependent upon which type of interactions are dominant. For example, for H_2O-N_2 collisions, the strongest interaction is dipole-quadrupole, yielding

$$\sigma^{DQ}_{if,J_2} = \pi b_0^2 \left(1 + s_{if,J_2}(b_0)\right) \tag{31}$$

where b_0 is an impact parameter related to the minimum distance between absorber and perturber during the interaction and s_{if,J_2} is related to the dipole moment of H_2O, the quadrupole moment of N_2, and the impact parameter. For self-broadened H_2O, the main interaction is dipole-dipole and similar calculations can be performed. This quantum treatment of line widths represents a large improvement over simple kinetic theory calculations. Such calculations, however, are most often scaled to agree with experimental results to obtain the highest degree of accuracy and are included in spectral line databases whenever accurate measurements are not available or possible.

Experimental studies of line widths can become surprisingly complicated for several reasons. One common complication is due to the overlapping and blending of adjacent spectral lines. Others include excessive experimental noise, badly-characterized instrument functions, incorrect "background" absorptions, and lack of characterization of the optical path. Some of these concerns have been reviewed by Gamache *et.al.*[**?**]. Furthermore, for some gases such as water vapor, experimental results from different investigators for the same spectral line lie well outside quoted uncertainties. Larger systematic and analysis errors, not inaccurate experimental spectra, are most likely responsible for these disagreements. The case for N_2-broadened water vapor line widths is investigated in detail in section **??**.

7.3 Lineshape Theories

The frequency dependence of the absorption coefficient is determined by the molecule's physical state and its environment. Broadening factors can be divided into three general classes. They are (1) natural broadening, (2) Doppler broadening, and (3) collision broadening. While natural and Doppler broadening can be described with relatively simple theoretical models, providing an accurate generalized collision broadening theory is a very challenging problem. Each of these are addressed below.

7.3.1 Natural Broadening

The *natural* lineshape is best described by considering a stationary, isolated molecule. If such a molecule is allowed to absorb radiation, undisturbed by any other form, it will eventually make a transition back to a lower level of internal energy. Consequently, the molecule has a limited lifetime at any given energy level. The resulting lineshape is given by

$$k_{nat}(\nu) = \frac{S}{\pi}\left(\frac{\gamma_{nat}}{(\nu-\nu_0)^2+\gamma_{nat}^2}\right) \tag{32}$$

where $\gamma_{nat} = 1/\tau_{nat}$ is the "natural" line width. Due to the relatively long lifetimes of these undisturbed molecules, γ_{nat} is very small, with values on the order of 10^{-5} cm^{-1}. For this reason, natural lineshapes are not observable under atmospheric conditions or with spectrometers of average resolution.

7.3.2 Doppler Broadening

The inhomogeneous *Doppler* lineshape is applicable to conditions encountered in the upper troposphere and stratosphere. In these cases, the temperature is assumed to be high enough to produce molecular motion, but the pressure is low enough so that the molecules experience no collisions; or at least are not subject to *strong* collisions which terminate the dipole moment oscillation. At pressures of about 5 Torr or less, the Doppler lineshape is predominant, with a typical line width of 0.001 cm^{-1} at 296 K. The molecular motion produces an apparent shift in the observed frequencies and such broadening is called Doppler broadening.

The shifted Doppler frequency, ν', for a molecule moving with a speed v_m along the direction of observation, relative to the observer, is given by

$$\nu' = \nu_0 \frac{\sqrt{1-(v_m/c)^2}}{1+v_m/c} \tag{33}$$

where ν_0 is the un-shifted frequency. For $v_m \ll c$, ν' can be approximated with a binomial expansion as

$$\nu' = \nu_0 \left(1 - \frac{v_m}{c}\right) \tag{34}$$

Therefore, for each v_m, there exists a corresponding shifted frequency. Given a Maxwell distribution of velocities within the gas, the density of molecules with velocity v_m is given by

$$dn = N \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} \exp\left(-\frac{m}{2kT} v_m^2\right) dv_m \tag{35}$$

N is the total number of molecules, m is the molecular mass, T is the temperature, and k is Boltzmann's constant. The corresponding absorption coefficient, $k_D(\nu)$, like the Boltzmann distribution, has a Gaussian form:

$$k_D(\nu) = \frac{S}{\gamma_D}\sqrt{\frac{\ln 2}{\pi}}\, e^{-\ln 2\left(\frac{\nu-\nu_0}{\gamma_D}\right)^2} \tag{36}$$

γ_D, the line width of the Doppler lineshape, is given by $\nu_0\sqrt{\frac{2kT\ln 2}{mc^2}}$. Notice how quickly the Doppler lineshape goes to zero far from the line center due to the negative exponential.

7.3.3 Collision Broadening

At pressures greater than $\sim$ 5 Torr, the collisions between molecules must be addressed. Collisions are the most important phenomenon to contribute to broadening at these higher pressures. In 1906 Lorentz showed that line broadening takes place when absorbing molecules or atoms collide. If one assumes that a collision takes place during the time in which radiation is being absorbed, the coherency of the wave train is interrupted. This interruption results in a broadening of the spectral line. Quantum mechanically, pressure broadening is caused by the broadening of the molecules' energy levels by fields produced by the colliding molecules. This is a complex subject, and exact solutions for the absorption coefficient are found only under certain approximations. The exact treatment of this problem requires the knowledge of the time-dependent quantum mechanical wavefunction of an ensemble of colliding molecules. In general, this has not been achieved to date and therefore the problem is often approached by developing empirical or semi-empirical models which simulate the system. In the following sections, a model leading to the standard Lorentz lineshape is presented, followed by descriptions of more complex techniques of dealing with collisional broadening.

7.3.4 Lorentz Lineshape

In the simplest treatment, the collisional lineshape is that of a Lorentzian. At high pressures, collisions occur often and it is unlikely that a molecule is allowed to oscillate undisturbed for its entire natural lifetime. Instead, the molecule is usually perturbed by many collisions. This model makes several assumptions which lead to a simple solution for the absorption coefficient. The molecule's dipole moment is assumed to be oscillating with frequency ν_0. When a collision occurs at time t, the oscillation terminates instantaneously. No natural damping is included because the time between collisions, t, is much less than the natural lifetime, τ_{nat}. In other words, $\exp(-t/\tau_{nat}) = 1$ for all times considered.

It is important to understand the assumptions which have been made for this model. One of the assumptions is called the *impact approximation*, which assumes that the time between collisions is much greater than the duration of a collision, τ_{dur}, and therefore, the behavior of the dipole moment during the collision is negligible. In this case, τ_{dur} is taken to be zero, corresponding to an instantaneous phase shift in the dipole moment. These types of collisions are also sometimes called *adiabatic* in that the system has no time to react to the collisions. The opposite of the impact approximation is called the *quasi-static* approximation, in which the collision durations are essentially assumed to be much larger than the time between collisions. This point is addressed when statistical lineshapes are discussed. Another assumption made here is that of *strong collisions*. A strong collision is taken to be an interaction which terminates the oscillation, leaving no memory regarding its orientation or other properties before the collision. On the other extreme, *weak* collisions are those which have little or no effect in disturbing the molecule. In this case, collisional effects are only felt as a damping effect after a large number of weaker impacts. Collisions are also assumed to involve only two molecules, and such collisions are referred to as *binary collisions*. One final assumption is that the molecules follow classical straight line trajectories between collisions. So in the Lorentz model, which is often called the billiard-ball model, colliding molecules can be thought of as quickly moving hard spheres which do not interact with one another until they actually touch each other.

Fourier analysis of this model wavetrain leads to a spectral distribution of the form

$$| F\{\mu(t)\} |^2 = \frac{\sin^2[2\pi(\nu - \nu_0)t/2]}{[2\pi(\nu - \nu_0)]^2} \tag{37}$$

This expression must be averaged over all possible values of t. From the kinetic theory of gases, the distance traveled between collisions, l, by a molecule of average velocity ν_m has a Poisson distribution:

$$p(l)dl = \frac{dl}{l_m} e^{-l/l_m} \tag{38}$$

where l_m is the mean free path. Using $dt = \frac{dl}{\nu_m}$, the distribution for the time between collisions is

$$p(t)dt = \frac{dt}{\tau_{col}} e^{-t/\tau_{col}} \tag{39}$$

where τ_{col} is the mean time between collisions. Using this distribution, the absorption coefficient, $k_L(\nu)$ is found to be

$$k_L(\nu) = \frac{S}{\pi}\left(\frac{\gamma_L}{(\nu - \nu_0)^2 + \gamma_L^2}\right) \tag{40}$$

where $\gamma_L = 1/\tau_{col}$ is the Lorentz line width. Within this billiard-ball model, τ_{col} is calculated as l_m/ν_m and has an average value of about 1.5×10^{-10} seconds[**?**]. This corresponds to a Lorentz line width of approximately 0.02 cm^{-1}, which is much larger than a typical Doppler width. Thus, whenever collisions are present, they provide the primary form of broadening.

This absorption coefficient is called the *Lorentz* lineshape. It has the same form as the natural lineshape; the only difference being the value of the line widths. It is useful to compare the Doppler and Lorentz lineshapes. The Doppler model assumes a Boltzmann velocity distribution, which goes smoothly to zero at large velocities. Its corresponding spectral distribution, therefore, also decays quickly in the far-wing (far from line center). This is not the case for the Lorentz model, which assumes instantaneous behavior during collisions. The effect of this unphysical temporal behavior is the placement of extremely high frequency components in the lineshape's spectral distribution. Consequently, $k_L(\nu)$ is too large in the far-wing, and the Lorentz model predicts too much absorption in this region.

Despite the apparent shortcomings of the model used for the Lorentzian line shape, it is very accurate for many applications. The Lorentz lineshape is accurate as long as two conditions are satisfied: (1) the spectral region of interest is not too far removed from the line center where the impact approximation results in the prediction of too much absorption, and (2) there exists no significant overlapping of adjacent spectral lines. The latter of these two conditions arises because the Lorentz theory assumes no transfer of intensity from one spectral line to another (often called "line mixing"). Experimental deviations from the Lorentz lineshape within ~2–4 γ_0 of line center of isolated lines have not been confirmed for systems of atmospheric interest.

7.3.5 Van Vleck - Weisskopf Lineshape

When describing the procedures used to calculate the Lorentz and natural lineshapes, the assumption that transitions occur at relatively high frequencies (i.e. infrared) was made. When computing the Lorentz lineshape, the Fourier transform of the dipole moment actually yields two terms – one centered about ν_0 and the other about $-\nu_0$. The lineshape should be written as [**?**]

$$k(\nu) = \frac{S}{\pi}\left(\frac{\gamma}{(\nu-\nu_0)^2+\gamma^2} + \frac{\gamma}{(\nu+\nu_0)^2+\gamma^2}\right). \tag{41}$$

This shape is more often used in the microwave region of the spectrum, where the second term of the sum is not negligible. For molecules active in the infrared region, however, ν_0 is large enough such that $(\nu-\nu_0) \ll (\nu+\nu_0)$ and the resulting lineshape can most often be safely approximated as Lorentzian. An exception is in "window" regions (far from any line centers). Another modification to the Lorentz model involves the behavior of the molecule directly after a collision. In the Lorentz model, we essentially assumed the wave-function experienced *random* phase shifts during collisions and immediately began oscillating again at its resonant frequency. The wave-function, however, does not experience a random reorientation, but should be distributed according to the Boltzmann distribution of the field when the collision occurs[**?**]. Following this approach leads to a slight modification of Equation 41:

$$k(\nu) = \frac{S}{\pi}\left(\frac{\nu}{\nu_0}\right)^2\left(\frac{\gamma}{(\nu-\nu_0)^2+\gamma^2} + \frac{\gamma}{(\nu+\nu_0)^2+\gamma^2}\right) \tag{42}$$

which is commonly called the *Van Vleck-Weisskopf* lineshape [**?**]. Equation 42 is an improvement over Equation 41 in that (41) predicts no absorption in the limit of zero resonant frequency, which is not observed experimentally. Equation 42 is also more acceptable in that it agrees with Debye's[**?**] relaxation theory in the same limit.

7.3.6 Van Vleck - Huber Lineshape

Another similar lineshape which was developed to satisfy the principle of *detailed balance* (discussed below) is[**?**]

$$k(\nu) = \frac{S}{\pi}\left(\frac{\nu}{\nu_0}\right)\frac{\tanh(hc\nu/2kT)}{\tanh(hc\nu_0/2kT)}\left(\frac{\gamma}{(\nu-\nu_0)^2+\gamma^2}+\frac{\gamma}{(\nu+\nu_0)^2+\gamma^2}\right) \tag{43}$$

which is called the *Van Vleck-Huber* lineshape.

7.3.7 Voigt Lineshape

Before going on to explain more elaborate models, the *Voigt lineshape* should be introduced. It does not introduce any new physical insight into broadening phenomenon, but is very useful computationally. The Voigt lineshape is the convolution of the Doppler and Lorentz lineshapes. For this reason, it assumes Doppler characteristics at low pressure and Lorentz characteristics at higher pressures. Therefore, one single expression for the lineshape can be used throughout a wide range of pressures. The Voigt lineshape, $k_V(\nu)$ is given by

$$k_V(\nu) = \frac{k_0 y}{\pi}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{y^2+(x-t)^2}dt \tag{44}$$

with

$$k_0 = \frac{S}{\gamma_D}, \quad y = \frac{\gamma_L}{\gamma_D}\sqrt{\ln 2}, \quad x = \left(\frac{\nu-\nu_0}{\gamma_D}\right)\sqrt{\ln 2} \tag{45}$$

The Voigt lineshape does assume there is no correlation between collision cross sections and the relative speed of the colloids. Again, for atmospheric systems this approximation appears to be quite accurate.

The VanVleck-Huber lineshape can be computed using the Voigt lineshape instead of the Lorentz lineshape.

7.3.8 General Techniques for Calculating Collisional Lineshapes

The lineshape models have been presented informally in order to provide general physical insight. However, for more advanced approaches, it is useful to understand the more formal techniques in which absorption coefficients are calculated. This is needed to help understand the deviations from Lorentz lineshapes. For more on this, one is refered to Dave Tobin's [**?**] thesis for a discussion, as well as for more references. In particular, his dissertation describes line mixing for carbon dioxide, as well as intermolecular forces and potentials used in lineshape calculations for water vapor.

8 Water vapor lineshape

When computing water wapor spectral lineshapes, the effects of the lines far away from the current region can be included in two ways : by directly individually adding on the far wings of each line to the current region, or just using the lines in the current region, plus a lump sum "continuum" contribution. As the far wings lineshapes might not be lorentz, and also to account for the possibility that the near wing lineshapes might also not be lorentz, a local lineshape definition is preferentially used, along with a continuum contribution.

The local lineshape is defined as the lorentz lineshape out to $\pm 25cm^{-1}$ minus the lorentz value at $25cm^{-1}$ away from line center. With this definition, the possibility that the lineshape near linecenter is itself not lorentz, can be modelled by including the effect into the continuum that has to be added on.

Non-Lorentz H_2O lineshapes also have a significant impact *within* the strong pure rotational and vibrational bands. This *in-band* continuum is particularly important for satellite infrared remote sensing of atmospheric H_2O profiles.

For well isolated pressure-broadened water vapor lines in the infrared, the Lorentz lineshape is very accurate near line center. However, if one uses a Lorentz lineshape, this generally overestimate the observed absorption in the far-wing atmospheric window regions and underestimates the absorption within the rotational and vibrational bands. This means that the actual water vapor lineshape is extremely sub-Lorentzian in the far-wing and at least somewhat super-Lorentzian in the intermediate and near-wing. Most experimental studies have focussed on the window regions and so the far-wing lineshape has been studied more than the near-wing (roughly within 5 to 25 cm^{-1} of line center).

Deviations of H_2O spectral lineshapes from Lorentz have been studied extensively for the atmospheric windows at 4 and 10 μm. In general, these deviations are observed to vary slowly with wavenumber and the anomalous absorption has become known as the water vapor continuum.

Several characteristics were found to be common to all window region continuum studies. In general, the continuum absorption [**?**, **?**, for example]: (1) does not change rapidly with wavenumber, (2) decreases rapidly with increasing temperature for pure water vapor, (3) is greater for self-broadening than for foreign broadening, (4) is more significant in regions of weak absorption than in regions of strong absorption, and (5) displays the pressure dependencies associated with gaseous absorption.

It is accepted that the deviations from the impact theory calculations in window regions are due to the non-Lorentz behavior of the far-wings of pure rotational and vibration-rotation water vapor absorption lines.

In the following sections, a review of previous studies of this continuum absorption is presented. These can be separated into two generally different approaches. The first, which is most often adopted in experimental studies, is to express the observed deviations from Lorentzian behavior through the use of *continuum coefficients*. With this method, the cumulative effects of all lines are characterized in a convenient form. The second approach, which provides more information about the shape of individual spectral lines, is used in most theoretical studies. χ-functions are often the end result of this approach.

8.1 A Definition of the Continuum

The definition proposed by Clough is widely used in atmospheric spectroscopy and radiative transfer, particularly in line-by-line codes such as FASCODE[**?**], GENLN2[**?**], LINEPAK[**?**], and LBLRTM. The "local" absorption for a single transition is defined as a Lorentz lineshape out to $\pm 25\text{cm}^{-1}$ from the line center, minus the Lorentz value at 25cm^{-1}. For several lines, the local absorption is expressed as[**?**]

$$k_{local}(\nu) = \nu \tanh(\beta\nu/2)\,\rho_{ref}\frac{T_{ref}}{T}P_{H_2O}L\sum_i \frac{S_i}{\pi}\begin{cases} \frac{\gamma_i}{\Delta\nu^2+\gamma_i^2} - \frac{\gamma_i}{25^2+\gamma_i^2} & \text{if } |\Delta\nu| \le 25\text{cm}^{-1} \\ 0 & \text{if } |\Delta\nu| > 25\text{cm}^{-1} \end{cases} \tag{46}$$

where $T_{ref} = 273.15K$, ρ_{ref} is the absorber number density per atmosphere at T_{ref}, $\beta = hc/kT$, and $\Delta\nu = \nu - \nu_i$. All of the continuum measurements presented in this work are consistent with Equation 46. This is actually a slight modification of Clough's definition[2], which also includes the negative resonance terms $(\frac{\gamma_i}{(\nu+\nu_i)^2+\gamma_i^2})$. In the infrared region (actually for $\nu > 25$ cm^{-1}), the two definitions are equal. *The continuum is then simply defined to be any observed absorption not attributable to the local absorption.* The continuum therefore includes far-wing absorption (beyond 25cm^{-1} from line center), absorption due to any near-wing (within 25cm^{-1}) non-Lorentz behavior, and the Lorentzian value at 25cm^{-1} within $\pm 25\text{cm}^{-1}$ of line center

[2]Equations 6 through 8 of Reference [**?**] do not reflect the local lineshape definition used in FASCODE. They actually lead to a χ dependent local lineshape, which is not used in the line-by-line codes.

(this is often called the “basement” term). This is illustrated in Figure 11 for a single absorption line. The “basement” term is a relatively minor part of the continuum and is introduced to ensure a smooth continuum for computational reasons.

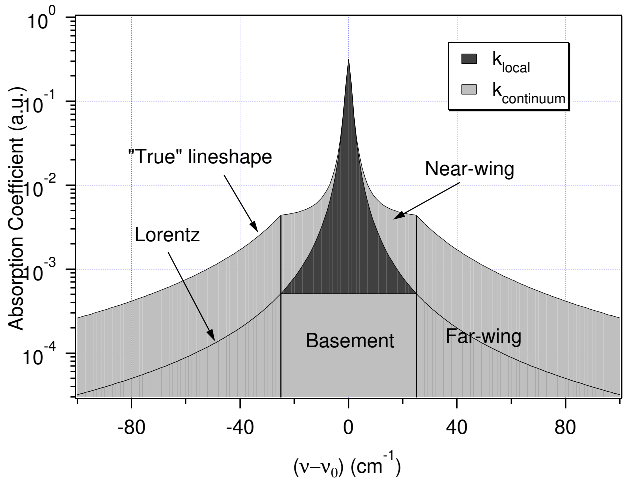

Figure 11: The local lineshape definition used in this work. The far-wing (beyond 25 cm^{-1}), near-wing (within 25 cm^{-1}), and basement components of the continuum are labeled.

With this definition of the local absorption, the continuum is always a positive quantity. The basement and far-wing components are certainly always positive. The near-wing component, which represents the difference between the actual lineshape and Lorentz within 25 cm^{-1}, is also positive because water vapor has a super-Lorentzian lineshape in this region. In fact, with this continuum definition, a non-zero continuum exists even for the Lorentz lineshape. A calculation of the total absorption coefficient and the continuum absorption (total minus local) using the Lorentz lineshape for the 0-4000 cm^{-1} region is shown in Figure 12. All of the high frequency components of the absorption are contained in k_{local} and the continuum is therefore a smoothly varying function, which can be stored in a look-up table for ease of computation.

In the line-wings, the displacement $\Delta\nu$ is much greater than the halfwidth γ and the Lorentz terms can be approximated as $\frac{\gamma_i}{\Delta\nu^2}$. This leads to a quadratic pressure dependence in the absorption coefficient on P_{H_2O} for self-broadened water vapor and a linear dependence on both P_{H_2O} and the broadening pressure, P_f, for foreign broadened water vapor. Since the continuum is mainly due to line wings, the total

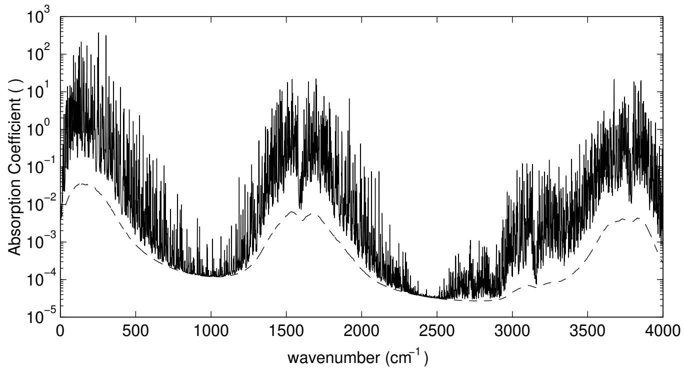

Figure 12: Absorption coefficient calculations using the Lorentz lineshape for 0-4000 cm^{-1} at ~2 cm^{-1} resolution. The total (solid curve) and continuum (dashed curve, as defined by Equation 46) absorption coefficients are shown. The calculations include the effects of all lines between 0 and 5000 cm^{-1}. Conditions are: 1 torr H_2O, 760 torr N_2, 12 m path length, 296K.

continuum absorption coefficient for all lines is formulated as [**?**, **?**]

$$k_{con}(\nu) = \nu \tanh(\beta\nu/2)\,\rho_{ref}\frac{T_{ref}}{T}P_{H_2O}L\frac{296}{T}\left(P_{H_2O}C_s^0(\nu,T) + P_f C_f^0(\nu,T)\right) \quad (47)$$

where C_s^0 and C_f^0 are the self- and foreign-broadened continuum coefficients at 296K and 1 atmosphere. To express experimental and theoretical results, the quantities C_f^0 and C_s^0 are often used.

Because absorption in the windows is very weak, all spectra gathering techniques require very long path lengths. The laboratory studies have focussed primarily on self- and nitrogen broadening at room temperature or above, while most atmospheric measurements have naturally looked at air-broadening at room temperature or below. Most of the results from these measurements are in accord with those of Burch *et.al.*, which are discussed below.

8.2 Laboratory Measurements of Burch *et.al.*

The most notable experimental studies of the water vapor continuum within the fundamental ν_2 band, are those of Burch and co-workers[**?**, **?**, **?**, **?**, **?**]. The work was carried out a number of years ago at relatively low spectral resolution (~0.5 cm^{-1} at 1500 cm^{-1}). Self and nitrogen broadened spectra were collected for a range of temperatures and pressures. Within the band, the wavenumber regions were chosen to be at the center of the so-called microwindows such that the absorption due to lines within $\sim$1cm^{-1} could be ignored and so the low instrument resolution would not distort the spectra. This led to a relatively small number of continuum measurements (13 points between 1400 and 1850 cm^{-1}). A compilation of these measurements is shown in Figure 13 for self-broadened water vapor and in Figure 14 for nitrogen-broadened water vapor. The continuum coefficients shown in these figures have been modified [3] from their published values to be consistent with Equation 46. Van Vleck-Weisskopf continuum coefficients are also shown for comparison. Obvious conclusions from these measurements is that the impact theory predicts too much absorption in the window regions at 4 and 10μm and not enough within the rotation-vibration bands. Due to a lack of other laboratory measurements and because these in-band regions are dominated by near-wing lineshapes, Burch's measurements have led to most of our present understanding of near-wing lineshapes and have formed the benchmark for theoretical comparisons within the ν_2 band. Furthermore, they have been used to produce empirical lineshapes which are used in operational radiative transfer codes and band models which in turn are used in Global Climate Models. The accuracy of these measurements is therefore crucial for several applications.

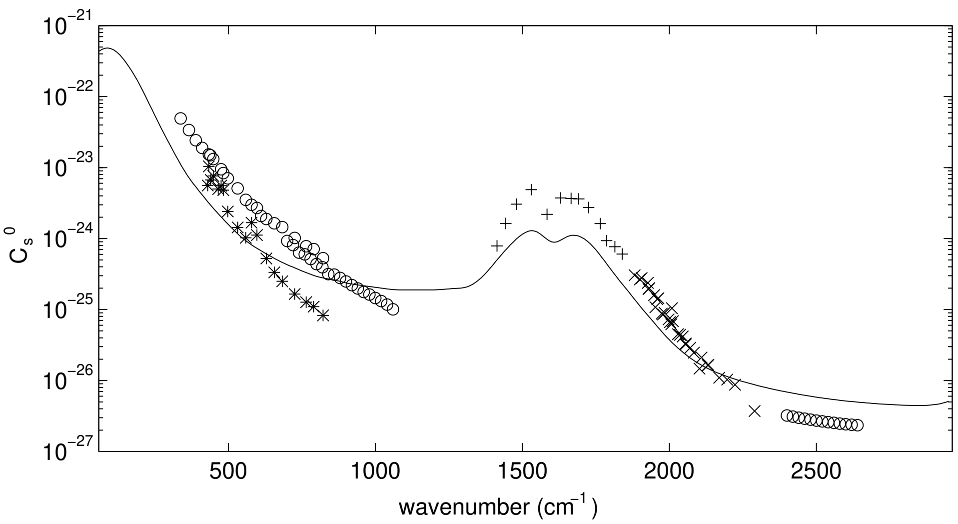

Figure 13: Burch's measurements of C_s^0 and impact theory calculations (solid curve) from 0 to 3000 cm^{-1}. Measurements are shown for temperatures of 296K (circles), 308K (+'s), and 322K (x's), and 430K (asterisks).

[3]Thanks to S. A. Clough for providing these data.

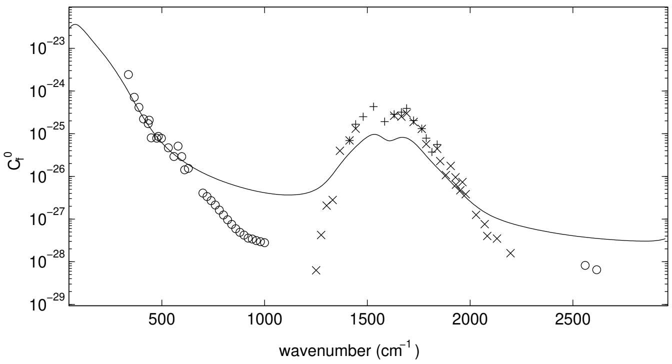

Figure 14: Burch's measurements of C_f^0 and impact theory calculations (solid curve) from 0 to 3000 cm^{-1}. Measurements are shown for temperatures of 296K (circles), 308K (+'s), and 353K (x's).

A good review of progress in the theoretical studies of the water continuum is is Tobin's thesis, and is not reproduced here.

8.3 Clough's CKD Continuum Models

Theoretical approaches have been oriented toward modeling far-wing non-Lorentz behavior in order to predict radiative transfer in the atmospheric windows and consequently are least accurate in the near-wing of the line, which is also more difficult to handle theoretically. The very near-wing lineshape is determined primarily by the long range interaction and the number of collisions per second experienced by the absorber. The far-wing is determined by the durations of the very fast, close collisions which are governed by the strong repulsive potential at small separation distances. The near/intermediate-wing lineshape, however, requires an accurate description of the absorber-perturber interaction for all intermediate time scales, as well as for all separation distances and angular orientations. This becomes a complicated two-body problem involving not only the long range interaction, but also other short-lived electrostatic interactions which arise at shorter separation distances.

Lacking an accurate intermediate-wing lineshape model, Clough *et.al.*[**?**, **?**] have developed empirical models of the continuum to obtain the best overall agreement with measurements in both the window regions and within the bands. These models are widely used in a variety of atmospheric applications. Clough's model expresses the absorption coefficient as

$$k(\nu) = \nu \tanh(\beta\nu/2) \sum_i \frac{S_i}{\pi} \left(\frac{\gamma_i}{(\nu-\nu_i)^2+\gamma^2}\chi(\nu-\nu_i) + \frac{\gamma_i}{(\nu+\nu_i)^2+\gamma^2}\chi(\nu+\nu_i) \right) \tag{48}$$

χ is an empirical function used to include non-Lorentz behavior and is given by

$$\chi = \left(\begin{array}{ll} 1-(1-\chi')\frac{(\nu\pm\nu_i)^2}{25^2} & \text{for } |\nu\pm\nu_i| \leq 25\text{cm}^{-1} \\ \chi' & \text{for } |\nu\pm\nu_i| \geq 25\text{cm}^{-1} \end{array} \right. \tag{49}$$

where χ' is determined empirically. These models, which are known as the CKD (Clough, Kneizys, Davies) continuum models, are shown in Figures **??** and **??**, along with the results of Rosenkranz and Ma and Tipping.

Note that there are separate chi functions for the self and foreign contributions to the continuum; hence

$$\chi = \frac{P(h20)\times\chi_{Pself} + P(for)\times\chi_{for}}{P(h20)+P(for)}$$

The functional form of Equation 49 was chosen because of the known near and far-wing behavior. The χ-functions were held fixed at 1 at line center, adjusted to some value in the intermediate wing, and then forced to decay exponentially in the far-wing. Because Burch's measurements included in-band as well as window region measurements, Clough's χ-functions were forced to assume the correct overall behavior in the near, intermediate, and far-wing. It should be noted that the form of the χ-functions between 0 and 25 cm^{-1} has no direct physical justification, but was adopted simply to interpolate between the line center and fitted region.

The original CKD model (CKDv0) was developed by fitting Equation 48 (with the use of line parameters from the 1986 HITRAN database) to match the laboratory continuum measurements of Burch, which are shown in Figures 13 and 14.

Burch's measurements and continuum coefficients computed with CKDv0 and by Ma and Tipping are shown in Figures 15 and 16. Also shown are two subsequent CKD models: CKDv1 and CKDv2.1. These newer models represent modifications to CKDv0 made to maintain agreement with measurements. For instance, C_s^0 of the CKDv1 model is given by

$$C_s^0 = C_s^0(\text{version0})\left(1 - 0.2333\frac{200^2}{(\nu-1050)^2+200^2}\right) \tag{50}$$

which results in a 23 percent reduction at 1050 cm^{-1}, yielding better agreement with HIS field measurements, yet remaining compatible with Burch's laboratory results. Thus, there are no χ-functions associated with the later CKD versions. Measurements which have led to CKDv2.1 are discussed in the next section.

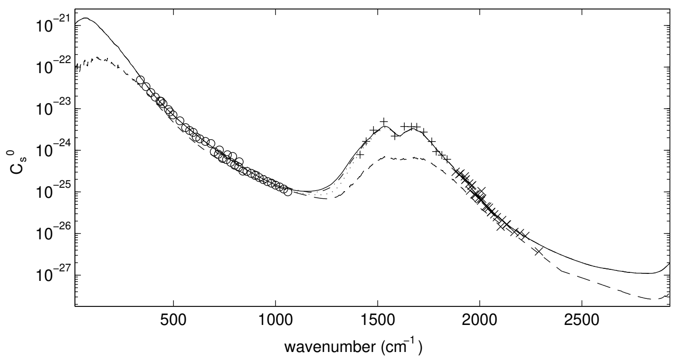

Figure 15: Self-broadened continuum coefficients of Ma and Tipping (dashed curve), Burch (circles, 296K; pluses, 308K; x's, 322K), CKDv0(solid curve), CKDv1(dotted curve), CKDv2.1(dash-dot curve).

The data shown in Figures 15 and 16 for Ma and Tipping's calculations were provided by Q. Ma. For these calculations, a Lorentz lineshape was used within 25 cm^{-1} of line centers due the inaccuracy of their χ-functions in the near wing. Thus, their coefficients do not include any non-Lorentz behavior within 25 cm^{-1}. Their coefficients have also been modified from their original form by this author to include the basement terms to achieve consistency with the continuum definition.

With an accurate description of the χ-function responsible for the in-band water vapor continuum, its relation to the observed absorption can be presented. In window regions, C_s^0 and C_f^0 vary smoothly with wavenumber since there are relatively few strong local absorption lines. Within strong absorption bands, however, the continuum may be influenced more by high frequency near-wing non-Lorentz behavior and these continuum coefficients may not vary as smoothly with wavenumber. Figure 17 shows the individual contributions of the far-wing, near-wing, and basement components to the total continuum (as described in Figure 11) in the ν_2 region based on the CKDv0 χ-function. Although the continuum is dominated by far-wing effects in the wings of the ν_2 band, the near-wing contributions dominate the continuum at the peaks of the P- and R-branches. By studying the continuum in this region, information about the near and intermediate wing lineshape can be determined.

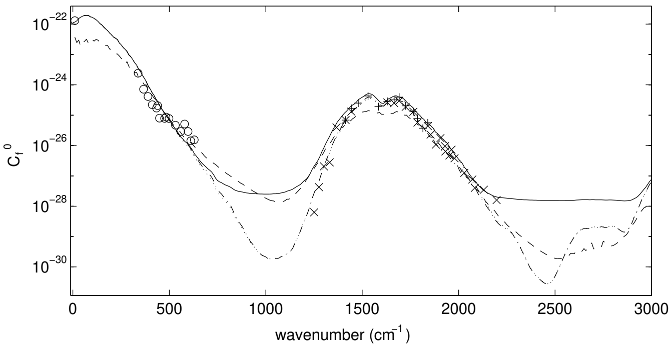

Figure 16: Nitrogen-broadened continuum coefficients of Ma and Tipping (dashed curve), Burch (circles, 296K; pluses, 308K; x's, 322K), CKDv0(solid curve), CKDv1(dotted curve), CKDv2.1(dash-dot curve).

8.4 Recent Field Measurements

Recent field measurements have led to modifications in the CKD models. Thériault *et.al.* [**?**] recently recorded atmospheric transmission spectra over a horizontal path of 5.7 km. These spectra were utilized to test the accuracy of the existing H_2O continuum. Their results suggest that, in the wings of the ν_2 band, the foreign component of the CKDv0 continuum had to be decreased by approximately a factor of 2 to recover good model-measurement agreement. The accuracy of these findings, however, is limited by the nature of the experiment and the difficulties in characterizing the optical path. These measurements are shown in Figure 18.

Atmospheric emission spectra from the University of Wisconsin's High Resolution Interferometer (HIS)[**?**, **?**, **?**] also show that large errors remain in CKDv0. These spectra are very sensitive to the continuum and provide an excellent test of the models when the atmospheric state is well characterized. Differences between observed and calculated high-resolution atmospheric radiances can be as large as 4K in brightness temperature in-between lines inside the ν_2 band when using CKDv0[**?**, **?**, **?**].

Clough recently modified his continuum model to obtain better agreement with several types of atmospheric observations [**?**, **?**, **?**]. In particular, the atmospheric studies suggested that the nitrogen-broadened continuum near ~1200 cm^{-1} was approximately a factor of 2 smaller than predicted by CKDv0. This improved model (CKDv2.1) reduces the maximum errors in HIS calculated radiances to ~2K in-between spectral lines in the 1200 cm^{-1} region [**?**]. However, these errors are

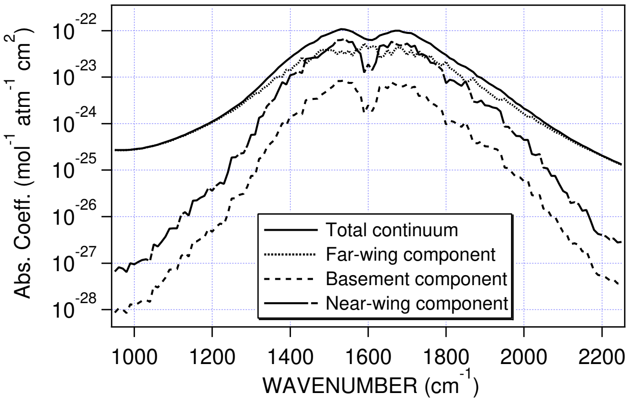

Figure 17: The individual contributions of the far-wing (beyond 25 cm^{-1}), near wing (within 25 cm^{-1}), and "basement" components to the total continuum absorption based on the CKDv0 χ-function.

highly dependent on wavenumber and the newer model also appears to have produced slightly worse results (as compared to CKDv0) further into the band, near the band center. Errors due to CKDv0 and CKDv2.1 for these measurements have been shown in Figure **??**.

9 Carbon dioxide lineshape

Remote sensing of atmospheric temperature and humidity from satellites is dependent on the ability to calculate observed radiances as a function of the atmospheric state. This so–called "forward problem" is at the heart of physically–based retrieval algorithms. Advanced infrared sounders such as the Atmospheric Infrared Sounder (AIRS)[**?**], which is scheduled to fly on the Earth Observing System (EOS) PM platform, will measure radiances in thousands of spectral channels between 3.7μm and 15μm. Physical retrieval algorithms use some subset of these channel radiances to determine global atmospheric temperature and humidity, as well as many other atmospheric and surface parameters.

The 4.3μm spectral region, which is dominated by the ν_3 vibrational band of CO_2, is particularly useful in retrieving atmospheric temperature profiles. This region and several Q-branches in the 15μm region are useful because the absorption varies strongly with wavenumber, thus allowing many levels of the atmosphere to be probed within a very narrow spectral range.

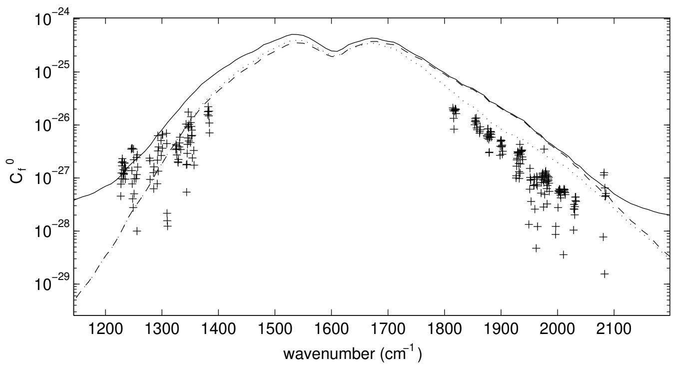

Figure 18: Nitrogen broadened continuum coefficients reported by Theriault. The CKD models are also shown: solid, CKDv0; dashed, CKDv1; dotted, CKDv2.1.

Deviations from Lorentzian lineshape behavior in CO_2 are observed primarily in two cases: (1) when there is significant overlap of adjacent spectral lines and (2) when the spectral region of interest is far from the line center. The Lorentz model fails in these two cases because (1) it neglects the interaction between spectral lines due to inelastic collisions while absorption is occurring and (2) it treats collisions between molecules as if they were instantaneous. These two phenomenon are frequently referred to as line-mixing and the duration-of-collisions effect, respectively. A lineshape which is accurate for the ν_3 band of CO_2 must include both effects.

The lineshape proposed here takes into account the effects of both line-mixing and non-zero collision times. The line-mixing model follows the basic formalism developed by Smith[**?**] and Rosenkranz[**?**]. In this work, instead of doing involved calculations of the the relaxation rates as in [**?**, **?**, **?**], a simple model to approximate the rates is used. Non-zero collision times are modeled with Birnbaum's autocorrelation approach[**?**].

The following section reviews the line-mixing formalism, which is (in principle) not specific to any spectral region or molecule. The application to the ν_3 $\Sigma \leftarrow \Sigma$ band of CO_2 is made in Section **??**.

9.1 Line Mixing

Deviations from the Lorentz lineshape in regions of overlapping spectral lines have been observed in many cases [**?**, **?**, **?**, **?**, **?**] . In particular, large deviations are found in infrared Q-branches, where the spectral lines are very closely spaced. Most attempts

[?, ?, ?] to account for line-mixing have used the impact approximation. This simply means that the theory treats the collisions between molecules as instantaneous and will therefore be accurate only in spectral regions close to the line centers.

Within an ensemble of colliding molecules, a single molecule is not a conservative system. Through interactions with the other molecules, its internal energy can be transformed into other forms of energy throughout the ensemble. For this reason, if the system is defined as a single molecule, the energy of the system does not have to be conserved. Quantum mechanically, this means that the Hamiltonian of the system does not have to be Hermitian and the system can have imaginary energies. The total Hamiltonian of the molecule is given by $H(t) = H_0(t) + H_1(t)$. In this expression, $H_0(t)$ is the Hamiltonian of the molecule without any interaction with its perturbers. $H_0(t)$ therefore has real eigenvalues and its eigenvectors are the stationary states of the molecule. $H_1(t)$ is the Hamiltonian representing the interaction of the molecule with its perturbers. It could, for example, represent the effect of inelastic collisions. $H_1(t)$ has energies which are partially imaginary.

Collision broadening is determined by the behavior of a molecule when making a transition between two known states. This implies that transitions should not be described by two independent functions which model each state separately, but by a wavefunction which represents the evolution from one state to another. This evolution is described by the evolution operator, $T(t)$. According to the Schroedinger equation, $T(t)$ is given by $e^{-\imath(H_0+H_1)\frac{t}{\hbar}}$. Let the initial and final states of a transition be given by Φ_i and Φ_f. Since both of these are stationary states of the molecule, they are solutions to $H_0\Phi = E_0\Phi$ and have totally real energies. For a transition of duration τ_t, which begins at time $t = 0$, the total wavefunction of the system at various times is:

$$\Psi(t = 0) \; = \; T(0)\Psi(0) \; \equiv \; \Phi_i \tag{51}$$

$$\Psi(0 \leq t \leq \tau_t) \; = \; T(t)\Phi_i \; = \Phi_i e^{-\imath(H_0+H_1)\frac{t}{\hbar}} \tag{52}$$

$$\Psi(t = \tau_t) \; = \; T(\tau_t)\Phi_i \; = \; \Phi_i e^{-\imath(H_0+H_1)\frac{\tau_t}{\hbar}} \; \equiv \; \Phi_f \tag{53}$$

Since $H_1(t)$ has imaginary eigenvalues, $\Psi(t)$ is subject to a damping term which drains energy from the system, therefore allowing for absorption. To calculate the absorption coefficient from this model, the autocorrelation function of the electric dipole moment $\mu(t)$ is used. The autocorrelation of $\mu(t)$ is given by $\phi(t) =< \mu(0) \cdot \mu(t) >$, where $<>$ denotes a statistical average over the entire system. Because $\mu(t)$ has contributions from many different elements of the system, $\phi(t)$ will, in general, involve a complicated sum. The absorption coefficient is then calculated by taking the Fourier Transform of $\phi(t)$ [?]. An expression for the correlation of the electric dipole moment of this system over the duration of a transition is written as [?]

$$\phi(t) = <\mu(0)\cdot\mu(t)> = \sum_{if} \rho_i [<\Phi_i(0)|\mathbf{d}|\Phi_f(0)> \cdot <\Phi_f(t)|\mathbf{d}|\Phi_i(t)>] \tag{54}$$

where the sum is taken over all possible states. **d** is the dipole moment operator of the system and ρ is the density matrix, which is related to the relative population of states and is assumed to be constant during a transition. Substituting the evolution operator into this expression, the time dependence of the correlation function is obtained. It is here that the impact approximation must be made. This allows the interaction Hamiltonian, $H_1(t)$, which is generally a fluctuating, time-dependent interaction, to be represented by an effective constant interaction [?]. Since H_1 is taken to be a constant in time, its effect on the correlation function is greatly simplified. If this assumption were not made, the time dependence of H_1 would be required. Instead, only the "time-averaged" interaction is needed and the details of its time dependence are ignored. With this approximation, the lineshape is then calculated by taking the Fourier transform of $\phi(t)$. Baranger [?] calculated the lineshape to be

$$k_{mix}(\nu) = \frac{\rho}{\pi}\mathbf{IM}\left(\sum_i <\Phi_i|[D(\nu - H_0 - H_1)^{-1}]|\Phi_i>\right) \tag{55}$$

where the sum is over all initial states and $D = \sum_f \mathbf{d}|\Phi_f><\Phi_f|\mathbf{d}$. If H_1 is diagonal, with diagonal elements of $\imath\gamma_j$, where γ_j are the Lorentz widths of the spectral lines, the resulting lineshape is the sum of many Lorentzians:

$$k(\nu) = \sum_j \frac{d_j^2 \rho_j\ \gamma_j}{(\nu - \nu_0)^2 + \gamma_j^2} \tag{56}$$

where $d_j^2\rho_j$ is essentially the line strength. This Lorentzian result shows that the model obeys the impact approximation. It also demonstrates the role of the off-diagonal elements of the interaction potential – they are the cause of interaction between spectral transitions. If the off-diagonal elements of H_1 are non-zero, intensity can be transferred from one line to another, with the "amount" of line-mixing determined by the magnitude of the corresponding off-diagonal element of H_1. This is why the phenomenon is often called line-mixing. Thus, a vibration-rotation band cannot be regarded as the simple sum of individual transitions, but must be treated as a complex interacting system. The interaction, which is discussed later, is provided by a bath of inelastic collisions. Inelastic collisions between an absorbing molecule and a perturber can cause the the absorber to gain or lose rotational energy. If this process occurs while a transition is occurring, it is possible that intensity is transferred from one spectral line to another.

The line-mixing absorption coefficient was re-written in doubled-state form better suited for computation of infrared spectra by E. W. Smith [**?**]. If the eigenvalues of H_0 are the line center frequencies (energy and frequency are used interchangeably here) and the interaction, H_1, is linearly proportional to the pressure P (binary collisions), $k_{mix}(\nu)$ can be re-written as[**?**, **?**, **?**]

$$k_{mix}(\nu) = \frac{N}{\pi}\mathbf{IM}\left(\sum_{j,k} d_j \ll j \mid [(\nu - \nu_0) - \imath P\mathbf{W}]^{-1} \mid k \gg d_k \rho_k\right) \tag{57}$$

where N is the molecular density of absorbers, d_j and d_k are the dipole moment matrix elements corresponding to the radiative transitions $|j \gg$ and $|k \gg$, ν is a diagonal matrix with $\ll j|\nu|k \gg = \nu\delta_{jk}$, ρ_k is a density matrix element that represents the population difference between the upper and lower levels of the transition $|k \gg$, and P is the gas pressure. **W** is the relaxation, or interaction, matrix with off-diagonal elements, $W_{j,k}(j \neq k)$, representing the rate, or magnitude, at which collisions transfer intensity from line k to j and its diagonal elements, $W_{j,j}$, representing the line widths of line $|j >>$.

Calculating an absorption coefficient for many transitions over a large spectral range using Equation 57 involves the inversion of a large matrix for each desired frequency. For this reason, $k_{mix}(\nu)$ was re-written as [**?**]

$$k_{mix}(\nu) = \frac{N}{\pi}\mathbf{IM}\left(\mathbf{d} \cdot \mathbf{G}(\nu)^{-1} \cdot \rho \cdot \mathbf{d}\right) \tag{58}$$

where $\mathbf{G} = \nu - \mathbf{H}$ and $\mathbf{H} = \nu_\mathbf{0} + \imath P\mathrm{W}$. **H** is diagonalized with a complex matrix **A** to get the diagonal matrix $\mathbf{L} = \mathbf{A^{-1}} \cdot \mathbf{H} \cdot \mathbf{A}$. **G** is also diagonalized by **A** and $k_{mix}(\nu)$ is written as

$$k_{mix}(\nu) = \frac{N}{\pi}\mathbf{IM}\left(\sum_i \frac{(\mathbf{d} \cdot \mathbf{A})_i(\mathbf{A^{-1}} \cdot \rho \cdot \mathbf{d})_i}{\nu - l_i}\right) \tag{59}$$

where l_i are the diagonal elements of **L**. In this form, the calculation of the absorption coefficient requires only one matrix inversion.

Using time-independent perturbation theory, the second order energy correction for an interaction potential, H_1, is given by $\sum_{m\neq k} \frac{|(\Phi_k, H_1\Phi_m)|^2}{E_k - E_m}$. Following this approach, and assuming that $PW_{jk}/(\nu_j - \nu_k)$ is small for all lines, Rosenkranz found the first-order approximation for $k_{mix}(\nu)$ to be [**?**, **?**]

$$k_{1st}(\nu) = \frac{N}{\pi}\sum_j S_j\left(\frac{P\gamma_j + (\nu - \nu_j)PY_j}{(\nu - \nu_j)^2 + (P\gamma_j)^2}\right) \quad \text{with} \quad Y_j = 2\sum_{k\neq j}\frac{d_k}{d_j}\frac{W_{kj}}{\nu_j - \nu_k} \tag{60}$$

where Y_j are first-order mixing coefficients. For a single transition, this lineshape is the sum of a Lorentzian and an asymmetric term. Far from the line centers, the asymmetric terms become proportional to ν^{-1}. In order for $k_{1st}(\nu)$ to go to zero in these regions, the sum of the coefficients must vanish. That is, detailed balance must be obeyed. In this context, Strow and Reuter[**?**] showed that detailed balance is obeyed if

$$\sum_j S_j Y_j = 0. \tag{61}$$

They also used this result to show that, in the far-wing limit, the ratio of mixing and Lorentz absorption coefficients is a constant[**?**]:[4]

$$\frac{k_{1st}}{k_L} = 1 + \frac{\sum_j S_j Y_j \nu_j}{\sum_j S_j \gamma_j} \tag{62}$$

This is a useful result because it allows the mixing lineshape to be calculated by simply multiplying the Lorentz lineshape by a constant in the far wing. $k_{1st}(\nu)$ is also useful because it provides a compact form for expressing the effects of line mixing. The amount of line mixing for a single line can be represented by the magnitude of the first-order mixing coefficients. Thus, lines which experience a large amount of mixing will have Y's of large magnitude and those which are Lorentzian will have Y's which are zero. The accuracy of the first order mixing absorption coefficient decreases with pressure. The largest errors in first-order mixing occur in spectral regions where lines overlap significantly. As long as this overlap is not too great or as long as the spectral region of interest is in the far-wing, however, the first-order approximation is accurate.

The only point left to be addressed is the determination of the off-diagonal, or mixing, terms of **W**. Direct calculations of these terms are very complex, involving a detailed knowledge of the intermolecular potentials and energy transfer during a collision. More often, empirical scaling laws based on energy changes caused by inelastic collisions are used to model the interactions. Several of these laws have been developed to model collisions for rotational-vibrational transitions, one of which is the exponential power energy-gap, or PEG scaling law. The PEG law models the energetically upward state-to-state inelastic collisional rates as a function of the rotational energy difference, $\Delta E_{j'j}$. An upward rate going from the state j to state j' is modeled as

$$K_{j'j} = a_1 \left(\frac{\mid \Delta E_{j'j} \mid}{B_0} \right)^{-a_2} \exp\left(\frac{-a_3 \mid \Delta E_{j'j} \mid}{kT} \right) \tag{63}$$

[4]The ratio of k_{mix} to Lorentz is also a constant in the far-wing.

where B_0 is the rotational constant and a_1, a_2, and a_3 are adjustable parameters which are discussed below. Other similar scaling laws are also employed. They include the modified exponential energy-gap, or MEG scaling law and the ECSL law [**?**, **?**]. All of these methods have shown to give similar results as the PEG law [**?**]. A more realistic model would not only depend on the energy difference between two levels, but would include rotational and vibrational factors as well.

Detailed balance is obeyed if

$$K_{jj'}(2j'+1)e^{-\frac{E_{j'}}{kT}} = K_{j'j}(2j+1)e^{-\frac{E_j}{kT}} \tag{64}$$

This relation essentially ensures that energy is conserved and gives the downward rates, $K_{jj'}$:

$$K_{jj'} = K_{j'j}\frac{2j+1}{2j'+1}e^{\frac{\Delta E}{kT}} \tag{65}$$

a_1, a_2, and a_3 of Equation 63 are determined by equating the width of a spectral line to the sum of all of the rates which limit the lifetime of that transition via a least-square fit to the known linewidths [**?**]. This is based simply on the fact that any rates which shorten the molecule's lifetime in a specific energy state broadens the spectral line. These rates include all of those which occur in either the lower or upper state of the transition. Since vibrational energies are much greater than rotational energies, only collisions between states within the same vibrational level are considered. Therefore, using line widths extracted from experimental data, a_1, a_2, and a_3 are determined by requiring

$$\gamma_j \equiv \mathbf{W_{jj}} = \sum_{j'\neq j} \text{All } K_{j'j} \text{ which limit the transition lifetime.} \tag{66}$$

This is obviously not the exact expression used in calculations; its details depend on the type of transition which is occurring. This is detailed later when specific symmetries are investigated. Equation 66 is valid only if elastic reorientation collisions and vibrational relaxations do not contribute significantly to the widths, which appears to be valid for CO_2-X systems.

The off-diagonal elements of **W** are then taken to be proportional to the corresponding collisional rates of **K**. These are the mixing terms. For two rotational levels which are energetically close, the collisional rate between them is relatively large and they experience mixing. On the other hand, if two levels are energetically far from each other, the corresponding **K** rates are negligible and no mixing occurs.

In order for line-mixing to occur, collisions which connect two spectral lines must occur in both the upper and lower vibrational states. If, for example, inelastic collisions occurred in only the upper vibrational state, the only effect would be to limit the transition lifetime, and therefore increase the line width. This point is significant for cases in which the collision rates in the upper and lower vibrational states are not equal.

Summarizing the calculational procedures, the relaxation rates, $K_{jj'}$, are first determined by adjusting a_1, a_2, and a_3 so that the sum of all relaxation rates which limit the lifetime of a transition equals the known line width. This is done using an equation similar to Equation 66. The off-diagonal elements of **W** are then taken to be proportional to the corresponding off-diagonal elements of **K**. The details of this step depend on the symmetry of the band and is discussed in Section **??**. The diagonal elements of **W** are equated to the line widths. The absorption coefficients are then calculated using full- (k_{mix}, Equation 59) or first- (k_{1st}, Equation 60) order mixing. This procedure is incorporated into a least-squares fitting algorithm in which an adjustable parameter that controls the magnitude of the off-diagonal mixing terms of **W** is adjusted to obtain optimum agreement with measured spectra.

9.2 Duration-of-Collision Effects

In addition to ignoring the effects of line mixing, the Lorentz model is inaccurate because it is based on the impact approximation. That is, the duration of the interaction between two colliding molecules is assumed to be negligible. Therefore, the Lorentz lineshape is too large at frequencies far from the line centers. A more accurate model treats the molecules as softer charge distributions with collisions therefore having finite durations. Just how long the collisions last is difficult to determine, but depends on how the molecular potentials interact during a collision. Assuming that the collisions have a finite duration eliminates the need for instantaneous derivatives in the temporal wave function and therefore decreases the high frequency components of the spectral distribution. This effect is shown in Figure 19.

This problem is approached by determining the behavior of the dipole moment during the collision. Birnbaum developed a theory using a versatile empirical correlation function [**?**, **?**] with known short and long time behavior and derived expressions for the absorption coefficient as a function of frequency for spectral transitions whose levels are perturbed by collisions. A flexible model was used:

$$\phi(t) = <\mu(0)\cdot\mu(t)> = \exp\left(\frac{\tau_{dur} - (\tau_{dur}^2 + y^2)^{\frac{1}{2}}}{\tau_{col}}\right) \tag{67}$$

with $y = (t^2 - \imath 2\tau_0 t)^{1/2}$ and $\tau_0 = \frac{\hbar}{2kT}$. τ_{dur} represents the mean duration of collisions and τ_{col} is the mean time between collisions. The lineshape function is then determined by Fourier analysis of $\phi(t)$. At long times, Equation 67 reduces to

$\exp\left([\tau_{dur} \pm \imath\tau_0 - |t|]\tau_{col}^{-1}\right)$ and the resulting line shape is Lorentzian with a width of τ_{col}^{-1}. In terms of the Lorentz shape ($k_L(\nu)$), a lineshape representing an average collision duration of τ_2 is given by [?]

$$k_B(\nu) = k_L(\nu)\chi_B(\nu) = k_L(\nu)A_m z K_1(z)\exp\left(\tau_2\gamma_j + \tau_0\Delta\nu\right) \tag{68}$$

with

$$z = \sqrt{(\gamma_j^2 + \Delta\nu^2)(\tau_0^2 + \tau_2^2)} \text{ and } \Delta\nu = \nu - \nu_j \tag{69}$$

where $K_1(z)$ is a modified Bessel function of the second kind, $\tau_0 = \frac{0.72}{T}$ [5], and A_m is a constant representing the effect of line-mixing far from band center. χ_B is independent of vibrational band except for the mixing factor A_m. Although τ_2 is expected to decrease with increasing temperature [?], no explicit temperature dependence is included in this application.

This "corrective" factor, $\chi_B(\nu)$, for the Lorentz line shape removes much of the far-wing absorption due inherently to the impact approximation. Since we are essentially multiplying two line shapes, this can be seen as taking the convolution of the Lorentzian wave train with the correlation function.

9.3 Combined Line-Mixing and Duration-of-Collision Lineshape

Both line-mixing and the duration-of-collision effect have shown to reduce the amount of absorption in the far-wing limit. Since the line-mixing theory is valid only under the impact approximation, the combined effects of line-mixing and duration-of-collision are approximated as if each effect were independent. In order to include the effects of line-mixing over the entire frequency range, $k_L(\nu)A_m$ in Equation (68) is replaced by the first order mixing absorption coefficient:

$$k(\nu) = \sum_i k_{1st}(\nu_i, \nu)\chi_B(\nu_i, \nu). \tag{70}$$

This represents the proposed lineshape model. The full line-mixing lineshape $k_{mix}(\nu)$ can also be implemented by using

$$k(\nu) = \frac{k_{mix}(\nu)}{\sum_i k_{Lor}(\nu_i, \nu)} \sum_i k_{Lor}(\nu_i, \nu)\chi_B(\nu_i, \nu)$$

if the first-order approximation is too inaccurate (which is seldom true for atmospheric applications). To show the sub-Lorentzian nature of this lineshape, ratios to the Lorentz profile are shown in Figure 20. Far from the ν_3 band center, line-mixing

[5] τ_0 and τ_2 have been converted to units of cm by multiplication by 2πc.

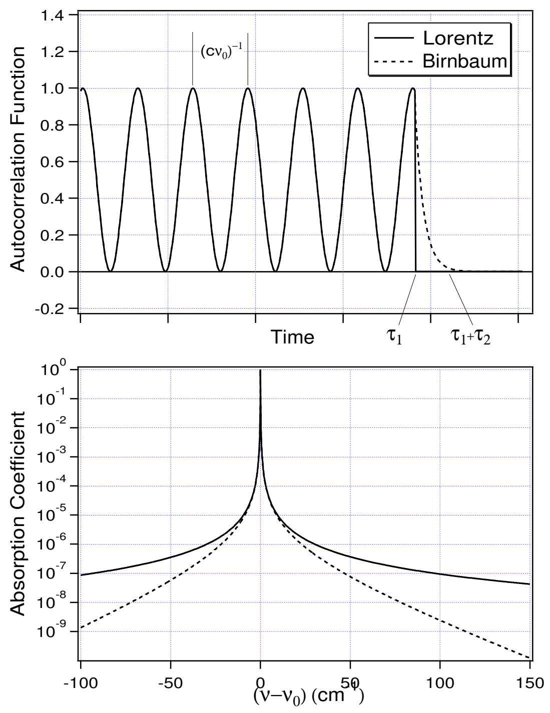

Figure 19: A cartoon (top panel) illustrating the Lorentz (solid curve) and finite duration-of-collision (dashed curve) models of the time development of the dipole moment autocorrelation function when a collision occurs. τ_1 is the time between collisions and τ_2 is the collision duration. The bottom panel shows the resulting lineshapes with the durations-of-collisions model computed with τ_2/τ_1=0.015.

has the overall effect of scaling the Lorentzian profile by a constant, approximately 0.10. Also in the far-wing, χ_B is well approximated by a decaying exponential. One obvious result of these calculations is that line-mixing is responsible for the majority of the observed sub-Lorentz behavior directly beyond the band-head.

Summarizing, Equation 70 represents the proposed lineshape model. It includes both line-mixing and duration-of-collision effects – the two phenomenon responsible for nearly all deviations from Lorentz lineshape behavior in CO_2. Because the variables ζ and τ_2 are built into the model, it is capable of producing a wide range of lineshape behavior. If, for example, ζ and τ_2 are set equal to zero, the resulting lineshapes are Lorentzian, while if $\zeta = 1$ and $\tau_2 \gg 0$, the resulting lineshape is extremely sub-Lorentzian in the far-wing.

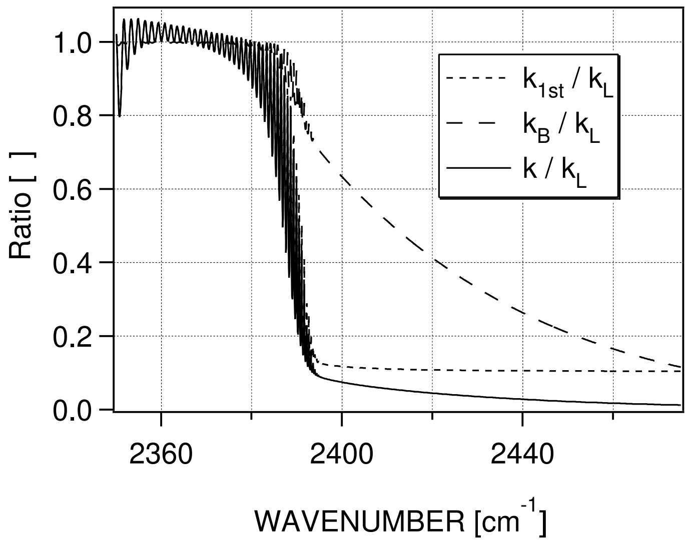

Figure 20: Ratios of the line-mixing (Equation 60), duration-of-collision (Equation 68), and combined (Equation 70) absorption coefficients to Lorentz for the fundamental ν_3 R-branch computed using ζ=1 and τ_2=0.0275. In the far wing, the ratios of line-mixing to Lorentz is a constant while the finite duration-of-collision effect leads to a decaying exponential.

APPENDIX : Computing the lineshapes These appendices duplicate some of the GENLN2 subroutines that have been used in this Matlab code, to compute lineshape parameters.

.1 qfcn.m

This function evaluates the correction to the overall line strength, due to the temperature differing from 296 K. The partition function can be quickly evaluated as a third order polynomial :

$$Q(T) = a + bT + cT^2 + dT^3$$

where the coefficients a, b, c, d depend on the gasID, isotope. The output from the function call is a vector containing the ratio Q(296)/Q(T) for the different isotopes.

```
function [qfcn]=q(A,B,C,D,E,lines,T);
```

```
% initialize coefficients vectors for qtips coefficients
a1 = ones(length(lines.ZISO),1);b1 = ones(length(lines.ZISO),1);
c1 = ones(length(lines.ZISO),1);d1 = ones(length(lines.ZISO),1);

% Assign coefficients according to isotope
no_isotopes = max(lines.ZISO);
for i = 1: no_isotopes
  ind = find(lines.ZISO == i);
  a1(ind) = a1(ind)*A(i);
  b1(ind) = b1(ind)*B(i);
  c1(ind) = c1(ind)*C(i);
  d1(ind) = d1(ind)*D(i);
end

% Evaluate partition functions at desired temperature and 296K
Qt   = a1 + b1*T   + c1*T^2   + d1*T^3;
Q296 = a1 + b1*296.0 + c1*(296.0^2) + d1*(296.0^3);

qfcn=Q296./Qt;
```

.2 broad.m

This function computes the line broadening, as a function of total and self pressures, and layer temperature. Pressures are in atmospheres.

The code starts out differently for CO2, for water and for the rest of the gases, as follows, for individual lines :
(a) water : if self broadening $\leq \epsilon$ then
self broadening = 5*air broadening
else
self broadening = self broadening
(b) others : if self broadening $\leq \epsilon$ then
self broadening = air broadening
else
self broadening = self broadening
The total broadening is then

$$brd = airbroadening * (press - press_self) + selfbroadening * press_self$$

$$brd = (296.0/T)^{pwr} * brd;$$

(c)CO2 : (done more rigorously for individual PQR lines in run6co2)
if self broadening $\leq \epsilon$ then
self broadening = 0.0
else
self broadening = self broadening
The total broadening is then

$$brd_for = airbroadening * (press - press_self) * (296.0/T)^{pwr}$$

$$brd_self = selfbroadening * (press - press_self) * (296.0/T)^{0.685}$$

$$brd = brd_self + brd_for$$

```
function [brd]=broad(press,press_self,press_ref,air,self,pwr,T,iGas)
%compute the broadening by combing air and self broadening
%remember units are in cm-1 per atm at 296 K, so we need the pressures
% press      = current AIRS pressure in atm
% press_self = current self pressure in atm
% press_ref  = current reference pressure in atm
% air        = air broadening cm-1/atm at 296 k
% self       = self broadening cm-1/atm at 296 k
% pwr        = power relationship to scale brd wrt temperature
% T          = temperature
%iGas        = GAS ID

%this eqn is from pg 31 of Genln2 manual

%%%brd=air*(press-press_self)/press_ref + self*press_self/press_ref;
%assume press_ref = 1.0 atm

if iGas ~= 2
  %this is the vectorised code
  dummysmall = (self < eps);
  dummybig   = (self >= eps);
  slfb=self.*dummybig;

  if (sum(dummysmall) > 0)
    if (iGas == 1)
      slfs=(5*air).*dummysmall;
    else
      slfs=air.*dummysmall;
```

```
      end
  else
    slfs=zeros(size(self));
    end
  slf=slfs+slfb;

  brd=air*(press-press_self) + slf*press_self;
  brd=(296.0/T).^(pwr).*brd;

else
  %this is the vectorised code
  dummysmall = (self < eps);
  dummybig   = (self >= eps);
  slfb=self.*dummybig;

  if (sum(dummysmall) > 0)
    slfs=air.*dummysmall;
  else
    slfs=zeros(size(self));
    end
  slf=slfs+slfb;

  brdf=(press-press_self)*air.*(296.0/T).^(pwr);
  brds=(press_self)*slf.*(296.0/T).^(0.685);
  brd=brdf+brds;
  end
```

.3 findstren.m

This subroutine finds the line center strength. It first changes the linestrength that is in the HITLIN database back to the units that are in the HITRAN databse by multiplying by Avogadros number * 1000

$$s00 = s0 * 6.022045e26$$

The line strength is then found from

$$S(T) = US(296)\frac{Q(296)}{Q(T)}s_b s_e$$

where U = gas amount in GENLN2 units of $kiloMoles/cm^2$, S(296) is the line strength read off the tape, Q(296)/Q(T) is the partition function correction, s_b is a Boltzmann factor accounting for lower state population at temperature T (and hence depending on energy of lower state E_{li}), and s_e accounts for detailed balance.

```
function  [strength]=find_stren(qfcn,v0,T,E_li,s0,amt)
%renormalises the strength based on eqn in pg 31 of Genln2 manual
%qfcn = Q(T)/Q(296)
%v0   = central wavenumber
%T    = temperature
%s0   = strength
%E_li = lower state energy
%amt  = gas amt (kilomolecules/ cm2)

s00=s0*6.022045e26;                     %or could do amt=amt*6.022e26
c2=1.4387863;                           %K/ cm-1  from Genln2 manual
sb=exp(-c2*E_li/T)./exp(-c2*E_li/296.0); %boltzman factor (distr at T)
se=(1-exp(-c2*v0/T))./(1-exp(-c2*v0/296.0)); %adjust for detailed balance

strength=amt*(qfcn'.*s00.*sb.*se);
```

This is the individual line strength used in the computations of the lines, whether they are lorentz or voigt or doppler or ......