Absorption coefficients, line collection and frequency grid

M. Kuntz

Abstract: An efficient algorithm of evaluating absorption coefficients line-by-line for the purpose of integrating the radiative transfer equation with emphasis on the Earth's atmosphere is presented. The main problems associated with such an algorithm are concerned with line shape, the treatment of lines outside the spectral range of interest, and the way in which spectral sampling is performed. These are considered together with methods improving speed of computation in general.

1 Introduction

The absorption coefficient is a function of temperature, frequency (wavenumber), pressure, gas volume mixing ratio and constants associated with all contributing line transitions. The contribution to absorption due to a single transition l centered at ν_l can be written as the product of the number density N_s of the molecular species s to which the spectral line belongs, the line intensity S_l, and a line shape factor $f_l(\nu, \nu_l)$. The monochromatic absorption coefficient at frequency ν of a given species thus is

$$\sigma_{s,l}(\nu) = N_s \sum_l S_l \, f_l(\nu, \nu_l), \tag{1}$$

where the summation is performed over all relevant lines. (To simplify notation we do not include all dependent variables in this and other equations). Eq. 1 refers to the absorption coefficient under consideration of local thermodynamic equilibrium (LTE). Corrections to account for non-LTE are performed a posteriori within the radiative transfer modeling (see Part X:'Non-LTE and radiative transfer').
Note that the calculation of absorption coefficients is extremely time consuming, in particular if one tries to calculate the spectral absorption coefficient by summing directly the contribution of each line to absorption. In this case the task soon becomes prohibitive for three reasons: (a) hundreds of thousands of vibrational and rotational transitions contribute to absorption; (b) each transition defines a spectral line that can contribute to absorption over a wide range of frequency; and (c) absorption coefficients must be sampled at a frequency interval that is sufficiently small to resolve the thinnest lines of interest.
Rather than summing directly the contributions to absorption, the method used in KOPRA relies on establishing for each line a set of grid points at which the difference between the analytic value of such a contribution and its interpolated approximation exceeds a limiting small value to be specified by the user; line profiles are evaluated only at these frequencies. The interpolation routine itself is based on three-point Lagrangian interpolation. The line profiles are then added to summations of line profiles calculated with the same spectral resolution. Local absorption coefficients are constructed when the summations are complete.
An important feature of the method is that the frequencies at which each line profile is evaluated are determined dynamically. The spacing between such points increases continuously with distance from line center. Narrow lines are evaluated on a finer scale than broad lines of the same intensity. Lines of stronger intensity are generally evaluated on more points than those of weaker intensity.
The method of approximation is subject to the following three sections which are part of a paper published in JQSRT in 1999 [11]. Following these sections the implementation of the algorithm in Fortran90 will be described together with a depiction of the user interface.

2 Calculation of the line intensity

The line intensity S_l of line l is determined as

$$S_l(T_{kin}, \nu) = S_l(T_0, \nu) \cdot \frac{Q(T_0)}{Q(T_{kin})} \, \frac{e^{-\frac{E_l''}{k_B T_{kin}}}}{e^{-\frac{E_l''}{k_B T_0}}} \, \frac{1 - e^{-\frac{hc\nu}{k_B T_{kin}}}}{1 - e^{-\frac{hc\nu}{k_B T_0}}} \tag{2}$$

where

k_B	Boltzmann constant
E_l''	lower state energy of transition l
$Q(T)$	LTE total internal partition function evaluated at T
$Q(T_0)$	LTE total internal partition function at T_0
$A_l(T_0)$	line intensity at reference temperature T_0
T_{kin}	kinetic temperature
T_0	reference temperature (296 K)
ν	wavenumber, usually approximated as $\nu_{0,l}$
$\nu_{0,l}$	central wavenumber of transition l

The lower state energy E'' is calculated from the lower state energy e'' in units of wavenumbers as given in spectroscopic databases by:

$$E_l'' = c\, e_l''\, h \tag{3}$$

The total internal partition function is determined by use of Gamache's parameterization [7]:

$$Q(T) = a_0 + a_1 T + a_2 T^2 + a_3 T^3 \tag{4}$$

where a_0, a_1, a_2, and a_3 are pretabulated coefficients stored in the file `hitmol.dat` (`$1.3` in the main KOPRA input file).

3 Efficient calculation of the line shape factor

Different line shape factors may be required for different atmospheric conditions. In the lower atmosphere, the shape of spectral lines is dominated by pressure broadening and can be represented most simply by the Lorentz line shape factor:

$$f_{Ll}(\nu, \nu_l) = \frac{1}{\pi} \frac{\alpha_{L,l}}{(\nu - \nu_l)^2 + \alpha_{L,l}^2}, \tag{5}$$

where $\alpha_{L,l}$ is the Lorentz half-width at half maximum of the line. At high altitudes, the shape of spectral lines is governed by Doppler broadening. The corresponding line shape factor is

$$f_{D,s,l}(\nu, \nu_l) = \frac{1}{\alpha_{D,s,l}\sqrt{\pi}} \exp\left[-\frac{(\nu - \nu_l)^2}{\alpha_{D,s,l}^2}\right], \tag{6}$$

where $\alpha_{D,s,l}$ is the Doppler half-width at $1/e$ of the maximum for species s. At intermediate altitudes, both pressure and Doppler broadening are important. They can be modeled using the Voigt line shape factor, a convolution of the Lorentz and Doppler line shape factors:

$$f_{V,s,l}(\nu, \nu_l) = \frac{1}{\alpha_{D,s,l}\sqrt{\pi}} \underbrace{\frac{y}{\pi} \int_{-\infty}^{\infty} \frac{\exp(-t^2)}{y^2 + (x - t)^2} dt}_{K_{V,s,l}(x,y)}, \tag{7}$$

where $x = (\nu - \nu_l)/\alpha_{D,s,l}$ is the distance from line center in units of the Doppler half-widths, $y = \alpha_{L,l}/\alpha_{D,s,l}$ is the ratio of the Lorentzian half-width to the Doppler half-width, and $K_{V,s,l}(x, y)$ is the Voigt profile function. The Voigt line shape factor reduces to the Lorentz and Doppler line shape factors, respectively, in the appropriate limits.

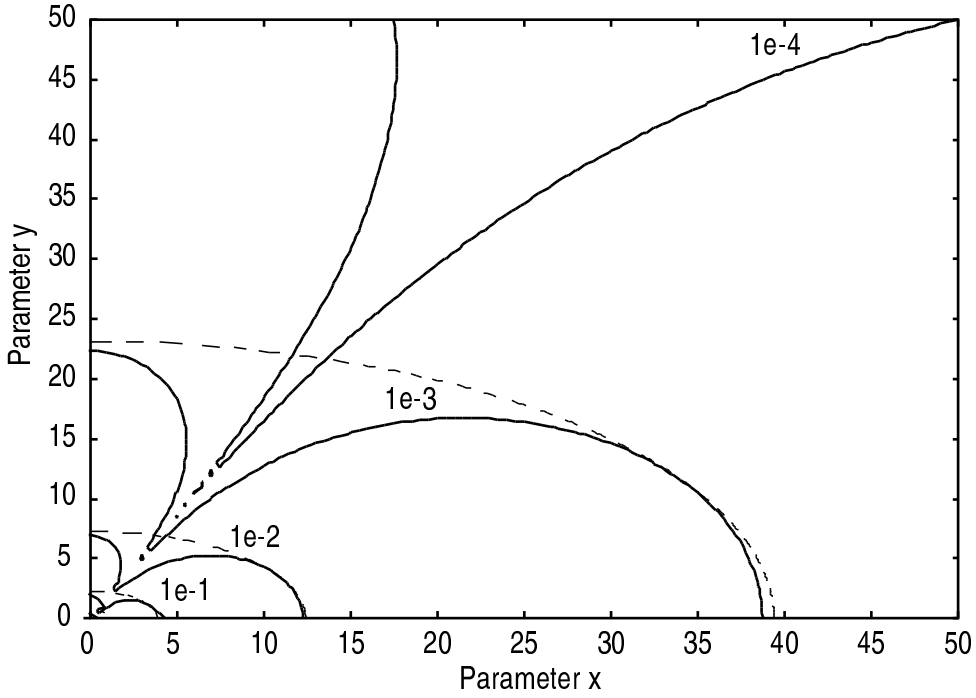

Figure 1: Contours of $|K_{L,s,l}(x,y) - K_{V,s,l}(x,y)|/K_{\mathrm{V}}(x,y)$, i. e. the relative error obtained by approximating the Voigt profile function by the Lorentz profile function. Dashed lines mark the regions outside which the Voigt profile function should be approximated by the Lorentz profile function. Numbers indicate the desired relative accuracy $\mathcal{E}$.

Since the integral in (7) cannot be evaluated analytically, the Voigt profile function $K_{V,s,l}(x,y)$ must be calculated numerically. It is thus desirable to substitute the Voigt line shape factor by the numerically less expensive Lorentz and Doppler line shape factors wherever possible. For deriving an appropriate criterion it seems reasonable to rewrite the Lorentz and Doppler line shape factors (2) and (6) as

$$f_{L,l}(\nu,\nu_l) = \frac{1}{\alpha_{D,s,l}\sqrt{\pi}} \underbrace{\frac{1}{\sqrt{\pi}} \frac{y}{x^2+y^2}}_{K_{L,s,l}(x,y)} \tag{8}$$

and

$$f_{D,s,l}(\nu,\nu_l) = \frac{1}{\alpha_{D,s,l}\sqrt{\pi}} \underbrace{\exp[-x^2]}_{K_{D,s,l}(x,y)}, \tag{9}$$

respectively, where $K_{L,s,l}(x,y)$ and $K_{D,s,l}(x,y)$ are the analogues to the Voigt profile function $K_{V,s,l}(x,y)$ in (7). Since the profile functions $K_{L,s,l}(x,y)$, $K_{D,s,l}(x,y)$, and $K_{V,s,l}(x,y)$ depend only on x and y it is much easier to compare the line shape factors in this form. Figure 1 displays contours of the relative error $|K_{L,s,l} - K_{V,s,l}|/K_{V,s,l}$ on a plane with coordinates x and y. Obviously, the larger the distance from the line center frequency x and the larger the ratio of the Lorentz to the Doppler half-width y is, the smaller the corresponding approximation error becomes. We therefore approximate the Voigt profile function by the Lorentz profile function outside the elliptical regions bounded by the dashed lines in Figure 1. The appropriate criterion may be written as

$$x^2 > 1.52/\mathcal{E} - 2.84y^2,$$

Table 1: Number of floating point operations (Flops) within each region. Exponentiation is considered as 10 floating point operation equivalents.

	Lorentz	Voigt				Doppler
Flops	Region 0	Region 1	Region 2	Region 3	Region 4	Region 5
New	3	7	15	19	79	11
Ref. [10]	7[1]	7	15	19	79	19[3] or 79[4]
Ref. [9]	23[1]	23	41	75	148	75[3] or 148[4]

[1,3,4] Numbers indicated correspond to the former regions 1, 3, and 4, respectively.

where $\mathcal{E}$ represents the desired relative accuracy. Analogous considerations apply for the derivation of an appropriate criterion for the Voigt and the Doppler line shape factor. Figure 2 shows contours of the relative error $|K_{D,s,l} - K_{V,s,l}|/K_{V,s,l}$ in the x-y plane. Note that the relative error is only small within a small region around $y = 0$ that is relatively complicated in shape. However, since the evaluation of the Voigt profile function for small values y is computationally expensive [9][?] (see also region 5 in Table 1), we use the Doppler line shape factor if

$$|x| < 2.15 - 2.53y/\mathcal{E},$$

as is indicated by the dashed lines in Figure 2. Using a more precise parameterization of the contours would be computationally too expensive.
In regions, where neither the Lorentz nor the Doppler line shape factors are appropriate, we use an accelerated implementation of the so-called Humlicek algorithm for the approximation of the Voigt line shape factor [10]. This algorithm involves dividing the x-y plane into four regions, where ratios of rational polynomials are used to achieve relative accuracy better than 1 part in 10^4. To perform the transition from the Voigt line shape factor to the Lorentz and Doppler line shape factors automatically, we add two further regions, region 0 and 5, to the original set of four regions. The borders of these regions are chosen according to the above criteria, and the profile function within these regions is evaluated using (8) and (9), respectively. The extent of saving of floating point operations within these two regions is indicated in Table 1.

4 The optimum set of sampling points

4.1 The method of approximation

For optimum speed, the rapid evaluation of the Voigt profile function must be augmented with a scheme that minimizes its use by reducing the resolution and therefore the number of function evaluations. Numerous strategies have been proposed to achieve this [2][4][5][13][14] . A common approach is to model the absorption coefficient spectra by using multiple frequency grids: fine grids are used to model the line shape factor near the line center frequency, while coarse grids are used for the line wings. However, reducing the resolution requires interpolation to higher resolution grids and introduces numerical errors, see Figure 3. The finer the frequency spacing Δx and the larger the distance from the line center frequency x is, the smaller the interpolation error $\delta K_{s,l}$ becomes.
The principle first proposed by Sparks [13] in order to keep these interpolation errors under control is to introduce an user-specified absolute accuracy coefficient $\delta\sigma$, to which the contribution of any line to the absorption coefficient (1) needs to

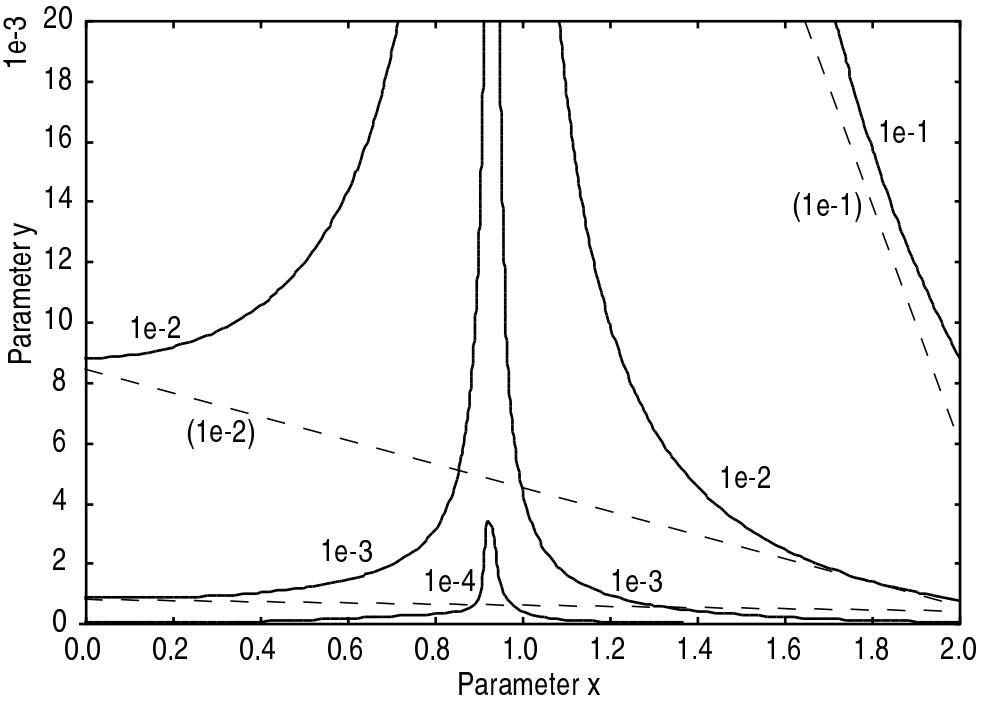

Figure 2: Same as Figure 1 except that the Lorentz profile function has been replaced by the Doppler profile function. Dashed lines mark the regions inside which the Voigt profile function should be approximated by the Doppler profile function.

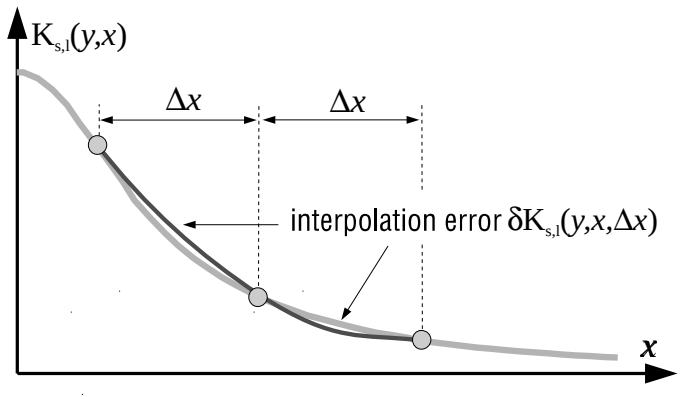

Figure 3: Approximating the profile function by a three-point Lagrangian interpolation.

be calculated to. Inserting Equations (1) and (7)-(9) this condition may be written as

$$\delta\sigma \geq \frac{N_s S_l}{\alpha_{D,s,l}\sqrt{\pi}}\delta K_{s,l},$$

or equivalently

$$\delta K_{s,l} \leq \frac{\alpha_{D,s,l}\sqrt{\pi}}{N_s S_l}\delta\sigma, \tag{10}$$

where $\delta K_{s,l}$ is the interpolation error that can be accepted for a given line. Note that this interpolation error is inversely proportional to the number density N_s of the molecular species s and the line intensity S_l. The stronger a given line is, the smaller the error bounds become within which the interpolation has to be performed to.
To determine the interpolation error associated with a given line and frequency spacing Sparks [14] suggests to approximate the profile function $K_{s,l}(x,y)$ by a simple three-point Lagrangian interpolation [1] :

$$\begin{aligned} K_{s,l}(x_0 + p\Delta x, y) \approx \\ &\frac{1}{2}p(p-1)K_{s,l}(x_0 - \Delta x, y) + \\ &(1-p^2)K_{s,l}(x_0, y) + \\ &\frac{1}{2}p(p+1)K_{s,l}(x_0 + \Delta x, y), \end{aligned}$$

or, setting $x = x_0 + p\Delta x$,

$$\begin{aligned} K_{s,l}(x, \Delta x, y) \approx \\ &\frac{1}{2}p(p-1)K_{s,l}(x - \Delta x - p\Delta x, y) + \\ &(1-p^2)K_{s,l}(x - p\Delta x, y) + \\ &\frac{1}{2}p(p+1)K_{s,l}(x + \Delta x - p\Delta x, y) \\ \equiv\ &\overline{K_{s,l}}(x, \Delta x, y). \end{aligned}$$

The interpolation error is then given by

$$\delta K_{s,l}(x, \Delta x, y) \equiv K_{s,l}(x, \Delta x, y) - \overline{K_{s,l}}(x, \Delta x, y).$$

The problem is to identify the frequencies and frequency spacings at which condition (10) is violated for a given line. Note that the maximum of the Voigt profile function $K_{s,l}(x = 0, y)$ and hence $\delta K_{s,l}(x = 0, \Delta x, y)$ is proportional to $1/y$ for $y \gg 1$ (Lorentz regime), while $\delta K_{s,l}(x = 0, \Delta x, y)$ remains almost constant for $y \ll 1$ (Doppler regime). We therefore multiply (10) on both sides with $y + 1 = (\alpha_{D,s,l} + \alpha_{L,l})/\alpha_{D,s,l}$ (this essentially corresponds to the Voigt half-width in units of the Doppler half-width) and consider the criterion

$$(y+1)\delta K_{s,l} \leq \frac{(\alpha_{D,s,l} + \alpha_{L,l})\sqrt{\pi}}{N_s S_l}\delta\sigma, \tag{11}$$

rather than (10). Condition (11) reduces to expressions (12) and (16) in the work of Sparks [14] in the appropriate limits. Figures 4-6 display contours of $\underset{-1\leq p\leq 1}{\mathrm{Max}}\,(y+1)|\delta K_{s,l}(x, \Delta x, y)|$ on planes with coordinates $\log_2(\Delta x/(y+1))$ and $\log_2|x/\Delta x|$ for different ratios of the Lorentz to the Doppler half-width, y. Note

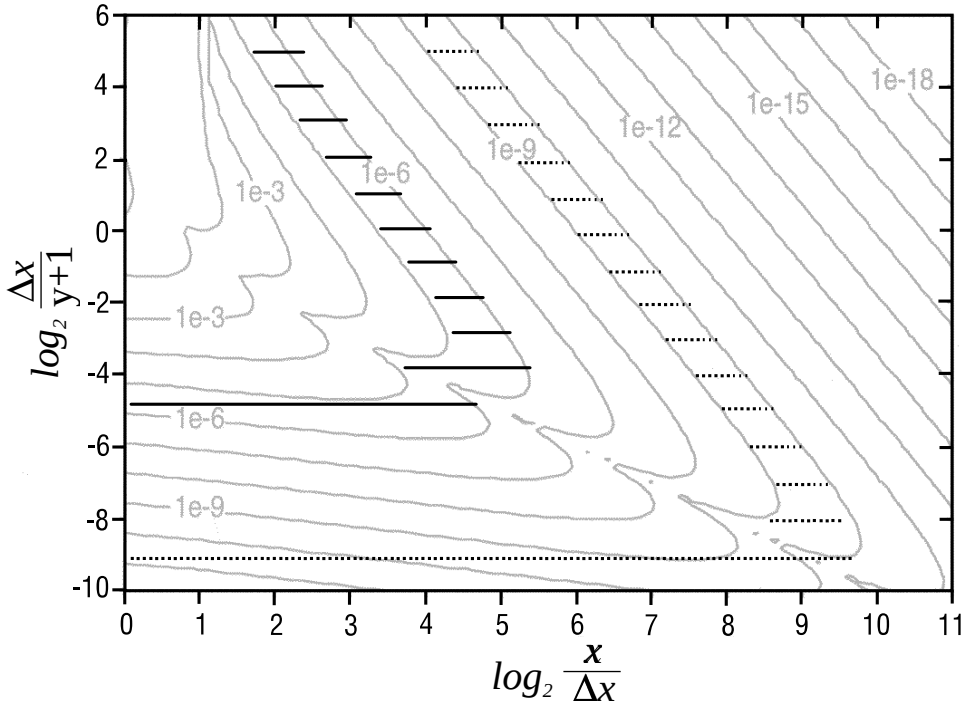

Figure 4: Contours of $\underset{-1\le p\le 1}{\text{Max}}\,(y+1)\delta K_{s,l}(x,\Delta x,y)$ for $y \gg 1$ (Lorentz regime). Also indicated is the optimum set of sampling points for a maximum acceptable interpolation error of $(y+1)\delta K_{s,l} \le 10^{-5}$ and $\le 10^{-9}$, respectively.

that $\Delta x/(y+1) = \Delta\nu/(\alpha_{D,s,l}+\alpha_{L,l})$ is the frequency spacing in units of the Voigt half-width and that $|x/\Delta x| = |\nu-\nu_l|/\Delta\nu$ is the distance from the line center frequency in units of frequency spacing; hence the two axes are not independent of each other. The logarithmic scale on the base two is chosen to account for the second order polynomial used for interpolation. Although this choice of axes seems rather complicated, it is in principle sufficient to note that the vertical axis essentially denotes frequency spacing while the horizontal axis indicates the number of sampling points that should be considered using a given frequency spacing. How this information can be used to determine the optimum set of sampling points for a given line is best illustrated considering examples.

4.2 Cutoff tables

Let us first examine the Lorentz regime, i.e. $y \gg 1$; the corresponding contours of $\underset{-1\le p\le 1}{\text{Max}}\,(y+1)\delta K_{s,l}(x,\Delta x,y)$ are shown in Figure (4). Let us further assume that a given line, indicated by the contour $\delta\sigma\sqrt{\pi}(\alpha_{D,s,l}+\alpha_{L,l})/N_sS_l$, needs to be interpolated with an interpolation error $\delta K_{s,l}$ so that $(y+1)\delta K_{s,l} \le 10^{-5}$. This implies a frequency spacing of $\log_2(\Delta x/(y+1)) = \log_2(\Delta\nu/(\alpha_{D,s,l}+\alpha_{L,l})) \sim -5$ near the line center frequency and the transition to a coarser frequency grid in a distance of $\log_2|x/\Delta x| = \log_2(|\nu-\nu_l|/\Delta\nu) \sim 4.7$, i.e. ~ 26, sampling points (indicated by the solid lines in Figure 4). It would be unwise to start with a finer frequency grid from the beginning since the above value bounds the contour from below and any frequency spacing finer than $\Delta x = (y+1)2^{-5}$ or $\Delta\nu = (\alpha_{D,s,l}+\alpha_{L,l})2^{-5}$ would result in an interpolation error smaller than the one that can be accepted. On the coarser frequency grid one should evaluate $2^{5.5}-2^{3.8}$, i.e. ~ 32,

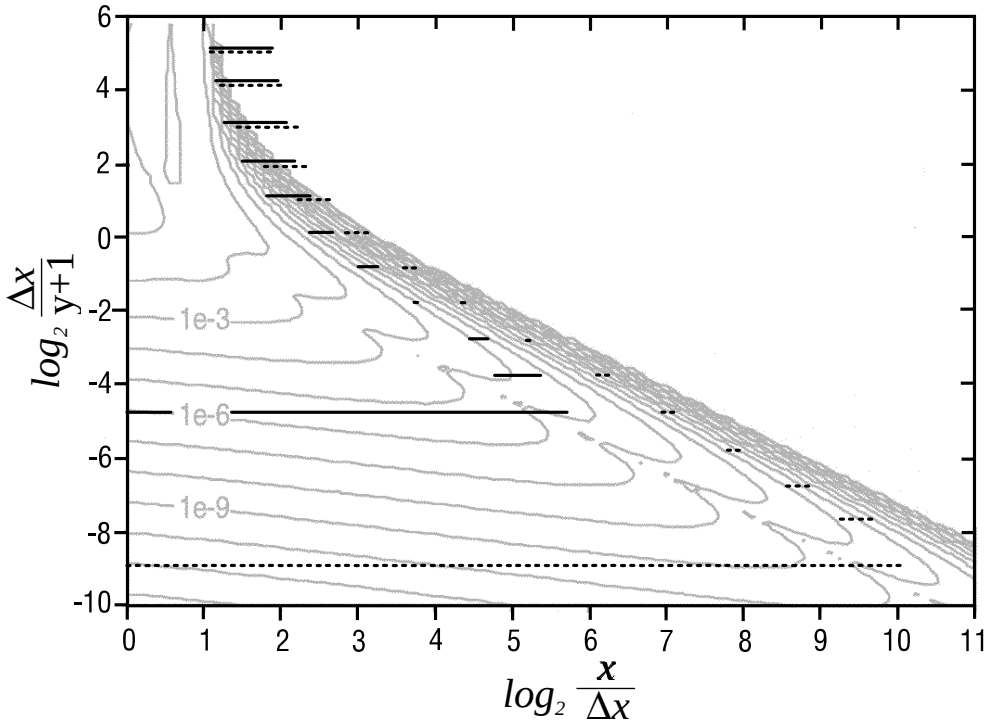

Figure 5: Same as Figure 4 except that the Lorentz regime has been replaced by the Doppler regime, i. e. $y \ll 1$.

sampling points before again proceeding to a coarser frequency grid and so on. The point is to proceed to a coarser frequency grid only if the interpolation error on the coarser grid falls below the maximum interpolation error that can be accepted for a given line. Note that due to the second order polynomial used for interpolation, the frequency spacing always increases by a factor of two.

Let us consider a stronger line that requires an interpolation error $\leq 10^{-9}$. In this case one would start with a much finer frequency grid from the beginning (corresponding to a frequency spacing of $\log_2(\Delta x/(y+1)) \sim -9$ near the line center frequency) and evaluate $\log_2 |x/\Delta x| \sim 10$, i.e. ~ 1000, sampling points before proceeding to a coarser frequency grid (indicated by the dotted lines in Figure 4).

The same procedure can in principle be applied to the Doppler regime, i.e. $y \ll 1$ (see Figure 5), so that tables of cutoffs may be constructed for both limits. These tables indicate over a wide range of possible interpolation errors and frequency spacings in which distance from the line center frequency (or equivalently after the consideration of how many sampling points using a given frequency spacing) one should proceed to a coarser frequency grid. For a given spectral line one can then evaluate the maximum interpolation error $\delta\sigma\sqrt{\pi}(\alpha_{D,s,l} + \alpha_{L,l})/N_s S_l$ and use these tables to define the appropriate Lorentz and Doppler cutoffs for any frequency spacing $\Delta x/(y+1) = \Delta\nu/(\alpha_{D,s,l} + \alpha_{L,l})$ desired. Note that either $\alpha_{D,s,l}$ or $\alpha_{L,l}$ are negligible in the appropriate limits. We will henceforth refer to these tables as the Lorentz and the Doppler cutoffs.

As we mentioned previously, in the real atmosphere both pressure broadening and Doppler broadening are important and can be modeled using the Voigt profile function. Since the latter, however, depends on both x and y, it is in principle necessary to establish a complete set of cutoff tables: one for each ratio of the Lorentz to the Doppler halfwidth y—an enormous waste of computer memory. Figure 6 shows a

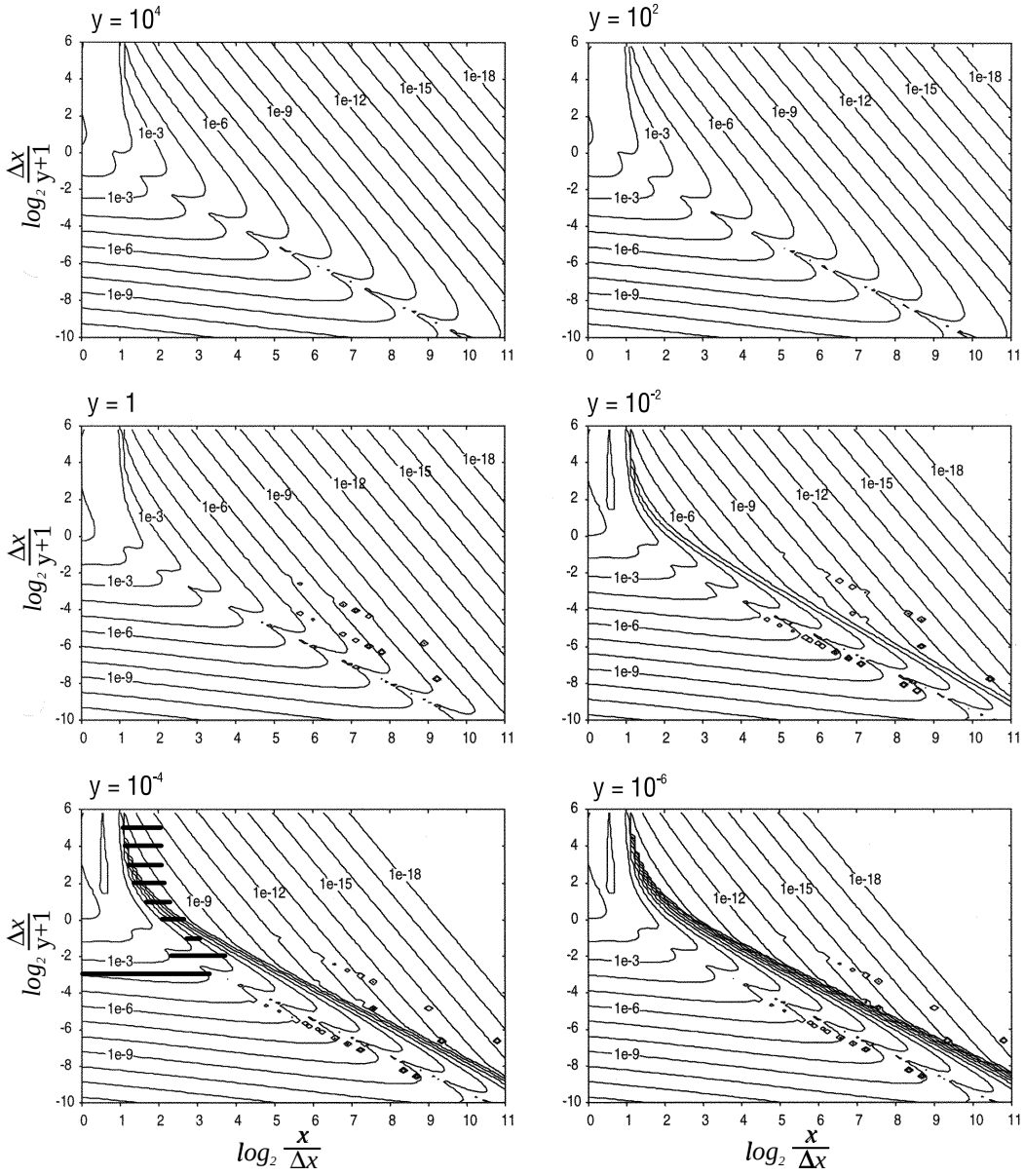

Figure 6: Contours of $\underset{-1\le p\le 1}{\text{Max}}\,(y+1)\delta K_{s,l}(x,\Delta x,y)$ for different ratios of the Lorentz to the Doppler half-width, y. Also shown is the optimum set of sampling points for a maximum permissible interpolation error of $(y+1)\delta K_{s,l} \le 10^{-3}$ for $y = 10^{-4}$.

subset of six contourplots (cutoff tables) for a selected sample of ratios y. Note that the contourlines change from a more Lorentz-like course for large values of y to a more Doppler-like course as y successively decreases (compare Figures 4 and 5).
To overcome this waste of computer memory Sparks [14] suggests to neglect the contribution of a spectral line to absorption on a given frequency grid if the Lorentz cutoff is equal to zero; otherwise, he takes advantage of the fact that the Voigt profile function reduces to the Lorentz and the Doppler profile functions in the appropriate limits and uses the greater of the Lorentz and the Doppler cutoffs for the approximation of the Voigt cutoff. The corresponding contourlines are shown in Figure 7. This approximation works quite well in the Lorentz regime, i.e. for $y > 1$, but fails for $y < 1$ (compare the contours shown in Figures 6 and 7).
Consider for example an maximum interpolation error of 10^{-3} and a ratio of the Lorentz to the Doppler half-width of 10^{-4} (Doppler regime). According to Figure 6 this implies a frequency spacing of $\log_2(\Delta x/(y+1)) \sim -3$ near the line center frequency and the calculation of $\log_2 |x/\Delta x| \sim 2^{3.3}$, i.e. ~ 10, sampling points before proceeding to a coarser frequency grid. The approximation of Sparks [14], however, results in a much finer frequency grid from the beginning (corresponding to a frequency spacing of $\log_2(\Delta x/(y+1)) \sim -7$, i.e. by a factor 16 finer) and requires the evaluation of $\log_2 |x/\Delta x| \sim 2^{6.7}$, i.e. ~ 104, sampling points before proceeding to a coarser frequency grid. Altogether this approximation results in the evaluation of ~ 390 sampling points compared to ~ 30 really necessary according to Figure 6. We mentioned already that the evaluation of the Voigt profile function for small values y is computationally extremely expensive so that any reduction of the number function evaluations in the Doppler regime results in a noticeable gain of computational speed). This incongruity becomes even worse if one prefers to use the simpler but less restrictive criterion (20) in the work of Sparks [14] instead of refering to the appropriate cutoff table. It should be mentioned that this approximation is not a source of computational error—quite the reverse is true. Absorption coefficients in the Doppler regime are calculated much more accurately than necessary. The aim, however, is to consider each spectral line with the same absolute accuracy $\delta\sigma$.
Instead we take advantage of the fact that the interpolation error of the Voigt profile function $\delta K_{V,s,l}(x, y, \Delta x)$ can be well approximated by the greater of the interpolation errors of the Lorentz and the Doppler profile function weighted by the squares of the ratio of the Lorentz to the Voigt half-width $y/(y+1) = \alpha_L/(\alpha_L+\alpha_D)$ and the ratio of the Doppler to the Voigt half-width $1/(y+1) = \alpha_D/(\alpha_L+\alpha_D)$, i.e.

$$\begin{aligned}\delta K_{V,s,l}(x, y, \Delta x) \approx \\ \max\big(&y^2/(y+1)^2 \delta K_{L,s,l}(x_L, y, \Delta x_L), \\ &1/(y+1)^2 \delta K_{D,s,l}(x_D, y, \Delta x_D)\big).\end{aligned} \tag{12}$$

Note that the auxiliary variables Δx_L and Δx_D for a given frequency spacing can be calculated from $\Delta x_L = \Delta x\, y/(y+1) = \Delta\nu\, \alpha_L/\alpha_D(\alpha_L+\alpha_D)$ and $\Delta x_D = \Delta x\, y/(y+1) = \Delta\nu/(\alpha_L+\alpha D)$, respectively. Putting (12) into (11) results in the criteria

$$(y+1)\delta K_{L,s,l}(x_L, y, \Delta x_L) \le \frac{(\alpha_{D,s,l}+\alpha_{L,l})^2\sqrt{\pi}}{\alpha_L N_s S_l}\delta\sigma \tag{13}$$

and

$$(y+1)\delta K_{D,s,l}(x_D, y, \Delta x_D) \le \frac{(\alpha_{D,s,l}+\alpha_{L,l})^2\sqrt{\pi}}{\alpha_D N_s S_l}\delta\sigma \tag{14}$$

for the determination of the auxiliary cutoffs x_L and x_D from the precalculated Lorentz and Doppler cutoff-tables, respectively. The greater of x_L and x_D is finally used for the approximation of the Voigt cutoff.

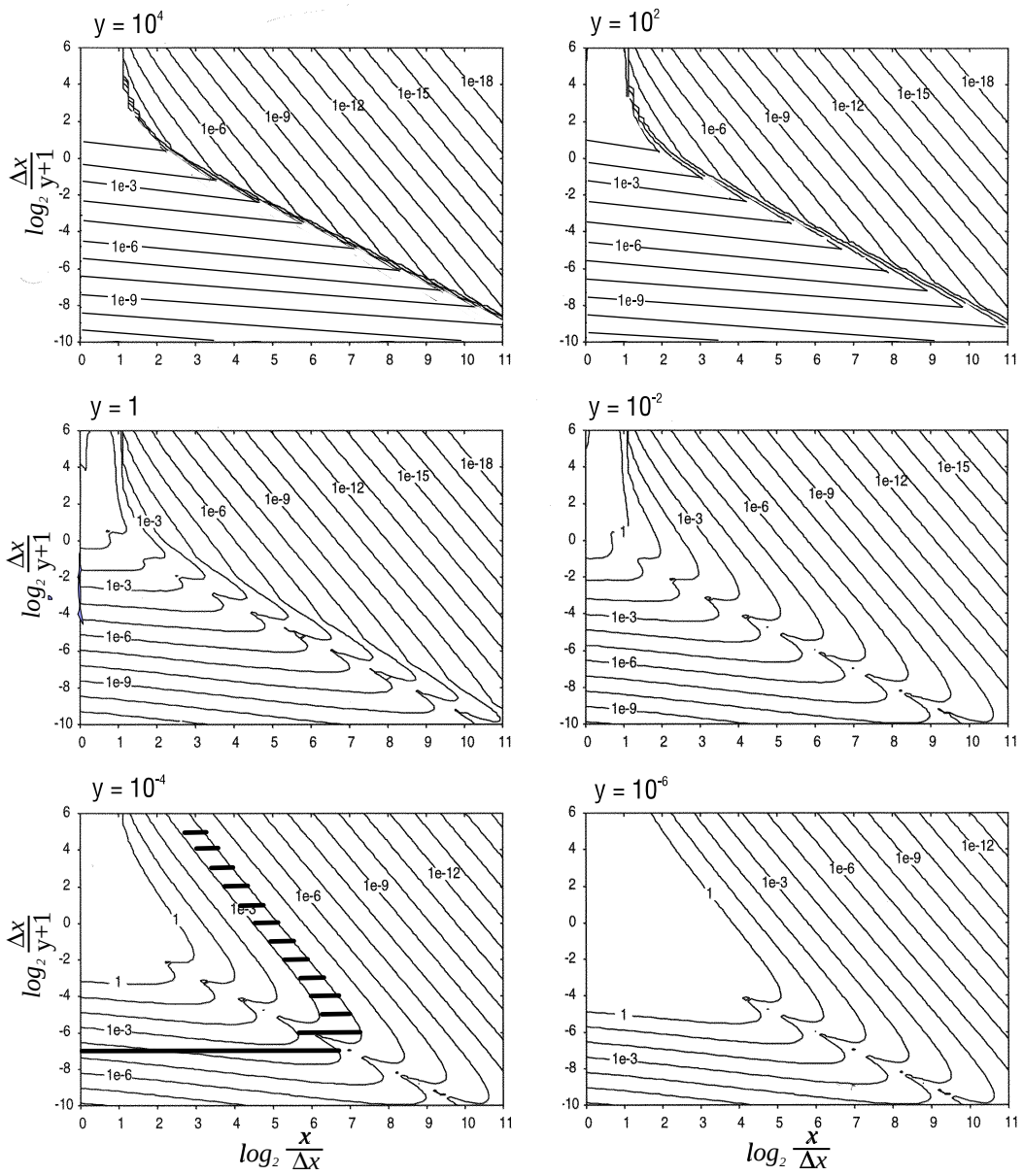

Figure 7: Sparks' [14] approximation for the determination of the Voigt cutoffs. Also shown is the resulting set of sampling points for $(y+1)\delta K_{s,l} \leq 10^{-3}$ and $y = 10^{-4}$.

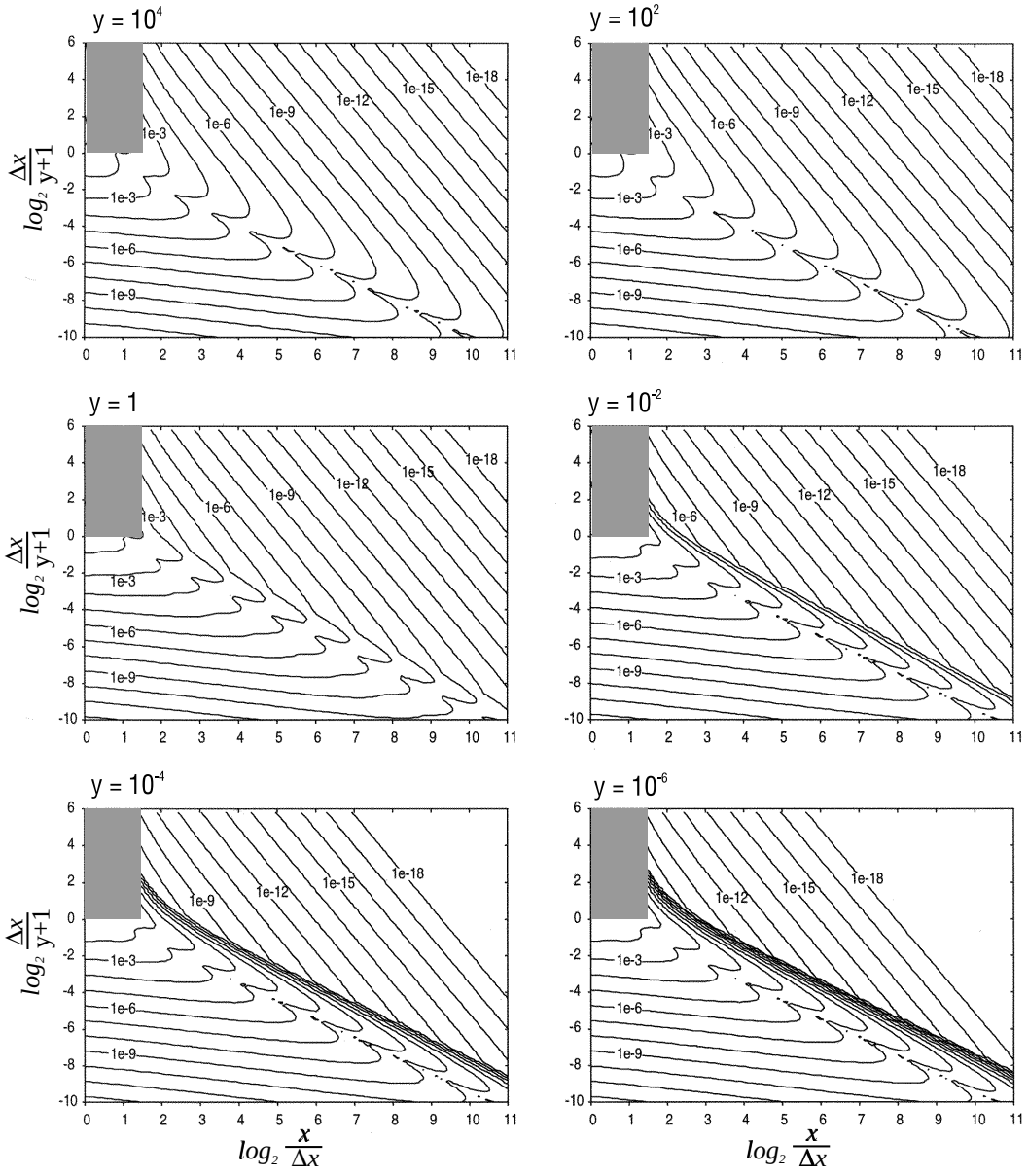

Figure 8: Proposed approximation for the determination of the Voigt-cutoffs. The light shaded regions indicate the domain of frequencies where lines are truncated.

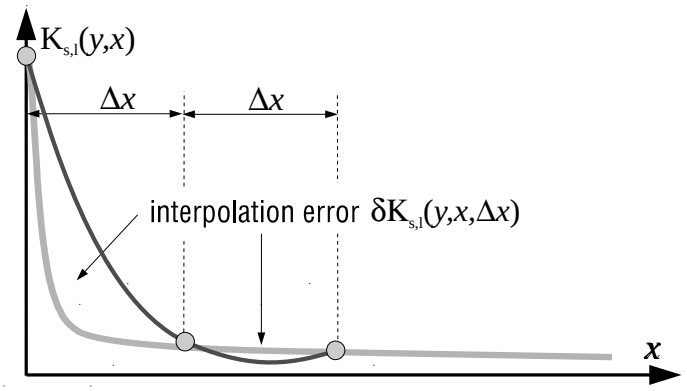

Figure 9: Approximating the profile function by a three-point Lagrangian interpolation for large frequency spacings Δx.

The quality of this new approximation is striking (compare the contours in Figure 8 with the true contours in Figure 6). Small deviations occurr only for $y \sim 1$. In addition, the evaluation of the new criterion requires only four extra floating point operations compared to the criterion proposed by Sparks [14]. This is certainly negligible compared to the achieved reduction of function evaluations. Furthermore, it should be mentioned that in the proposed algorithm the use of cutoff-tables is minimized so that all sampling points that have to be considered on a given frequency grid are determined with a single reference to the corresponding cutoff-tables, whereas Sparks [14] examines each sampling point individually.

4.3 Line rejection and truncation

Another question that has not yet been mentioned is which lines should be retained in the model and which lines may be rejected. Furthermore, we need a criterion to decide in which distance from the line center frequency spectral lines may be truncated without substantial loss of accuracy. This is especially true, if the far wings of lines are incorporated into models of the continuum and need not be taken into account explicitly [3]. Sparks [14] provides only two alternatives: to reject a spectral line completely, or to consider it over the entire frequency range of interest. In the proposed algorithm lines are only rejected, if $K_{s,l}(x = 0, y)$ is smaller than the maximum interpolation error that can be accepted for a given line or if the transition frequency of the line lies beyond a user-specified distance from the endpoints of the frequency range of interest. This distance is set equal to 25 cm^{-1}. Other lines may only be truncated.

A closer inspection of Figures 4-6 reveals that all contours converge to $\log_2(x/\Delta x) = 1$ for increasing values $\log_2(\Delta x/(y+1))$. This implies that the interpolation error becomes negligible only in a distance $x > 2\Delta x$, regardless how large the frequency spacing is. This can be understood if one recalls that each contour represents the maximum interpolation error within a region $\pm\Delta x$ around a given distance x from the line center frequency. The maximum interpolation error thus approaches $K_{s,l}(x = 0, y)$ for increasing frequency spacings Δx, as is shown in Figure 9. This also explains why the algorithm of Sparks [14] evaluates a contribution to absorption for each spectral line retained in the model over the entire frequency range of interest, rather than truncating it beyond a certain distance. In the proposed algorithm, we assume that the central part of each spectral line is considered on grids of finer frequency spacing so that spectral lines are truncated if $\log_2(x/\Delta x) \leq 1.5$

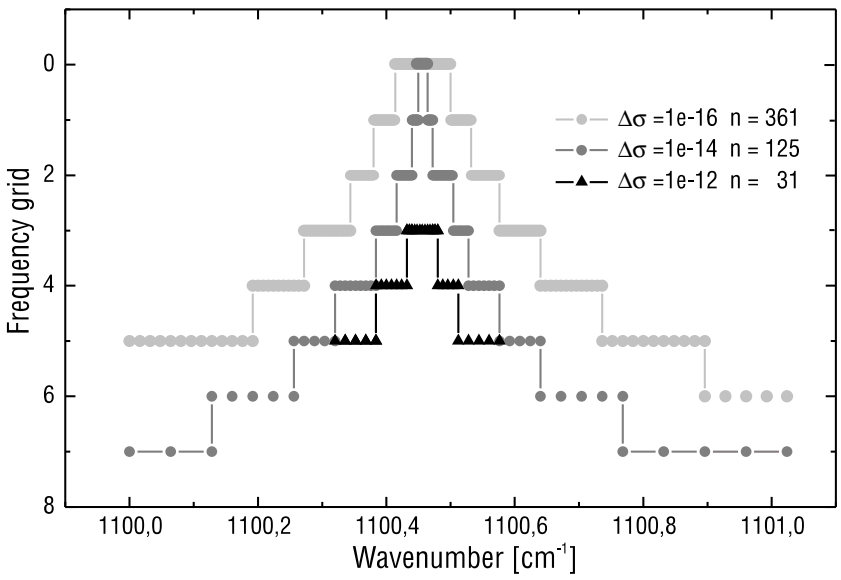

Figure 10: Representations of a given spectral line of maximum absorption $\sigma = 10^{-10}$ cm^{-1} for different absolute accuracies $\Delta\sigma$ on a set of successively finer frequency grids. The absolute accuracies $\Delta\sigma$ and the number of function evaluations are indicated.

and $\log_2(\Delta x/(y+1) \geq 0)$, as is indicated by the light-shaded regions in Figure (8). This assumption is strengthened by the fact that the frequency spacing of the finest frequency grid is usually chosen according to the Doppler half-width at high altitudes. Nevertheless, the truncation of lines may also be suppressed as an option so that each spectral line retained in the model is considered over the entire frequency range of interest on at least three sampling points. Test calculations revealed that the above criterion meets the requirement of negligible truncation errors quite well, especially if the frequency range of interest is small (microwindow approach) and the far wings are incorporated into models of the continuum.

Figure 10 displays different representations of a given spectral line with a maximum absorption coefficient of $\sigma = 10^{-10}$ cm^{-1} for different absolute accuracies $\Delta\sigma$. Note that the frequency spacing is finest near the line center frequency and successively increases with increasing distance from it. For the lowest accuracy $\delta\sigma = 10^{-12}$ cm^{-1}, 31 sampling points are adequate, while 361 sampling points are necessary to achieve an absolute accuracy $|\delta\sigma| \leq 10^{-16}$ cm^{-1}. It should be mentioned that for the lowest accuracy the spectral line has been truncated.

An example of the contribution of a single spectral line to absorption is shown in Figure 11 together with a reference spectrum calculated on an equidistant frequency grid of uniform intervals. The resulting interpolation error is shown below. The aim was an absolute accuracy $|\delta\sigma| \leq 10^{-16}$ cm^{-1} which could obviously be achieved over the entire frequency range of interest. Furthermore, it is interesting to note how the interpolation error always increases when the algorithm proceeds to a coarser frequency grid.

5 Efficient summation of spectral lines

The summation of spectral lines follows much of the methodology suggested by Fomin [5] and is performed on a set of successively finer frequency grids. Each of these grids is made up of intervals that consist of three sampling points due to the second order polynomial used for interpolation, see Figure 12. It should be

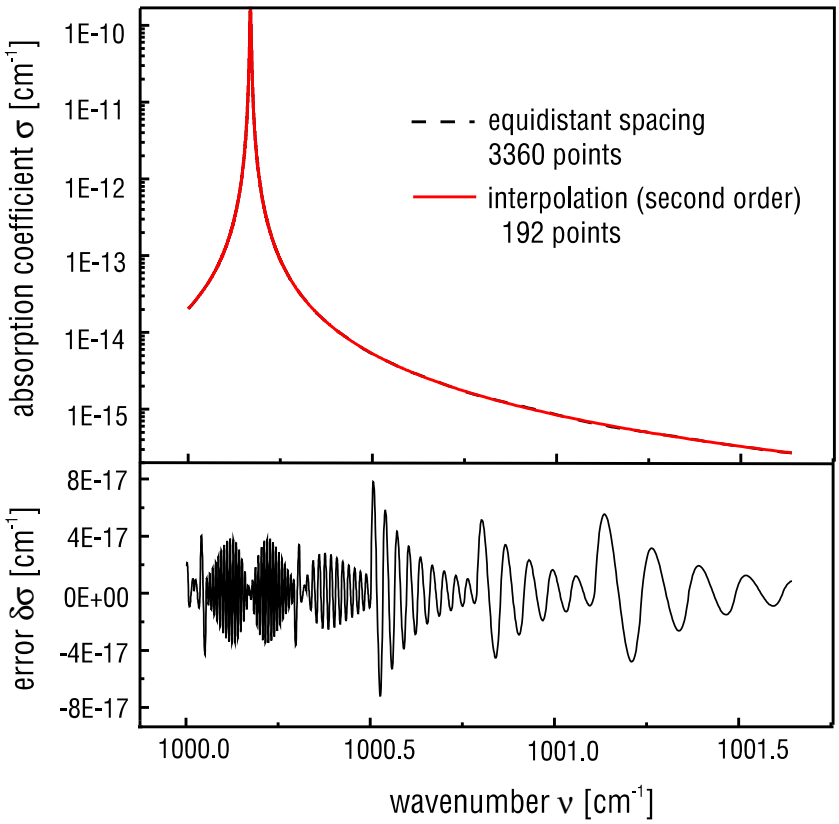

Figure 11: (top) Contribution of a single spectral line to absorption applying the new method of approximation together with the reference spectrum evaluated on an equidistant frequency grid of uniform intervals. (bottom) The resulting interpolation error.

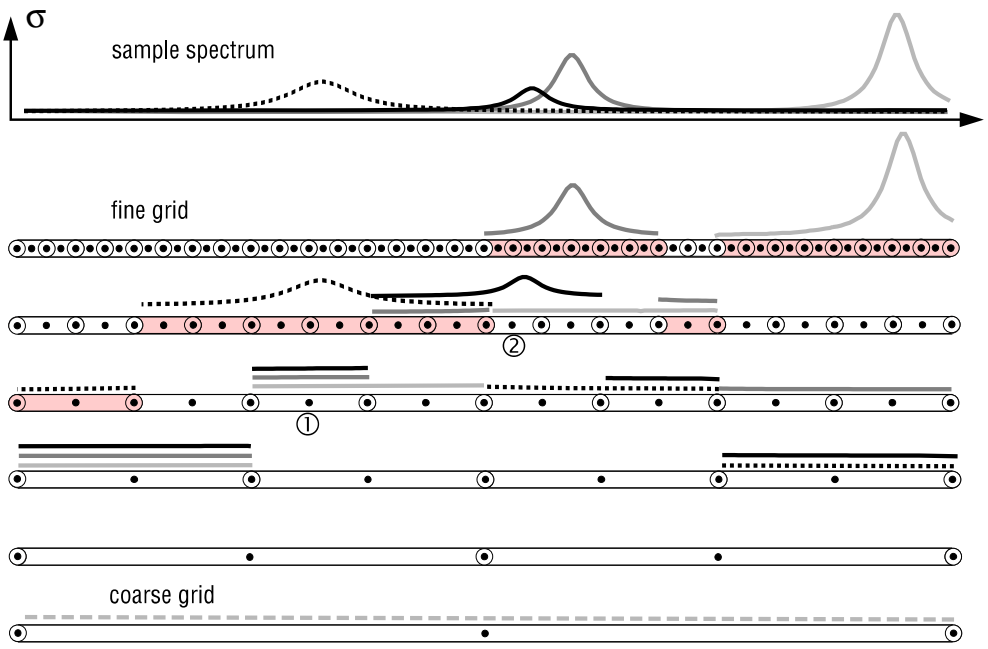

Figure 12: Illustration of the proposed interpolation technique on a set of successively finer frequency grids. The resulting quasi-uniform frequency grid is indicated by the light shaded intervals.

mentioned that the two endpoints must be considered individually for each interval (i.e. it takes two storage locations). This makes possible to consider each spectral line independently from the others so that each spectral line may be evaluated on its own optimum set of sampling points. The procedure is similar to the technique that has been applied in the widely used algorithm of Clough and Kneizys [2] or Gordley et al. [8] and in particular in the work of Fomin [5], except that the number of frequency grids for each line is determined dynamically.

The summation of spectral lines is performed in such a way that for each line the contribution to absorption is added up in the appropriate intervals. For example consider the interval marked by ① in Figure 12, to which three spectral lines of different strengths and half-widths contribute. Note that only those spectral lines, whose transition frequencies lie far outside the frequency range of interest, contribute to the interval residing on the coarsest frequency grid. Because each succeeding frequency grid is by a factor of two finer than the preceeding one, it is possible to move back to a coarser frequency grid only from every second interval. This may sometimes result in the calculation of an extra interval on a given frequency grid as for example for the light shaded spectral line in the intervals marked by ① and ②.

Having completed the summation of spectral lines, absorption coefficients are calculated in a straightforward manner starting on the coarsest frequency grid and moving progressively to grids of finer frequency spacing while interpolating and adding the coarser intervals to the appropriate intervals on the succeeding finer frequency grid. Note that by using an interpolation scheme with constant coefficients (e.g. Lagrangian interpolation) the contribution of each interval to the succeeding intervals can be calculated as a linear combination of the contributions to absorption evaluated on the current frequency grid. The interpolation step thus is computationally efficient, especially since it has to be performed only once after the summation over relevant lines has been completed.

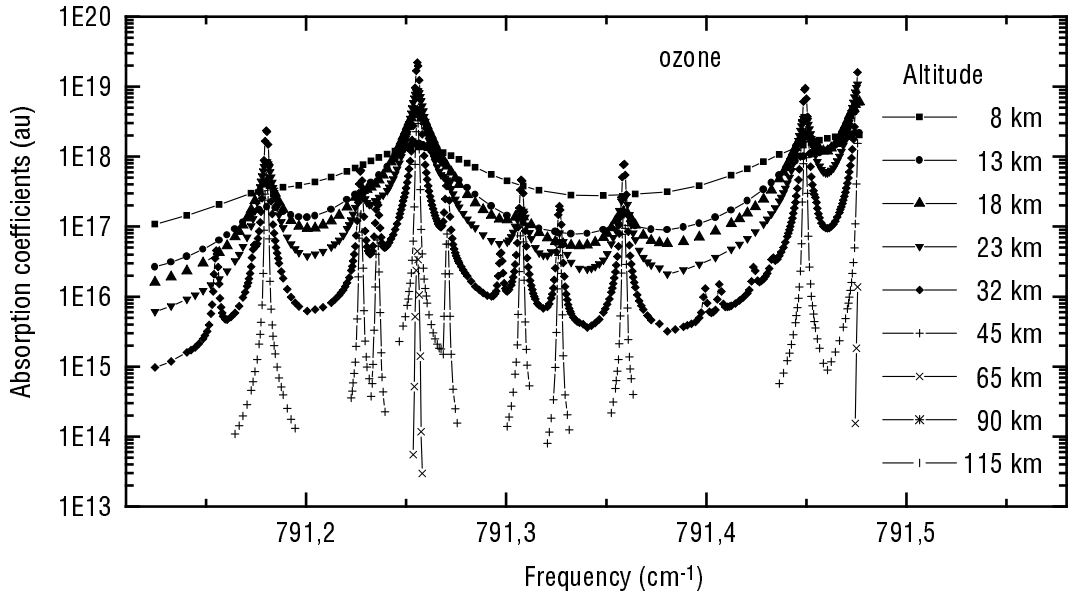

Figure 13: Absorption coefficients of ozone at different altitudes calculated for a limb sounding geometry and a tangent height of 8 km.

Rather than terminating the interpolation only if the finest frequency grid is reached, the interpolation may readily be terminated at the finest intervals for which a contribution to absorption has been evaluated. In this case one ends up with a nonuniform or nonequidistant frequency grid which is, however, based on an equidistant fine grid of uniform intervals, indicated by the light shaded intervals in Figure 12. Let us henceforth refer to this nonequidistant grid as "quasi-uniform" frequency grid. Compared to an equidistant frequency grid of uniform intervals this quasi-uniform grid saves valuable computer memory. Furthermore, the use of quasi-uniform frequency grids presents additional opportunities for optimizing the radiative transfer calculations, because the integration is performed only on the sampling points of the quasi-uniform frequency grid. Details of this integration are, however, beyond the scope of this section. A sequence of O_3 absorption coefficients calculated on quasi-uniform frequency grids at altitudes varying between 8 and 115 km is shown in Figure 13. For the calculation a limb sounding geometry with a tangent height of 8 km has been assumed. Note that the number of sampling points always decreases where the absorption coefficient varies only slowly with frequency. At high altitudes the contributions to absorption are extremely weak and so that spectral lines are truncated.

Figure 14 is a conceptional depiction of the basic steps involved in the proposed algorithm for calculating absorption coefficients. Firstly, the optimum set of sampling points for a given line is determined from the appropriate cutoff tables. Secondly, the sampling points are sorted into increasing order to allow for the rapid evaluation of the Voigt line shape factor. Thirdly, the contributions to absorption are collected into the appropriate intervals residing on a set of successively finer frequency grids. Finally, the coarser intervals are interpolated to the uniform or quasi-uniform frequency grid, respectively.

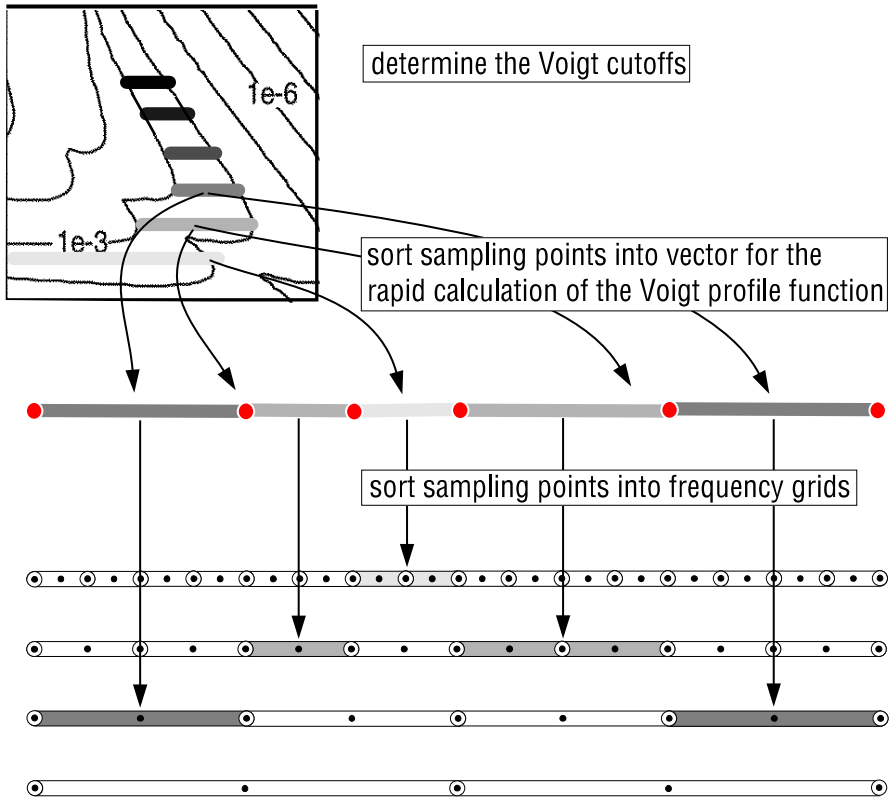

Figure 14: Conceptual depiction of the proposed technique for the efficient calculation of absorption coefficients line-by-line.

6 Implementation of the algorithm

The proposed algorithm has been implemented in ADDLIN, a Fortran90 module of $\sim$ 1000 lines of code. The objective in developing this module was to accommodate flexibility and simplicity in use without substantial loss of accuracy and efficiency. The result was a clear and modular approach that can easily be adapted as a research tool for specific needs separately from its specific use within KOPRA.
Besides the standard calculation of spectral absorption coefficients ADDLIN has been generalized to the application of more sophisticated line shape models. In general, using a different line shape model will require defining new tables of cutoffs analogous to the Lorentz and Doppler cutoffs discussed before. Considering line-mixing, for example, requires defining a separate table of cutoffs since the line shape factor decays asymptotically as $|\nu - \nu_l|^{-1}$ rather than $|\nu - \nu_l|^{-2}$. In practice, however, it may often prove sufficient to continue to use the Lorentz and Doppler cutoffs, especially if line mixing is combined with a line shape model that consists of a Lorentzian profile modified by a corrective factor $\chi_l(\nu, \nu_l)$ (often $\chi_l(\nu, \nu_l)$ takes the form of a decaying exponential with adjustable coefficients determined empirically). ADDLIN has been designed to support both line mixing and corrective $\chi_l(\nu, \nu_l)$-factors basing on the precalculated Lorentz and Doppler cutoffs. Moreover, considering line mixing requires the evaluation of the complex probability function which is closely related to the Voigt profile function. Let us henceforth refer to the complex probability function as $f_l(\nu, \nu_l)$; the Voigt profile function is referred to as real part of the complex probability function, $\mathrm{Re}(f_l(\nu, \nu_l))$, while the imaginary part is referred to as $\mathrm{Im}(f_l(\nu, \nu_l))$. For the calculation of the complex probability function ADDLIN uses the new implementation of the Humlicek algorithm described in section 3 except that the algorithm has been modified to provide also the imaginary part $\mathrm{Im}(f_l(\nu, \nu_l))$ as an option. The following paragraphs are a summary of how absorption coefficients are calculated within ADDLIN if line mixing and/or corrective $\chi_l(\nu, \nu_l)$-factors are taken into account.

Ordinary absorption coefficient: If neither line mixing nor corrective $\chi_l(\nu, \nu_l)$-factors are considered monochromatic absorption coefficient is calculated as

$$\sigma_{s,l}(\nu) = N_s \sum_l S_l \,\mathrm{Re}(f_l(\nu, \nu_l))\,. \tag{15}$$

This is the same expression as (1) on page 44.

Corrective $\chi_l(\nu, \nu_l)$-factors: If line profiles are to be modified by a corrective $\chi_l(\nu, \nu_l)$-factor, the absorption coefficients is evaluated as

$$\sigma_{s,l}(\nu) = N_s \sum_l S_l \,\chi_l(\nu, \nu_l)\,\mathrm{Re}(f_l(\nu, \nu_l))\,, \tag{16}$$

where the corrective $\chi_l(\nu, \nu_l)$-factor has to be provided by the user in form of an external subroutine (for details see the following interface section).

Line mixing: If line mixing is taken into account, the absorption coefficient is calculated as

$$\sigma_{s,l}(\nu) = N_s \sum_l S_l^* \,\mathrm{Re}(f_l(\nu, \nu_l^*)) + y_l^* \,\mathrm{Im}(f_l(\nu, \nu_l^*))\,, \tag{17}$$

where the summation is performed over all lines of the line-mixing branch. The y_l^* are coefficients that have to be provided by the user. KOPRA provides two different line mixing models for the calculation of these coefficients. Note, that depending

on the line mixing model used, the linestrengths, line center frequencies and line widths may also be modified. This is indicated by the asterisks in (17). For more details on line-mixing and the calculation of the y_l^* coefficients see Part VI: 'Line mixing' in this issue.

Corrective $\chi_l(\nu,\nu_l)$-factors and line mixing:
If both corrective $\chi_l(\nu,\nu_l)$-factors and line mixing are taken into account the calculation of absorption coefficients requires two steps. In a first step an averaged corrective $\tilde{\chi}(\nu)$-factor for the whole line-mixing branch is determined from

$$\tilde{\chi}(\nu) = \frac{\sum_l S_l^* \, \chi_l(\nu,\nu_l^*) \, \mathrm{Re}(f_l(\nu,\nu_l*))}{\sum_l S_l^* \, \mathrm{Re}(f_l(\nu,\nu_l^*))}.$$

Note, that the averaged corrective $\tilde{\chi}(\nu)$ may be determined either from the original or the modified line-strengths, line-widths, and line center frequencies (the latter marked by an asterisk). ADDLIN uses the modified line-strengths, line-widths, and line center frequencies as indicated in the above equation. In a second step the averaged corrective $\tilde{\chi}(\nu)$-factor then is multiplied with the result of the pure line-mixing absorption coefficient to give

$$\sigma_{s,l}(\nu) = N_s \, \tilde{\chi}(\nu) \sum_l S_l^* \, \mathrm{Re}(f_l(\nu,\nu_l*)) + y_l \, \mathrm{Im}(f_l(\nu,\nu_l^*)) \, . \qquad (18)$$

Note that the two steps are performed automatically, if ADDLIN is invoked appropriately via its user-interface.

7 User interface

The calculation of absorption coefficients within ADDLIN is performed in an almost automated way which enables workers to perform computations quickly without detailed knowledge of the calculation methods involved. Workers may communicate with ADDLIN only via 9 subroutine invocations, which will be described in detail in the following:

`allocate_cutoff(cutdop, cutlor)` Allocates memory for a number of internal arrays and initializes some auxiliary variables. Reads the appropriate look-up-tables for the Doppler- and Lorentz limit, respectively, from the files specified in the strings `cutdop` and `cutlor`. This is the first subroutine to be called by the user. It should be invoked once before the calculation of absorption coefficients starts.

`deallocate_cutoff` This subroutine is the exact counterpart of `allocate_cutoff` and should be invoked once after the calculation of absorption coefficients has been completed.

`allocate_grid (fmin, fmax, fdel)` Determines the spectral range of interest (`fmin`, `fmax`) and the desired spectral resolution `fdel`. Allocates memory for the internal set of successively finer frequency grids. The subroutine should be invoked each time a new spectral range is considered. The units of `fmin`, `fmax`, and `fdel` should be consistent with the units of `alphaL`, `alphaD`, and `f0` in the subroutines `add_lines`, `add_lines_chi`, `add_lines_lm`, and `add_lines_chilm`.

`deallocate_grid` This subroutine is the counterpart of `allocate_grid` and deallocates the arrays specific to a given spectral range of interest. It should be

invoked each time after the calculation of absorption coefficients for a given spectral range has been completed.

`add_lines (nlines, iflag, alphaL, alphaD, ratio, f0, fac, S, dSdT)`
Calculates the absorption coefficient according to (15), i.e. `add_lines` performs the summation of the `nlines` spectral lines with Lorentz half-widths `alphaL`, Doppler half-widths `alphaD`, line center frequencies `f0`, and line strengths `S`. Note that `alphaL` and `alphaD` are defined as the half-width at 1/2 of the maximum of the profile (rather than $1/e$ of the maximum as in the definition of $\alpha_{D,s,l}$ in section 3). As an option, the user may also pass the temperature derivatives of the line strenghts `dSdT`, in which case `add_lines` also provides the temperature derivative of the spectral absorption coefficient as an option. It should be mentioned that the temperature derivative of the absorption coefficient is calculated by summing the temperature derivatives of the line strengths. Temperature dependences of the line-widths e.g. are not taken into account. Test calculations revealed, however, that this approximation is accurate within 5% or better.

`iflag` determines how the calculation of the absorption coefficient is performed in the detail.

- If `iflag = 0`, weak spectral lines may be rejected and those lines retained in the model may be truncated at their line wings.
- If `iflag = 1`, all spectral will be retained in the model and will be considered over the entire frequency range of interest on at least three sampling points. The truncation of lines is suppressed.

`ratio` determines the accuracy to which spectral lines are considered in the model:

- If `ratio` is chosen to be negative, the spectral absorption coefficient is calculated by summing directly the contribution of each line to absorption at each frequency. Note that the calculation of this reference absorption coefficient may be extremely time consuming.
- If `ratio` is chosen to be ratio of the desired absolute accuracy $\delta\sigma$ to which the contribution to absorption of a given line is to be calculated to the number density of the species N_s, i.e. `ratio` $= \delta\sigma/N_s$, the absorption coefficient is calculated in such a way that the absolute accuracy of the contribution to absorption of each spectral line is better than $\delta\sigma$.
- `Ratio` may also be chosen to be the ratio of the desired absolute accuracy $\delta\tau$ to which the contribution of a single line to the optical depth of a given atmospheric layer is to be calculated to the column density $N_s\Delta s$, i.e. `ratio` $= \delta\tau/(N_s\Delta s)$, where Δs is the optical path-length through the layer. In this case the absorption coefficient is determined in such a way that the absolute accuracy to which the contribution of each spectral line is calculated times the optical path-length Δs yields $\delta\tau$.

`fac` is a factor the absorption coefficient is multiplied with.

- If `fac` is set equal to the volume mixing ratio of a given species, the result will be the absorption coefficient itself.
- If `fac` is set equal to the column density, the result will be the optical depth of the given atmospheric layer.
- If `fac` is set equal to 1, the result will be the absorption coefficient divided by the column density.

```
interface                                      ! interface to the user-defined
  subroutine sub(x,n,T,y,alphaL,alphaD)        !   external subroutine SUB
    use precis_m                               ! definition of precision
    integer, intent(in)  :: n                  ! number of sampling points
    real(dp), intent(in) :: T                  ! temperature
    real(dp), optional   :: alphaD,alphaL      ! Doppler and Lorentz half widths
    real(dp), dimension(*), intent(in)   :: x ! distance from line center
    real(dp), dimension(*), intent(inout):: y ! ratio Lorentz to Doppl. halfwidth
  end subroutine sub
end interface
```

Figure 15: Subroutine interface for the corrective $\chi_l(\nu, \nu_l)$ factor to be provided by the user. For the definition of x and y see section 3.

Note, that the units of `ratio` (e.g. cm^2) and `alphaL`, `alphaD`, and `f0` (e.g. GHz) must be consistent with the units of the line strength `S` (e.g. $GHz\,cm^2$) and the units of `fmin`, `fmax`, and `fdel` (e.g. GHz) in `allocate_grid`. Furthermore, the user should take care that if `add_lines` is invoked several times in succession, the optional parameter `dSdT` is either specified or not specified in all subroutine invocations.

`add_lines_chi (nlines, iflag, alphaL, alphaD, ratio, f0, fac, S, sub T, dSdT)` Same as `add_lines` except that each spectral line is multiplied by a corrective $\chi_l(\nu, \nu_l)$-factor according to (16). This corrective $\chi_l(\nu, \nu_l)$-factor) has to be provided by the user in form of an external subroutine, which then is used as an argument (`sub`) in the invocation of `add_lines_chi` and should be declared according to the interface block shown in Figure 15. The input variable `T` is the temperature in K and has to be provided because the user-defined $\chi_l(\nu, \nu_l)$-factor may be temperature-dependent. If the optional parameter `dSdT` is passed, `add_lines_chi` calculates temperature derivatives as an option.

`add_lines_lm (nlines, iflag, alphaL, alphaD, ratio, f0, fac, S, ycoef, dSdT)` Same as `add_lines` or `add_lines_chi` except that line mixing is taken into account according to (17). For this purpose the user has to provide a dimensionless y_l^* coefficient (`ycoef`) for each spectral line of the line-mixing branch (for the calculation of these y_l^* coefficients see Part VI: 'Line mixing' in this issue or [*Funke et al.* 1997]). As an option `add_line_lm` may also calculate the T-derivative of the absorption coefficient depending on whether the optional argument `dSdT` is passed or not.

`add_lines_chilm (nlines, iflag, alphaL, alphaD, ratio, f0, fac, S, sub T, ycoef, dSdT)` takes into account both line mixing and a corrective $\chi_l(\nu, \nu_l)$ factor according to (18). Note, that `add_lines_chilm` performs the two necessary steps, i.e. the calculation of the averaged corrective $\chi_l(\nu, \nu_l)$ factor and its following multiplication with the line mixing absorption coefficient automatically. This enables workers to perform computations quickly without detailed knowledge of the calculation methods involved. Workers only have to make sure, that each invocation of `add_lines_chilm` comprises all spectral lines of a given line-mixing branch. I.e. a given line-mixing branch may not be split into bundles of lines and `add_lines_chilm` invoked several times for each of these bundles. Again, as an option `add_lines_chilm` calculates the temperature derivative of the absorption coefficient, if the optional argument `dSdT` is passed.

`interpolate_grid (iabcomx, iabco, abco, dabcodT, equidistant)` performs the final interpolation of all contributions to absorption calculated on the set

of successively finer frequency grids either to the quasi-uniform frequency grid or the uniform fine grid, depending on whether the optional logical input parameter `equidistant` is passed and true or not. The output variable `iabcomx` indicates how many sampling points are necessary to store the absorption coefficient while `iabco` is a pointer on an integer vector that contains the sampling points. Note, that all sampling point are stored as integers and that the real frequency of the `i`th sampling point is `fmin+iabco(i)*fdel`. The output variables `abco` and `dabcodT` are pointers on the spectral absorption coefficient and its temperature derivative, respectively. The temperature derivative is only calculated if the optional variable `dabcodT` is passed and if the `dSdT` has been passed on in the preceding invocations of `add_lines`, `add_lines_chi`, `add_lines_lm`, or `add_lines_chilm`.

Bibliography

[1] Abramowitz, M. A. and I. A. Stegun, Eds., Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series 55, 1968.

[2] Clough, S. A. and F. X. Kneizys, Convolution Algorithm for the Lorentz function, Appl. Optics, 18, 2329-2333, 1979.

[3] Clough, S. A., F. X. Kneizys, and R. W. Davies, Line shape and the water vapor continuum, Atmospheric Research, 23, 229-241, 1989.

[4] Edwards, D. P., Modeling of the Atmosphere, SPIE Modeling Atmos., 928, 94-116, 1988.

[5] Fomin, B. A., Effective interpolation technique for line-by-line calculations of radiation absorption in gases, J. Quant. Spectrosc. Radiat. Transfer, 53, 663-669, 1995.

[6] Funke, B., G. P. Stiller, T. von Clarmann, G. Echle, and H. Fischer, CO_2 line mixing in MIPAS limb emission spectra and its influence on retrieval of atmospheric parameters, J. Quant. Spectrosc. Radiat. Transfer, 59, 215-230, 1998.

[7] Gamache, R.R., R.L. Hawkins, L.S. Rothman, Total internal partition sums for atmospheric molecules in the temperature range 70-2005 K: Atmospheric linear molecules, J. Mol. Spectrosc. 142, 205-219, 1990.

[8] Gordley, L. L., B. T. Marshall, and D. A. Chu, LINEPAK: Algorithms for modeling spectral transmittance and radiance, J. Quant. Spectrosc. Radiat. Transfer, 52, 563-580, 1994.

[9] Humlicek, J., Optimized Computation of the Voigt and complex probability fuction, J. Quant. Spectrosc. Radiat. Transfer, 27, 437-444, 1982.

[10] Kuntz, M., A new implementation of the Humlicek algorithm for the calculation of the Voigt profile function, J. Quant. Spectrosc. Radiat. Transfer, 75, 819-824, 1997.

[11] Kuntz, M. and M. Höpfner, Efficient line-by-line calculation of absorption coefficients, J. Quant. Spectrosc. Radiat. Transfer, 63, 97-114, 1999.

[12] Schreier, F., The Voigt and complex error function: A comparison of computational methods, J. Quant. Spectrosc. Radiat. Transfer, 48, 743-762, 1992.

[13] Sparks, L., Accelerated line-by-line calculation of spectral absorption coefficients with high numerical accuracy, in Optical Remote Sensing of the Atmosphere, 2, 1995 OSA Technical Digest Series, 68-70, Optical Society of America, Washington DC, 1995.

[14] Sparks, L., Efficient line-by-line calculation of absorption coefficients to high numerical accuracy, J. Quant. Spectrosc. Radiat. Transfer, 57, 631-649, 1997.