ELSEVIER

Journal of Quantitative Spectroscopy & Radiative Transfer 107 (2007) 173–192

Journal of Quantitative Spectroscopy & Radiative Transfer

www.elsevier.com/locate/jqsrt

Rapid and accurate calculation of the Voigt function

Kendra L. Letchworth, D. Chris Benner*

Department of Physics, College of William and Mary, Box 8795, Williamsburg, VA 23187 8795, USA

Received 8 September 2006; received in revised form 17 January 2007; accepted 18 January 2007

Abstract

Two general techniques significantly improve both the accuracy and speed of spectral line shapes that use only the complex Voigt function. Such line shape functions include the Voigt line profile, which is closely related to the real part of this function. The first technique is a new algorithm, which trades a small amount of RAM (at most 1.49 megabytes) for a considerable gain in speed and accuracy. The accuracy is one part in 10^6 of the function itself. In addition, the derivatives of the function with respect to x and y are returned with an accuracy of 0.5%. The algorithm is up to nine times faster than the Drayson or Humlíček algorithms and about two orders of magnitude more accurate. A second independent improvement is a set of simple criteria to decide when an evaluation is required and when a value of zero can be assumed. This section alone can reduce the computation time by three orders of magnitude or more and it can be applied to algorithms using any Voigt approximation.

Keywords: Voigt function

1. Introduction

The Voigt function is widely used for the calculation of spectral line profiles and in other areas of physics [1]. The Voigt profile is a convolution of the Lorentz and Doppler line profiles and can also be generalized to include the effect of Dicke narrowing in the hard collision limit [2]. The Voigt function is closely related to the complex error function and inclusion of the corresponding imaginary part allows for calculation of the spectral line profile including line mixing [3,4] as well as derivatives of these profiles [5]. Many of the numerous algorithms for the computation of the Voigt function with or without the imaginary part have been summarized by Refs. [1,5–7].

The numerical simulation of atmospheric or laboratory spectra is often limited by the speed and accuracy of the Voigt function approximation used. For an infrared spectrum of typical atmospheric gases that includes gas at pressures low enough that the Doppler width is larger than the Lorentz width, the lines are well characterized only if the wavenumber grid is of the order of $10^{-3}\,\mathrm{cm}^{-1}$ or smaller. Simulation of a $100\,\mathrm{cm}^{-1}$ wavenumber interval with thousands of spectral lines requires more than 10^8 evaluations of the Voigt function. This computational burden may be exacerbated by the inclusion of multiple atmospheric layers at different temperatures and pressures or the inclusion of numerous laboratory experiments, each performed

*Corresponding author. Tel.: +1 757 221 3531; fax: +1 757 221 3540.
E-mail address: dcbenn@wm.edu (D.C. Benner).

0022-4073/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jqsrt.2007.01.052

Nomenclature

A	absorption
A_m	absorption cutoff
e	base of natural logarithms
f	function which is a portion of an integrand
i	counter in sum of Gauss–Hermite quadature [in Eqs. (7)–(9) and Eqs. (19)–(20)]
i	$\sqrt{-1}$ [Eqs. (2) and (25)]
j	counter over Lagrange polynomial indices
k	counter over Lagrange polynomial indices
K	real part of W (Voigt function)
K_1	first order quadrature approximation of K
L	imaginary part of W
L_1	first-order quadrature approximation of L
n	order of quadrature in approximations
p	Lagrange polynomials
S	spectral line intensity
S_m	spectral line minimum intensity for which lines are included
t_i	zero points of Hermite polynomials
t	variable of integration in Voigt profile
T	transmission through the gas
u	mass path of gas
v	value at which polynomial is to be determined
v_0,v_1,v_2	tabulated values of polynomial
V	Voigt profile
w_i	weights of Hermite polynomials
W	function closely related to the complex error function
x	$\sqrt{\ln 2} \times (v - v_0)/\alpha_D$
x_0	x-values at precomputed points in Taylor expansion
y	$\sqrt{\ln 2} \times \alpha_L/\alpha_D$
y_0	y-values at precomputed points in Taylor expansion
z	$x + iy$
α_D	Doppler half-width
α_L	Lorentz half-width
Δv	$(v - v_1)/(v_2 - v_1)$
Δx	$x - x_0$
Δy	$y - y_0$
ϵ	relative error
ν	spectral frequency
ν_0	spectral line center frequency
π	ratio of circumference to diameter of a circle

under different temperature and pressure conditions. Each layer in the atmosphere or each spectrum requires more than 10^8 evaluations of the Voigt function. In fitting situations, derivatives with respect to a very large number of spectral line parameters or physical conditions could increase this calculation by a factor of 10 to even 1000. Furthermore, a number of iterations may be required to produce a good fit.

The transmission at any point on a line profile is most sensitive to the normalized line shape where the optical depth is unity. Since optical depth is a product of line intensity, gas amount and the normalized line shape, this can occur at any point on the spectral line profile. There are very different applications of spectroscopy in which almost any point on any profile can be important. At one optical depth a given relative

change in the normalized line shape affects the transmission by an amount smaller by a factor of $1/e$. The required accuracy in the line shape function is the relative accuracy of the line shape function times $1/e$ and divided by the noise level over the signal. Since there is generally more than one point contributing to the measurement, the contribution of the noise tends to average out somewhat and the line shape function relative accuracy required to match within the noise is close to the inverse of the signal to noise. Of course, weak lines do not have a point near one optical depth and most of the points in a strong line will not be near one optical depth, so most evaluations of the line shape function do not require this much accuracy, but when devising an algorithm it is impossible to know which points will need the full accuracy. For the Fourier transform spectrometer at the National Solar Observatory solar telescope at Kitt Peak, the signal to noise can be as high as 7×10^4 [8], so it is sometimes required that the relative error in the line shape function be smaller than 10^{-5} (or perhaps even less for other instruments). Atmospheric spectra now can measure to a precision of 10^{-3} [9], requiring laboratory data at levels somewhat better. Many available Voigt function algorithms strive only for the accuracy needed by the author and only then in line profile calculations practical in their particular field.

Analyses of our laboratory data have depended primarily upon the Drayson [10] algorithm when the imaginary portion of the Voigt function is not required and the Humlíček [11] algorithm when the imaginary part is required. For each of these routines small changes were made to optimize their speeds, but their numerical accuracy was not changed. As shown by Schreier [5], the Drayson routine is comparable in speed or somewhat faster than the Humlíček routine, but has a relative accuracy that is three times worse. The worst relative accuracy, other than near the Doppler limit, for each of these routines is of the order of one part in 10^4, at times not acceptable. In some regions (especially the quadrature regions) these errors are systematic and do not benefit from averaging errors of opposite signs. Fig. 1 displays the relative accuracy of a typical NO line simulated in a retrieval for HALOE [12]. The line has a Lorentz width of $0.06\,\mathrm{cm}^{-1}$/atm, a Doppler width of $0.002\,\mathrm{cm}^{-1}$ and the atmospheric pressure at the top of the retrieved portion of the atmosphere is about 10^{-9} atm. For the figure y (defined later in Eq. (4)) is 2×10^{-8} and the vertical axis is the value of the computational error of the Voigt function relative to the function value itself. The relative accuracy of the Voigt evaluation as a function of distance from line center is plotted for several popular algorithms. For $|x| < 3$, (where x, defined later in Eq. (3), is the distance from line center in units close to the Doppler width) all of the algorithms do well since a relative error a few times smaller than the inverse of the signal to noise ratio is all that is required. For $|x|$ between 3 and 5.5, though, the Drayson [10] and especially the Pierluissi [13] algorithms have very significant errors. The Pierluissi [13] approximation returns values five orders of magnitude too small at x just over 3. As the Lorentz width is decreased even farther, even the Humlíček routine reports values in error by orders of magnitude [14]. The algorithm presented in this paper always returns a value with a relative error smaller than 10^{-6} and, unlike most other routines, the relative error tends toward zero as the Doppler limit is approached. Thus, there is no requirement that the accuracy of the algorithm be checked for a particular application very different than ours. Such accuracy is always sufficient for spectra with signal to noise ratio less than 2 or 3×10^5. Optionally, the derivatives of the Voigt function are returned with a relative accuracy of 0.5% at a cost in computation time of, at most, a factor of 1.7.

The remainder of this paper will present two techniques for speeding Voigt calculations with one of these also making them more accurate. The first technique, presented in Sections 2–5, is a new algorithm for the evaluation of the Voigt function which takes advantage of the increased RAM available in current computers compared to those in use two or three decades ago when many of the Voigt algorithms were published. The second technique, presented in Section 6, is a technique to reduce the number of function evaluations required without significant deterioration in the accuracy of the results. Section 6 is independent of Sections 2–5 and can be implemented with any Voigt algorithm.

2. Basic equations

The Voigt profile, $V(x,y)$, can be defined in terms of the complex Voigt function, $W(x,y)$, closely related to the complex error function [5].

$$V(x,y) = \frac{\sqrt{\ln 2}}{\alpha_D \sqrt{\pi}} Re[W(x,y)]. \tag{1}$$

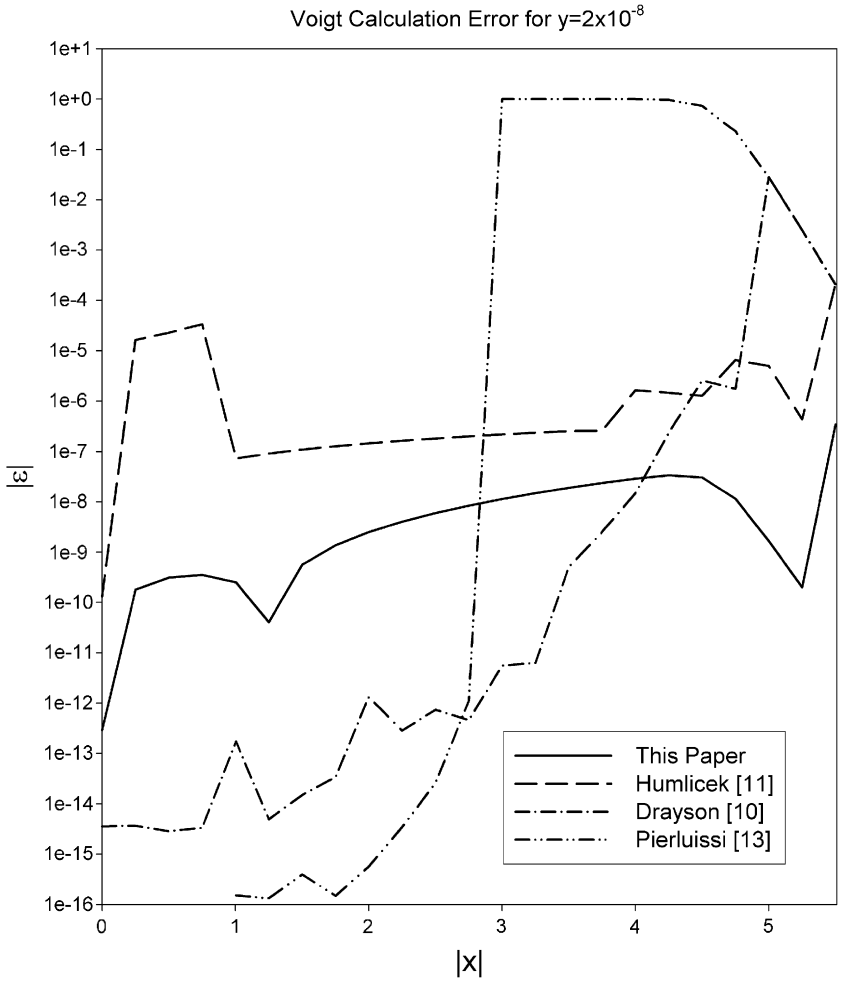

Fig. 1. Errors in the calculation of various approximations to the real part of the Voigt function as a function of $|x|$ for $y = 2 \times 10^{-8}$, near the Doppler limit. The "true" value of the function is taken as the Maple^{TM} calculation.

Here, α_D is the Doppler half-width of the line as defined by Benner et al. [4] and the function $W(z)$ is defined in Eq. (7.1.4) of Abramowitz and Stegun [15] and Eq. (5) of Schreier [5].

$$W(z) = \frac{\mathrm{i}}{\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{-t^2}}{z - t} \, \mathrm{d}t \tag{2}$$

Here z is $x + \mathrm{i}y$, i is $\sqrt{-1}$ and t is the variable of integration. For the sake of simplicity, $W(z)$ can be broken into its real, $K(x, y)$, and imaginary, $L(x, y)$, parts. The other variables on the right side of Eq. (1) are defined in terms of more basic quantities in Eqs. (3) and (4).

$$x = \sqrt{\ln 2} \frac{\nu - \nu_0}{\alpha_D}, \tag{3}$$

$$y = \sqrt{\ln 2} \frac{\alpha_L}{\alpha_D}. \tag{4}$$

In these equations ν is the spectral frequency, ν_0 is the spectral frequency at line center and α_L is the Lorentz half-width of the spectral line.

Computation of both the real and imaginary parts of W (the Voigt function) is helpful since its imaginary part can be used for the computation of line mixed spectral profiles [4,16] and the Rautian line shape which combines Lorentz broadening, Doppler broadening and Dicke narrowing in the hard collision limit [2].

Additionally, the derivatives of $W(x,y)$ can be used to calculate the derivatives of all of these profiles with respect to various spectral line parameters and physical conditions of the gas. The calculation of the derivatives of $W(x,y)$ is described in more detail in Section 5. Thus, the Voigt function is useful for more than just the Voigt profile; it is the basis of other, more accurate profile calculations. Some published Voigt algorithms such as Humlíček [11] include the evaluation of the imaginary part of $W(x,y)$ but many such as Drayson [10] and Pierluissi [13] do not.

In the low-pressure limit (y small), the real part of the Voigt function approaches the Doppler limit [17].

$$K(x,0) = \mathrm{e}^{-x^2}. \tag{5}$$

In the high-pressure limit, the real part of the Voigt function approaches the Lorentz limit [17].

$$K(x, y \to \infty) = \frac{1}{\sqrt{\pi}}\left(\frac{y}{x^2+y^2}\right). \tag{6}$$

3. Numerical approximations

The evaluation of the Voigt integral in Eq. (2) can be slow and inaccurate. Numerous approximations have been proposed to supply a compromise between the speed and accuracy of evaluation. In almost all cases a subroutine call is made for each pair of x, y values, as defined in Eqs. (3) and (4), that is each spectral point for each spectral line. However, in molecular spectroscopy there are generally a large number of calls to this subroutine with the same value of y and different values of x. This corresponds to the calculation of points over a single spectral line profile. Furthermore, the values of x are typically evenly spaced. For the algorithm presented in this work, a single call to a subroutine calculates all of these values for a spectral line by specifying the grid of x values and the single y value in the input. This operational change allows for significant improvements in the algorithm, which are otherwise not possible. For each approximation region, the portion of the computation depending upon y alone is calculated only once for the spectral line. Furthermore, calculation of the real and imaginary parts of the function together allows some computed values for each x to be shared.

Three types of approximations are used. The choice of method for a given case is based upon the values of x and y. Two of them are further subdivided into various levels of approximation.

3.1. Gauss–Hermite quadrature

Gauss–Hermite quadrature is a rapid way to evaluate an integral of the form of Eq. (2) by assuming that the integrand can be well approximated by $\exp(-t^2)$ times a polynomial, $f(t)$. Since $\exp(-t^2)$ appears in the integrand on the right side of Eq. (2), one would expect this to be a good approximation as long as the rest of the integrand behaves as a low-order polynomial.

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2} f(t)\,\mathrm{d}t \cong \sum_{i=1}^{n} w_i f(t_i). \tag{7}$$

The weights, w_i, and the zero points of the Hermite polynomials, t_i, are tabulated in Ref. [15] as a function of the order of quadrature, n. Odd-order quadrature is advantageous due to the fact that one of the values of t_i is always zero and $f(0)$ is calculable with fewer arithmetic operations. Further simplification is possible given that the remaining evaluation points come in pairs, which have identical weights and whose values of t_i are the negatives of each other. For $K(x,y)$, the integral is evaluated with Eq. (8).

$$K(x,y) \cong \frac{w_{(n+1)/2}\,y}{\pi(y^2+x^2)} + \frac{1}{\pi}\sum_{i=1}^{(n-1)/2} \frac{2w_i y(y^2+x^2+t_i^2)}{(y^2+x^2+t_i^2)^2 - 4x^2t_i^2}. \tag{8}$$

Similar manipulation results in the following evaluation for $L(x,y)$.

$$L(x,y) \cong \frac{w_{(n+1)/2}\,x}{\pi(y^2+x^2)} + \frac{1}{\pi}\sum_{i=1}^{(n-1)/2} \frac{2w_i x(y^2+x^2+t_i^2-2t_i^2)}{(y^2+x^2+t_i^2)^2 - 4x^2t_i^2}. \tag{9}$$

These calculations are very efficient despite the apparent complexity of Eqs. (8) and (9). Instead of storing the values of t_i in the routine, the values of t_i^2 and $2t_i^2$ are stored instead. For a series of evaluations with only a change in x, y^2 is unchanged and needs to be computed just once for the whole series. The series of values of $y^2+t_i^2$ are the same for a series of evaluations with the same y. The denominators inside the two summations are the same and can be reused. The intermediate value $y^2+x^2+t_i^2$ in the calculation of those denominators can be used in the numerators within the sums. Instead of storing the weights, w_i, storing $w_{(n+1)/2}/\pi$ and $2w_i/\pi$ eliminates a number of multiplication and division operations. The expression $w_{(n+1)/2}/[\pi(y^2+x^2)]$ need only be calculated once for the two equations. After the initial setup for a series of evaluations with the same y values, third-order quadrature ($n=3$) only requires 17 floating point calculations for evaluation of both the real and imaginary parts of the Voigt function at each x value. Without these shortcuts, the same calculation requires 50 floating-point calculations at each x value.

In the various regions of the x,y plane in which Gauss–Hermite quadrature is used, different orders of quadrature are required for a given accuracy. Lower quadrature orders compute faster, but are less accurate. For our algorithm, first, third, fifth and seventh orders are used as displayed in Fig. 2 and tabulated in Table 1.

3.2. Taylor series expansion about precomputed points

Refs. [18,19] have precomputed a grid of values for $K(x,y)$ for various x, y pairs and interpolated in these tables. For values of x and y both smaller than about 5, the behavior of the function is very nonlinear and this interpolation is not always accurate. A table storing precomputed values of the complex function $K(x,y)+\mathrm{i}L(x,y)$ and its derivatives with respect to x and y provides a better solution. The interpolation then expands the closest point in a Taylor series. The required number of stored function values and derivatives is much smaller for Taylor expansion than for standard spline interpolation due to the much larger spacing allowable between the precomputed points and some relationships among the derivatives.

Computation time and storage space for the partial derivatives are saved by successive applications of three mathematical relations.

$$\frac{\partial K(x,y)}{\partial x}=\frac{\partial L(x,y)}{\partial y}, \tag{10}$$

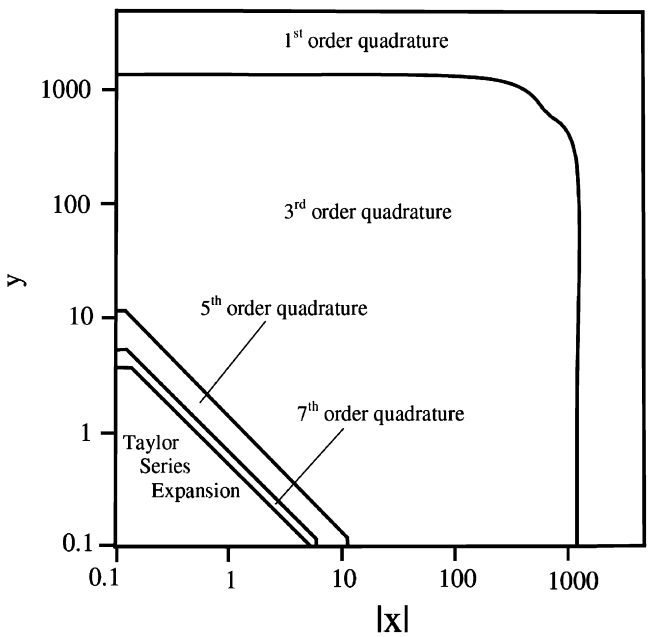

Fig. 2. Approximation regions for the real and imaginary parts of the Voigt function.

Table 1
The boundaries between the approximation regions in Figs. 2 and 3.

Approximation region	Boundaries
Third-order Taylor	$2.905 \leqslant y < 3.75$ and $\lvert x\rvert \leqslant 0.08$
Series (real)	$1.895 \leqslant y < 2.905$ and $\lvert x\rvert \leqslant 0.06$
Lagrange interpolating	$0.705 \leqslant y < 1.895$ and $\lvert x\rvert \leqslant 0.04$
Polynomials (imaginary)	$0.007 \leqslant y < 0.705$ and $\lvert x\rvert \leqslant 0.02$
Lagrange interpolating	$y \leqslant 2 \times 10^{-5}$ and $\lvert x\rvert \leqslant 5.75$
Polynomials	$2 \times 10^{-5} \leqslant y < 0.007$ and $\lvert x\rvert \leqslant 0.01$
Third-order Taylor series	
Spacing 0.08	$2.905 \leqslant y < 3.75$ and $0.08 \leqslant \lvert x\rvert \leqslant 5.3-y$
Spacing 0.06	$1.895 \leqslant y < 2.905$ and $0.06 \leqslant \lvert x\rvert \leqslant 5.3-y$
Spacing 0.04	$0.705 \leqslant y < 1.895$ and $0.04 \leqslant \lvert x\rvert \leqslant 5.3-y$
Spacing 0.02	$0.007 \leqslant y < 0.705$ and $0.02 \leqslant \lvert x\rvert \leqslant 5.3-y$
Spacing 0.01	$2 \times 10^{-5} \leqslant y < 0.007$ and $0.01 \leqslant \lvert x\rvert \leqslant 5.4$
Seventh order	$y \leqslant 2 \times 10^{-5}$ and $5.75 \leqslant \lvert x\rvert \leqslant 6.1$
Gauss–Hermite Quadrature	$2 \times 10^{-5} \leqslant y < 0.007$ and $5.4 \leqslant \lvert x\rvert \leqslant 6.1$
	$0.007 \leqslant y < 0.67$ and $5.3-y \leqslant \lvert x\rvert \leqslant 6.1$
	$0.67 \leqslant y < 3.75$ and $5.3-y \leqslant \lvert x\rvert \leqslant 6.77-y$
	$3.75 \leqslant y < 5.5$ and $\lvert x\rvert \leqslant 6.77-y$
Fifth order	$y < 0.67$ and $6.1 \leqslant \lvert x\rvert \leqslant 13.3$
Gauss–Hermite Quadrature	$0.67 \leqslant y < 1.7$ and $6.77-y \leqslant \lvert x\rvert \leqslant 13.3$
	$1.7 \leqslant y < 5.5$ and $6.77-y \leqslant \lvert x\rvert \leqslant 14.76-y$
	$5.5 \leqslant y < 13.1$ and $\lvert x\rvert \leqslant 14.76-y$
Third order	$y < 1.7$ and $13.3 \leqslant \lvert x\rvert \leqslant 1235–0.001111y^2$
Gauss–Hermite Quadrature	$1.7 \leqslant y < 13.1$ and $14.76-y \leqslant \lvert x\rvert \leqslant 1235–.001111y^2$
	$13.1 \leqslant y < 410$ and $\lvert x\rvert \leqslant 1235–.001111y^2$
	$410 \leqslant y < 570$ and $x \leqslant 1400–.0021y^2$
	$570 \leqslant y < 1230$ and $\lvert x\rvert \leqslant 800–.00035y^2$
First order	$y < 410$ and $\lvert x\rvert \geqslant 1235–.001111y^2$
Gauss–Hermite Quadrature	$410 \leqslant y < 570$ and $\lvert x\rvert \geqslant 1400–0.0021y^2$
	$570 \leqslant y < 1230$, $y \geqslant 1230$ and $\lvert x\rvert \geqslant 800–0.00035y^2$

$$\frac{\partial K(x,y)}{\partial y} = -\frac{\partial L(x,y)}{\partial x}, \tag{11}$$

$$\frac{\partial^2 K(x,y)}{\partial x^2} = -\frac{\partial^2 K(x,y)}{\partial y^2}. \tag{12}$$

By repeated application of these three relations, it can be shown that for a third-order Taylor series expansion, only the following need be tabulated: $K(x,y)$, $L(x,y)$, $\partial K(x,y)/\partial x$, $\partial K(x,y)/\partial y$, $\partial^2 K(x,y)/\partial x^2$, $\partial^2 K(x,y)/\partial x \partial y$, $\partial^3 K(x,y)/\partial x^3$ and $\partial^3 K(x,y)/\partial y^3$. To third order the interpolation for $K(x,y)$ finally takes the following form:

$$K(x,y) \cong K(x_0,y_0) + \Delta x \frac{\partial K(x_0,y_0)}{\partial x} + \Delta y \frac{\partial K(x_0,y_0)}{\partial y} + \Delta x \Delta y \frac{\partial K^2(x_0,y_0)}{\partial x \partial y} + \frac{1}{2}\frac{\partial^2 K(x_0,y_0)}{\partial x^2}(\Delta x^2 - \Delta y^2) + \frac{\Delta x}{6}\frac{\partial K^3(x_0,y_0)}{\partial x^3}(\Delta x^2 - 3\Delta y^2) - \frac{\Delta y}{6}\frac{\partial K^3(x_0,y_0)}{\partial y^3}(\Delta y^2 - 3\Delta x^2). \tag{13}$$

Here x_0 and y_0 are the x and y values of the precomputed point and $\Delta x = x - x_0$ and $\Delta y = y - y_0$. The third-order interpolation for $L(x,y)$ takes the following form:

$$L(x,y) \cong L(x_0,y_0) - \Delta x \frac{\partial K(x_0,y_0)}{\partial y} + \Delta y \frac{\partial K(x_0,y_0)}{\partial x} + \Delta x \Delta y \frac{\partial^2 K(x_0,y_0)}{\partial x^2} + \frac{1}{2}\frac{\partial^2 K(x_0,y_0)}{\partial x \partial y}(\Delta y^2 - \Delta x^2) + \frac{\Delta x}{6}\frac{\partial^3 K(x_0,y_0)}{\partial y^3}(\Delta x^2 - 3\Delta y^2) - \frac{\Delta y}{6}\frac{\partial^3 K(x_0,y_0)}{\partial x^3}(\Delta y^2 - 3\Delta x^2). \tag{14}$$

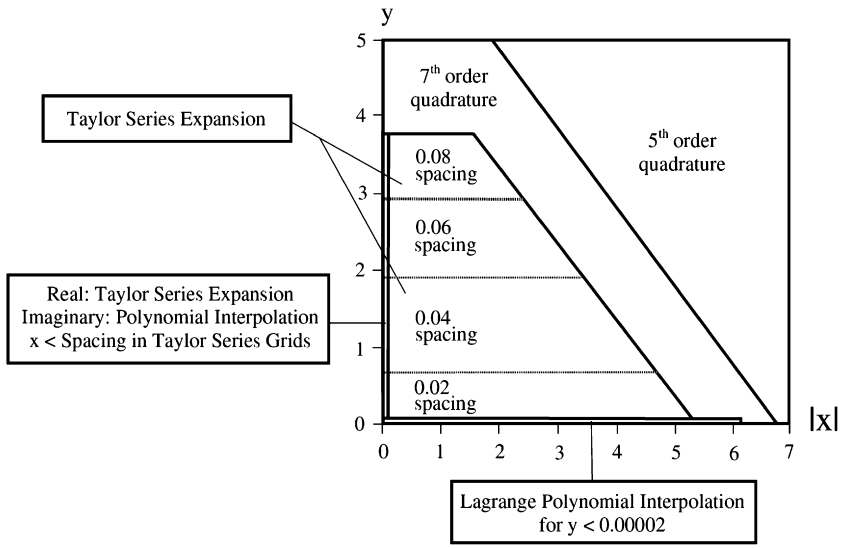

Fig. 3. Approximation regions for the real and imaginary parts of the Voigt function for $y<5$ and $|x|<7$.

The spacing between the tabulated values of x_0 and y_0 is varied as necessary to meet the accuracy requirements. Fig. 3 indicates the regions of the x, y plane in which this approximation is used for our algorithm and the spacing between tabulated points in different regions. The required memory of the precomputed tables for the Taylor series expansion method is quite minimal, about 890 kilobytes.

3.3. Lagrange interpolating polynomials

In a few cases where either x is very small and y is less than 3.75 or y is very small and x is less than 5.75, the above approximations fail to accurately evaluate the function. Many of the popular Voigt routines fail badly in these same regions, so an approximation must be carefully chosen. Since these cases are uncommon, speed of computation is not as high a priority and Lagrange interpolation [chapter 25 of Ref. [15]] in a precomputed table of values is used. This table requires only 597 kilobytes of RAM. The general form is a sum over a series of ratios of polynomials.

$$p(v) \cong \sum_{j=0}^{n} p(v_j) \prod_{\substack{k=0 \\ k\neq j}}^{n} \frac{v - v_k}{v_j - v_k}. \tag{15}$$

Here v is the value at which the function $P(v)$ is to be determined and n is the order of interpolation. The subscripts refer to grid points at which the function is precalculated. The grid points are defined equally spaced in x and y, resulting in a considerable simplification of Eq. (15). A second-degree polynomial was chosen and was simplified to the form of Eq. (16).

$$p(v) \cong p(v_1) + \frac{1}{2}\Delta v[p(v_2) - p(v_0)] + \frac{1}{2}\Delta v^2[p(v_2) + p(v_0) - 2p(v_1)]. \tag{16}$$

Here v_0, v_1, and v_2 are grid points in the table such that $v_0<v_1<v_2$, Δv is $(v-v_1)/(v_2-v_1)$ and when Lagrange interpolating polynomials in y are used near the Doppler limit (for $y<2\times10^{-5}$), y_1 is always equal to 1×10^{-5}, y_0 is zero and y_2 is 2×10^{-5}. Otherwise, v_1 is the tabulated point closest to v. This approximation is used for interpolation in both x and y, meaning that four polynomials of the form of Eq. (16) are required to evaluate each of the real and imaginary parts.

4. Evaluation of the real and imaginary parts of the Voigt function

4.1. Regions of approximation

The regions in which the various approximations are used to compute the real and imaginary parts of the Voigt function are presented in Figs. 2 and 3, as well as in more quantitative form in Table 1. Fig. 2 displays the region for each approximation on a log–log plot, showing the overall trends; as x and y become small, higher orders of quadrature, and eventually Taylor series expansion become necessary. No interpolation by Lagrange polynomials is present on this graph, as the regions are too small to appear on this scale. In Fig. 3, the region close to the x–y origin is shown in greater detail, making it possible to view the Lagrange interpolation regions. The Lagrange interpolation regions are not even to scale on this graph; they should be even smaller than they actually appear. At this scale it is also possible to view the different precomputed grid point spacings used in Taylor series expansion; again, the spacing becomes more dense as y becomes small. Many previous functions use mathematical techniques, which fail for extremely low values of y [14], where the function is approaching the Doppler limit. The Lagrange polynomial interpolation technique has an advantage for small y values because it converges smoothly to the Doppler profile. Also, despite the number of approximation regions, the process of calculating which approximation to use has been streamlined so that each spectral line is evaluated in as few different approximations as possible.

4.2. Accuracy

Accuracy for all routines was determined relative to a more accurate standard. The standard used most commonly was CADRE, a numerical integration program whose relative accuracy was set to 10^{-12} [20]. Calculations with the commercial program Maple^{TM} [21] with sixteen-digit precision were used to verify the accuracy of CADRE. It was determined that for values of y less than 10^{-5} CADRE became unreliable, so Maple^{TM} was used in this region. The relative error $|\epsilon|$ displayed in Figs. 4 and 5 is the difference between the approximation and the standard divided by the standard for the real and imaginary parts of the Voigt function calculated using the routine presented in this paper. The absolute value of the error is always less than 10^{-6} of the value of the function. Though these graphs only present data for $0 \leqslant |x| \leqslant 10$ and $0 \leqslant y \leqslant 10$, the error is the same or better in the remainder of the x–y plane. Similar graphs for many other common Voigt routines are displayed by Schreier [5] and are useful for comparison to the algorithm presented in this paper. These graphs show that the Humlicek algorithm has relative error for both the real and imaginary parts approximately two orders of magnitude larger than the routine presented in this paper. Schreier [5] also shows that for the real part of the Voigt function, the Drayson [10] routine has a maximum error a few times worse than the Humlíček [11] routine and the Pierluissi [13] routine is about two orders of magnitude worse. Both of these routines perform more poorly than the Humlíček [11] routine near the Doppler limit (low values of y).

4.3. Speed

In addition to being orders of magnitude more accurate than many common Voigt routines such as Refs. [10,11,13], our approximation is also significantly faster. Table 2 displays the ratio of the time of calculation for Refs. [10,11] and the simple Lorentz profile in Eq. (5) divided by the time of calculation for the algorithm presented in this paper. The first column of ratios corresponds to a typical calculation for a strong, broad, Lorentz line. Similarly, the last column corresponds to a typical spectral line near the Doppler limit where the Lorentz approximation is worst. For this comparison the Drayson [10] and Humlíček [11] functions are modified such that they are only called once per spectral line and precalculate all possible quantities involving y. This is an improvement in speed of up to a factor of two over the functions as published. For each approximation 10^9 points, evenly spaced throughout the regions indicated, with equal numbers of different x- and y-values were computed. Our function runs two to three times faster than Humlíček in all regions, as well as up to a factor of nine times faster than Drayson. Compared to the routines as originally published, the tabulated timing ratios are up to a factor of two larger. In regions where the Lorentz profile is close to the Voigt profile, our function is only a factor of two slower, but more accurate, than the Lorentz profile.

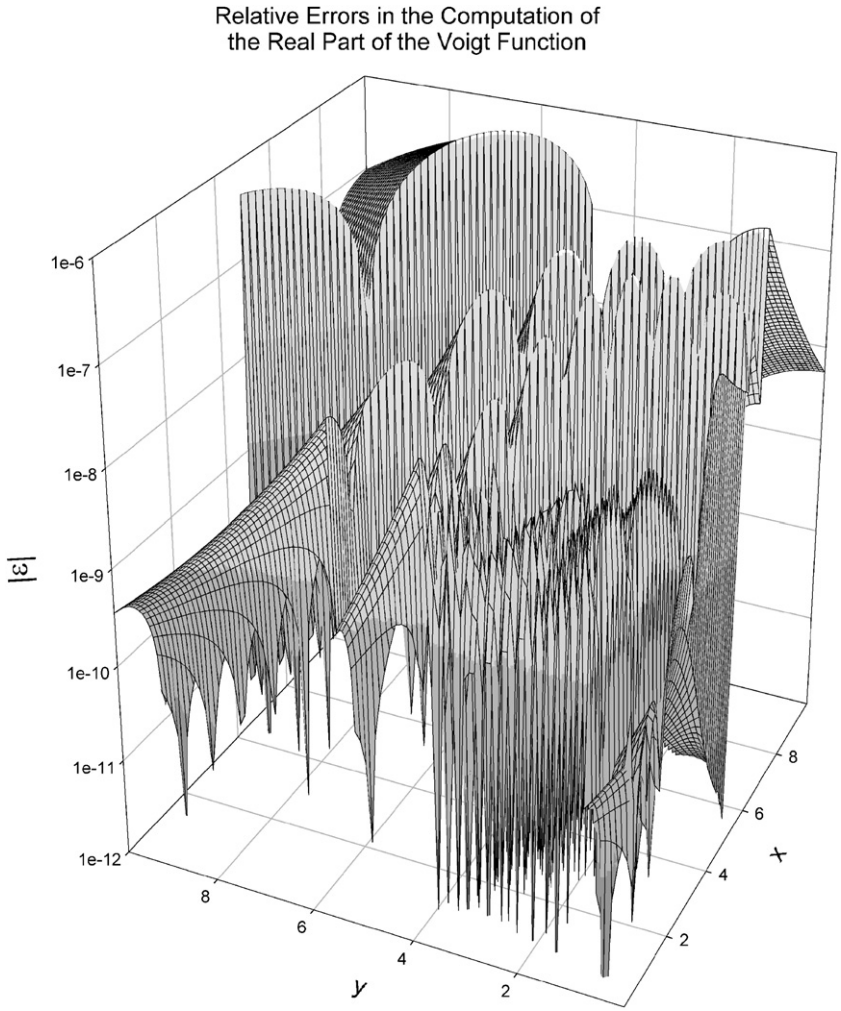

Fig. 4. Relative error $|\epsilon|$ in the calculation of the real part of the Voigt function for $0 \leqslant |x| \leqslant 10$ and $0 \leqslant y \leqslant 10$ for the algorithm presented in this paper.

Situations in which any approximation is faster than the Lorentz function is extremely rare. This means that order of magnitude gains in speed are not to be expected in the future and explains why it has been so hard to write faster routines than those written decades ago. Even the fast Fourier transform techniques (FFT) (see Ref. [5] for a review of these) must do a calculation at least comparable in difficulty to the Lorentz profile before doing the transforms and multiplications of the Doppler and Lorentz sections. This means that under no circumstances can the FFT method go faster than the Lorentz evaluation; for almost all cases they will be far slower once the FFT time is averaged over all of the evaluations of the Voigt function. For small values of x and y where the Lorentz profile is such a poor approximation that it is almost never acceptable, our approximation is less than a factor of five slower than the calculation of the Lorentz profile. The Drayson [10] routine and the simple Lorentz profile calculation do not calculate the imaginary part of the Voigt function and cannot be used for applications such as line mixing and the Rautian profile. For small y values where line mixing is less important so that only the real part is likely to be needed, Drayson's [10] algorithm is almost an order of magnitude slower than our algorithm. Where Drayson's algorithm does best compared to ours is where the imaginary part of the Voigt function is most likely to be required.

Table 3 displays a comparison of speed between our approximation and other common Voigt approximations, all implemented to be called once per (x, y) pair, using the same testing procedure as in Table 2. Our function is not optimized for this method, but it is still as fast or even slightly faster than the

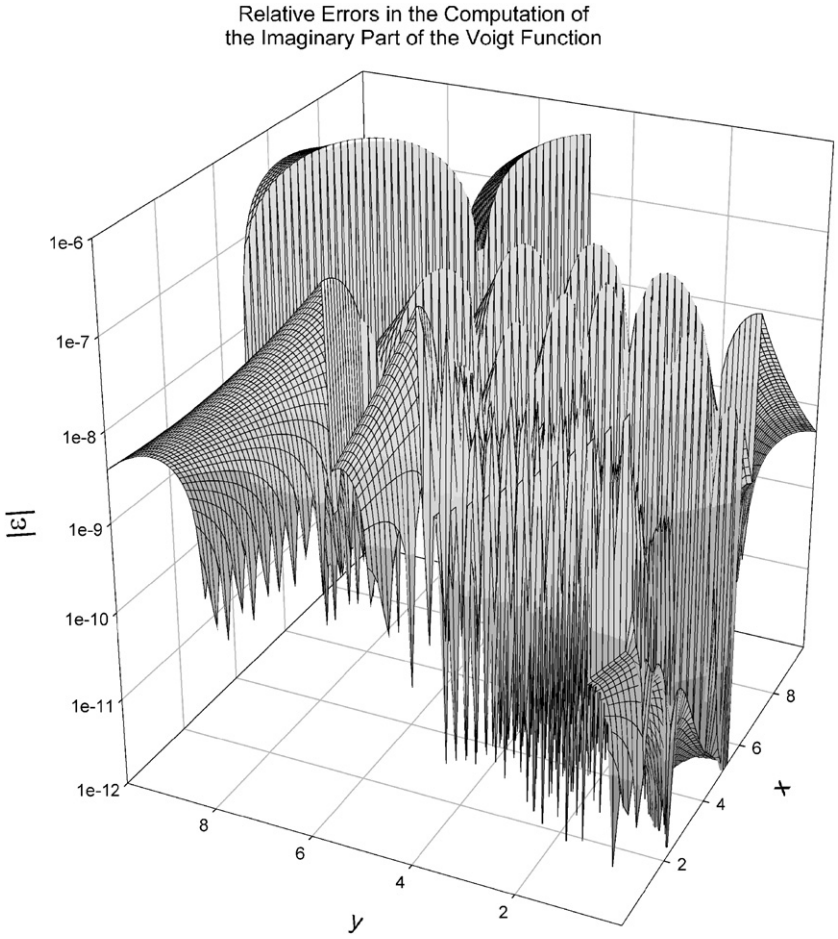

Fig. 5. Relative error $|\epsilon|$ in the calculation of the imaginary part of the Voigt function for $0 \leqslant |x| \leqslant 10$ and $0 \leqslant y \leqslant 10$ for the algorithm presented in this paper.

Table 2
Ratio of calculation time for the Voigt function to that for the algorithm in this paper for some published and well known approximations

	$0 \leqslant x \leqslant 1000,\ 0 \leqslant y \leqslant 1000$	$0 \leqslant x \leqslant 10,\ 0 \leqslant y \leqslant 10$	$0 \leqslant x \leqslant 5,\ 0 \leqslant y \leqslant 1$
Humlíček [11]	2.2	2.11	3.29
Drayson [10][a]	1.07	1.92	8.71
Lorentz profile[a]	0.52	0.29	0.21

For this comparison one subroutine call is performed per spectral line for all algorithms.

[a]This comparison is for the time taken for calculation of both the real and imaginary parts of the Voigt function for the algorithm in this paper, but only for the real part for the other algorithm.

Drayson [10] and Humlíček [11] routines. Though the Pierluissi [13] routine (as modified by Ref. [6]) is slightly faster than the present routine for many cases, it is also very inaccurate, with a relative error of a few percent for values of y only slightly less than 1. For a more comprehensive comparison of the accuracy and speed of other common Voigt subroutines, see Ref. [5].

The breakdown of time spent on calculation for each Voigt function evaluation in each approximation region of the present Voigt routine is shown in Table 4, and the time spent in each approximation region of the Humlíček [11] algorithm is shown in Table 5. These times are expressed as a ratio of the computation time

Table 3
Ratio of calculation time for the Voigt function using published approximations to that for the approximation in this paper when the subroutine call is performed for each (x,y) pair

	$0\leqslant x\leqslant 1000,\ 0\leqslant y\leqslant 1000$	$0\leqslant x\leqslant 10,\ 0\leqslant y\leqslant 10$	$0\leqslant x\leqslant 5,\ 0\leqslant y\leqslant 1$
Humlíček [11]	1.93	1.17	1.57
Drayson [10][a]	1.03	1.00	3.04
Pierluissi [13][a]	0.83	0.75	1.90

[a]This comparison is for the time taken for calculation of both the real and imaginary parts of the Voigt function for the algorithm in this paper, but only for the real part for the other algorithm.

Table 4
Ratios of calculation time for the present Voigt function for the various approximation regions pictured in Figs. 2 and 3 to that for the first order Gauss–Hermite quadrature region

Taylor expansion (real) and Lagrange interpolation (imaginary)	8.00
Lagrange interpolating polynomials	6.71
Taylor series expansion	4.35
Seventh-order Gauss–Hermite Quadrature	4.06
Fifth-order Gauss–Hermite Quadrature	2.88
Third-order Gauss–Hermite Quadrature	1.88
First-order Gauss–Hermite Quadrature	1.00

Table 5
Ratios of calculation time for the Humlíček [11] Voigt function for the various approximation regions to that of the first order Gauss–Hermite approximation in this work

Seven term expansion	30.1
Five term expansion	6.71
Four term expansion	4.59
Two term expansion	3.82

divided by that for the first order Gauss–Hermite quadrature in Table 4. For both routines the time increases steadily with the degree of approximation used. The fastest average time per evaluation for the Humlíček [11] approximation is about a factor of four slower than the fastest approximation in our approximation. Also, the slowest calculation using our function is about four times faster than the slowest calculation of Humlíček. The slowest region in the present algorithm, involving Lagrange interpolation, is required much less frequently than Humlíček's seven-term expansion, which is invoked for any evaluation with y less than about 0.75. All of our approximation regions other than the rare Lagrange interpolation are faster than at least three of the four Humlíček regions. This table is useful in conjunction with Table 1 and the equivalent information from Humlíček [11] to compare speeds at any combination of points in the x, y plane.

5. Derivatives of the Voigt function

Ref. [5] presents equations for the derivatives of the real part of the Voigt function in terms of the real and imaginary parts of itself.

$$\frac{\partial K(x,y)}{\partial x} = 2[yL(x,y) + xK(x,y)], \tag{17}$$

$$\frac{\partial K(x,y)}{\partial y} = 2\left[yK(x,y) + xL(x,y) - \frac{1}{\sqrt{\pi}}\right]. \tag{18}$$

Eqs. (17) and (18) are mathematically exact, however when implemented on a computer, round off error becomes a large problem in some cases. When $yL(x,y) \approx xK(x,y)$ in Eq. (17) and when $yK(x,y)+xL(x,y) \approx 1/\sqrt{\pi}$ in Eq. (18), the subtraction of two nearly equal terms causes a large decrease in relative accuracy. In the regions where Gauss–Hermite quadrature is employed, the problem could be solved by manipulating the quadrature equations. When the expressions for $K(x,y)$ and $L(x,y)$ in Eqs. (8) and (9) are substituted into Eq. (17), several leading terms in the quadrature equations cancel each other, reducing the roundoff error, and yielding Eq. (19) for nth-order quadrature.

$$\frac{\partial K(x,y)}{\partial x} \cong -\frac{4}{\pi}\sum_{i=1}^{(n-1)/2}\frac{w_i x y t_i^2}{(y^2+x^2+t_i^2)^2-4x^2t_i^2}. \tag{19}$$

When the quadrature expressions for $K(x,y)$ and $L(x,y)$ are substituted into Eq. (18), some of the terms in the quadrature equations add together to equal $1/\sqrt{\pi}$ and cancel the $1/\sqrt{\pi}$ term in Eq. (18), producing Eq. (20) for nth-order quadrature.

$$\frac{\partial K(x,y)}{\partial y} = -\frac{2}{\pi}\sum_{i=1}^{(n-1)/2} w_i t_i^2 \frac{y^2+t_i^2-x^2}{(y^2+x^2+t_i^2)^2-4x^2t_i^2}. \tag{20}$$

The cancellation of the largest terms render the results of Eqs. (19) and (20) more accurate than Eqs. (17) and (18) for the Gauss–Hermite quadrature regions. However, in the region where first-order quadrature is employed Eqs. (19) and (20) are not sufficient to describe the derivatives because the numerical result of substituting the first-order approximations for $K(x,y)$ and $L(x,y)$ into Eqs. (17) and (18) produces the value of zero for both derivatives. In this first-order region, $K(x, y)$ is approximated by $K_1(x, y)$ as defined in Eq. (21).

$$K_1(x,y) = \frac{1}{\sqrt{\pi}}\frac{y}{x^2+y^2}. \tag{21}$$

$L(x,y)$ is approximated by $L_1(x,y)$ as defined in Eq. (22).

$$L_1(x,y) = \frac{1}{\sqrt{\pi}}\frac{x}{x^2+y^2}. \tag{22}$$

To find $\mathrm{d}K(x,y)/\mathrm{d}x$ and $\mathrm{d}K(x,y)/\mathrm{d}y$, the partial derivatives of Eq. (21) are taken with respect to x and y. The expressions simplify to Eqs. (23) and (24).

$$\frac{\partial K(x,y)}{\partial x} \cong -\frac{2}{\sqrt{\pi}}K_1(x,y)L_1(x,y), \tag{23}$$

$$\frac{\partial K(x,y)}{\partial y} \cong -\frac{1}{\sqrt{\pi}}[K_1(x,y)-L_1(x,y)][K_1(x,y)+L_1(x,y)]. \tag{24}$$

For the first-order quadrature region where $K_1(x,y) \approx L_1(x,y)$, this method still suffers round-off error. For $0.9y<x<1.1y$ a method was employed which simplifies the two term Humlíček expansion by canceling the leading terms was substituted when $|x|+y>15$. The equation in complex arithmetic for the two-term Humlíček [11] expansion (shown in Eq. (25)) is separated into the real and imaginary parts.

$$K(x,y)+\mathrm{i}L(x,y) \cong \frac{1}{\sqrt{\pi}}\frac{y-\mathrm{i}x}{0.5+(y-\mathrm{i}x)^2}. \tag{25}$$

The real and imaginary parts are then substituted for $K(x,y)$ and $L(x,y)$ into Eqs. (17) and (18) to yield Eqs. (26) and (27).

$$\frac{\partial K(x,y)}{\partial x} \cong -\frac{1}{\sqrt{\pi}}\frac{2xy}{(x^2+y^2)^2+y^2-x^2+0.25}, \tag{26}$$

$$\frac{\partial K(x,y)}{\partial y} \cong \frac{1}{\sqrt{\pi}}\frac{x^2-y^2-0.5}{(x^2+y^2)^2+y^2-x^2+0.25}. \tag{27}$$

Eqs. (17) and (18) are employed to calculate derivatives wherever Lagrange interpolating polynomials or Taylor series expansions are used to determine the Voigt function. Fig. 6 displays where each approximation is used.

The absolute value of the relative error, $|\epsilon|$, of these derivative calculations remains less than 0.5% in all regions where $|x|+y>15$. CADRE and Maple™ were again used as the standard to which our routine was compared. The relaxed error requirement of 0.5% was chosen because calculation of the derivatives to this accuracy is a peripheral task, which can be achieved along with calculation of the Voigt function, adding little additional calculation time. In addition, since the calculation of the derivatives is generally used for sensitivity studies, error analysis or in nonlinear fitting of a spectrum, a high relative accuracy is usually not required. For $|x|+y<15$, at values where $yK(x,y)+xL(x,y)\approx 1/\sqrt{\pi}$, $\partial K(x,y)/\partial y$ approaches zero. The relative accuracy of 0.5% is not achieved when very close to this limit, however an absolute error of 10^{-7} is always met at points where the relative error in $\partial K(x,y)/\partial y$ is greater than 0.5%. A maximum absolute accuracy criterion for these points is useful because the values of $\partial K/\partial y$ at the nearby values of x are much larger, so the derivatives are essentially zero in comparison. Error analyses and fitting routines are thus insensitive to these points and it is only essential that the derivatives be small in comparison to those of the neighboring points in x. The relative error in $\partial K(x,y)/\partial x$ always meets the 0.5% criterion.

The addition of the derivatives to the calculation increases the computation time of the Voigt subroutine by 70% at most, the largest impact being where the function itself is calculated most rapidly. Versions of the subroutine with and without the derivative calculation have been written.

These derivatives are useful for a number of applications. When simulating a spectrum, it is often useful or necessary to find derivatives to either fit or find the sensitivity of the spectrum to various physical parameters (such as temperature or pressure) or spectral line parameters (such as intensity or Lorentz half-width). The required derivatives generally include a partial derivative of the form $\partial K(x,y)/\partial q$, where q is

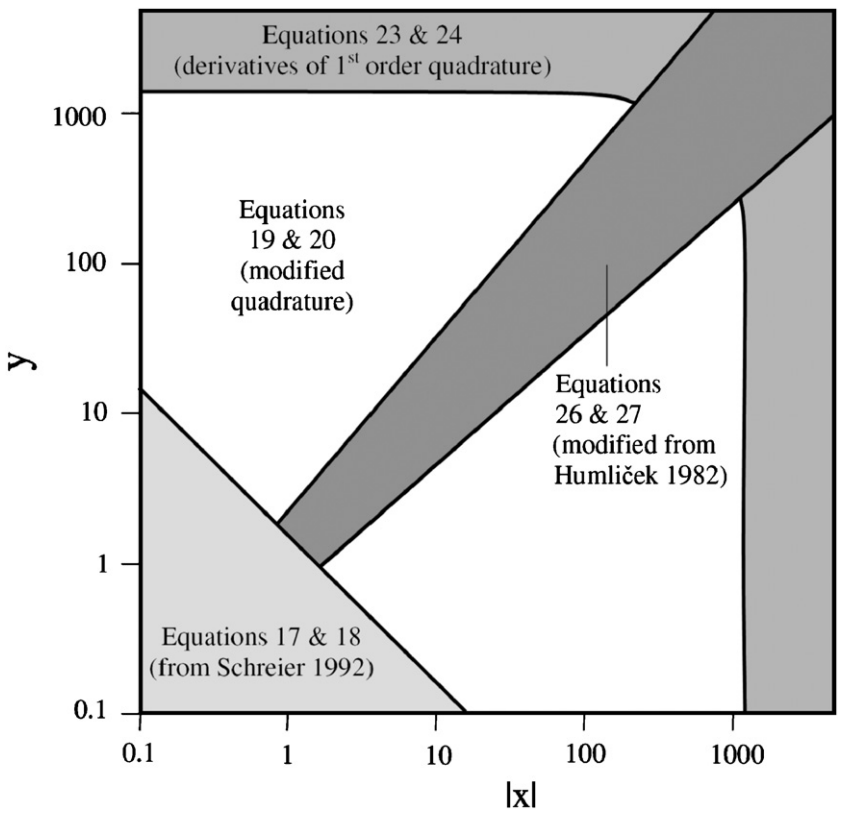

Fig. 6. Approximation regions for the derivatives of the real part of the Voigt function with respect to the parameters x and y.

any one of these parameters in the simulation. This partial derivative may be found by means of the chain rule [22].

6. Truncation of calculation

Calculation of spectral line profiles is a computationally intense, time-consuming portion of a line-by-line spectral simulation. The previous sections have shown how each evaluation of the Voigt function can be made more rapid. A second, completely independent approach is to try to minimize the number of evaluations of the Voigt function required. This minimization can be done using any Voigt profile approximation, not just that presented in this work. Spectral lines whose line centers contribute absorption below some cutoff do not contribute appreciably to the final results. Applying such a spectral line intensity cutoff can, in some cases, reduce the computation time by large factors when the number of spectral lines is rapidly increasing as the intensity of the lines is decreasing. The required cutoffs are not easy to calculate since there is no simple analytic expression for the absorption at the center of a Voigt profile and the required cutoffs are a function of numerous experimental and spectroscopic parameters.

Another means of reducing time for the computation of spectral line profiles is to truncate the calculation of the profile at some distance from the line center and simply assume it to be zero for all points farther from the line center. Far from line center the absorption becomes weak and, eventually, negligible. Simple criteria for how far from line center it is necessary to calculate a Voigt line would be useful. Criteria other than those listed here may become important if the shape of the spectral line far from line center is not Voigt. Such complications are not considered here.

6.1. The simplest techniques

The calculation of the extent of the spectral line profile to be calculated must not take a significant amount of time compared to the evaluation of the Voigt function itself. The easiest technique is to calculate all spectral lines to the same distance in wavenumber from line center. The appropriate maximum distance from line center is that required by the spectral line in the list of lines, which has the strongest wings at the physical conditions of the gas. While this criterion is straightforward and the specification of the distance from line center is rapid, it is clear that if there are spectral lines ranging over orders of magnitude in intensity, many spectral lines are being calculated much farther from line center than necessary. Also, there may be a large range among the half-widths of the various spectral lines that make it unnecessary to calculate the narrower lines as far from line center as the broader ones. Furthermore, one often calculates several spectra and the appropriate criterion may be different under different physical conditions.

A somewhat better criterion is to calculate all spectral lines to the same number of half-widths from line center. Generally, the Lorentz half-width is the basis for this criterion since it dominates the wings of a spectral line in almost all cases. This simple choice solves the problem of spectral lines of different widths, but not the problem of spectral lines of different intensities. In most cases it is the spectral line intensities that vary the most.

These techniques are more efficient than the calculation of all spectral lines throughout the entire calculated spectrum, but they leave a great deal of room for improvement. The following sections will assume that the spectrum calculated is an absorption spectrum, but an extension to emission spectra is trivial; substituting a minimum emission for the minimum absorption in the final equations is all that is required.

6.2. Absorption cutoff

It is clear that a criterion which would truncate each spectral absorption line when its absorption, A, decreases to a specific value, A_m, would be an improvement over these simple techniques. The appropriate value of A_m will be discussed in Section 6.3, but it is sufficient here to say that it must be small enough that the excluded calculations must never get close to the noise level. The problem is that every spectral point calculated must be tested and the test itself requires a major fraction of the calculation time for the Voigt function. A more general approach to this technique is required. First, consider pure Lorentz and pure Doppler lines. After that the Voigt profile may be considered more easily.

6.2.1. Lorentz profiles

For the Lorentz line profile the transmission, $T(v)$, is the limit where the thermal motions of the molecules do not affect the profile. In practice, this line profile dominates above some gas pressure and beyond some point in the wings of the line. Eq. (28) uses the definition of Eq. (6) for the Lorentz limit of the real part of the Voigt function to form the Lorentz profile, but defines it in terms of more spectroscopic quantities.

$$T(v) = 1 - A = \exp\left\{-\frac{Su\alpha_L}{\pi[(v - v_0)^2 + \alpha_L^2}\right\}. \tag{28}$$

Here S is the spectral line intensity, u is the mass path of the gas. Taking the natural logarithm of both sides of Eq. (28):

$$\ln(1 - A) = -\frac{Su\alpha_L}{\pi[(v - v_0)^2 + \alpha_L^2]} \cong -A. \tag{29}$$

For cases where $A \ll 1$, the error in the approximation on the far right side of Eq. (29) is less than $\frac{1}{2}A^2$.

6.2.1.1. Criterion for complete exclusion of a Lorentz spectral line. Any spectral line whose line center has less absorption than A_m can be excluded from the calculation altogether. The choice of value for A_m will be discussed in Section 6.3, but generally it is smaller than the inverse of the signal to noise ratio of the experimental spectrum. In general, A_m is very small compared to one, so that the approximation on the far right side of Eq. (29) applies very well. $v = v_0$ applies at the center of the spectral line, so substituting this into Eq. (29) and solving for S yields the minimum intensity of a spectral line, S_m, that has an absorption of at least A_m at its center.

$$S_m \cong \frac{\pi\alpha_L A_m}{u}. \tag{30}$$

A spectral line with less than this intensity can be excluded from the calculation. For a given gas in a given spectrum, this cutoff intensity is just a constant times the Lorentz width of the spectral line.

6.2.1.2. Truncation distance from line center. For spectral lines with intensities greater than S_m, the distance from the center of the line corresponding to an absorption of A_m serves as the truncation point in the calculation of the spectral line wings. Since the Lorentz profile is symmetric about the line center, the distance is the same on both sides of the spectral line. Once again, A_m is assumed to be much smaller than one, so the right side of Eq. (29) is accurate enough for these purposes. The distance the profile must be calculated from line center is found by substituting A_m for A in Eq. (29) and solving for $v - v_0$.

$$v - v_0 \cong \pm\sqrt{\alpha_L\left(\frac{Su}{\pi A_m} - \alpha_L\right)}. \tag{31}$$

This value is zero for $S = S_m$ and imaginary for $S < S_m$, so the line must be tested for whether it can be completely excluded first. This expression is more complicated than that in Eq. (30), but since it only needs to be calculated once for each spectral line in each spectrum and need not be calculated at all for spectral lines with $S < S_m$, the required computation time is small compared to that for the Voigt evaluations.

Eq. (31) also provides another means of excluding lines altogether. If the spectral line center is outside the simulated spectral interval by more than the distance indicated in Eq. (31), the line can be excluded.

6.2.2. Doppler profiles

At low pressure, the Doppler profile may dominate. Rewriting Eq. (5) in terms of spectroscopic parameters, the transmission, $T(v)$, takes the form of Eq. (32).

$$T(v) = 1 - A = \exp\left\{-\frac{Su}{\alpha_D\sqrt{\pi}}\exp\left[-\left(\frac{v - v_0}{\alpha_D}\right)^2\right]\right\}. \tag{32}$$

Taking the natural logarithm of both sides and applying the same approximation as Eq. (29) for the left side yields a more useful form.

$$A \cong \frac{Su}{\alpha_D\sqrt{\pi}}\exp\left[-\left(\frac{v-v_0}{\alpha_D}\right)^2\right]. \tag{33}$$

For small values of A the error on the left side of this equation is less than $\frac{1}{2}A^2$ and for the purposes of determining the range over which to calculate the spectral line, this is a very good approximation.

6.2.2.1. Criterion for the complete exclusion of a Doppler spectral line. The central absorption of a weak Doppler spectral line ($A \ll 1$) is found by substituting zero for $v-v_0$ in Eq. (33). If A_m is substituted for A and the equation is solved for S, the result is the minimum intensity, S_m, of a line that needs to be included in the calculation.

$$S_m \cong \frac{\sqrt{\pi}\alpha_D A_m}{u}. \tag{34}$$

If a spectral line has less than this intensity, it need not be included in the computation at all. This is a simple test, which for a given spectrum, amounts only to a constant times the Doppler width of the spectral line.

6.2.2.2. Distance from line center required. The Doppler profile for a spectral line is symmetric about the line center, so a cutoff can be applied at the same distance from line center on both sides of the line. By also substituting A_m for A and solving for $v–v_0$ yields the distance from line center that corresponds to the cutoff absorption level.

$$v - v_0 \cong \alpha_D\sqrt{\ln\left(\frac{Su}{\sqrt{\pi}\alpha_D A_m}\right)}. \tag{35}$$

This expression is the most complicated yet, but it need only be calculated once for each spectral line for each spectrum. The expression yields imaginary values if $S<S_m$, but spectral lines that cause such a situation are eliminated by the test of Eq. (34). If a spectral line center is outside the spectral region being calculated and $v–v_0$ is less than the distance of the center of the spectral line from the nearest edge of the simulated spectrum, the line can be completely ignored.

6.2.3. Voigt line profiles

The Voigt line profile is the convolution of the Lorentz and Doppler line profiles. Consideration of the properties of each of these profiles allows one to determine criteria for the cutoff of Voigt spectral lines.

6.2.3.1. Criterion for the complete exclusion of a Voigt spectral line. The Lorentz and Doppler profiles are normalized and the line center of each represents the maximum absorption for that spectral line. The absorption at the center of a Voigt profile is always less than the absorption at the center of both the Doppler profile and the Lorentz profile from which it is constructed. The result is that if the spectral line intensity is less than either the criterion in Eq. (30) or the criterion in Eq. (34), then the Voigt profile can also be excluded from the line-by-line calculation. Mathematically, these two criteria may be combined into a single expression.

$$S_m \cong \frac{\sqrt{\pi}A_m}{u}\mathrm{Max}[\alpha_D, \sqrt{\pi}\alpha_L]. \tag{36}$$

Any line with intensity less than S_m need not be calculated since it will have negligible contribution to the overall spectrum. Lines with intensities just above S_m may also have an absorption less than A_m, but these rare cases will have a central absorption only slightly less than A_m and there is no harm in calculating them. These cases are most likely to occur when the Lorentz and Doppler widths are within a factor of a few.

At a given temperature, the Doppler width for spectral lines of a given isotopologue is proportional to the wavenumber of the center of the spectral line, so for a limited spectral interval the Doppler width is almost constant. Thus, a single value of the Doppler width is often sufficient for this test for a considerable spectral

region. Generally, one would use the wavenumber of the center of the spectral interval for the calculation of the Doppler width, but one could be conservative and use the low wavenumber end. This eliminates the need to calculate the Doppler width for spectral lines that are not to be included in the calculation.

Since the range of Lorentz widths is usually limited, one could eliminate the computation of the Lorentz widths of many lines by using a mean value, or even the smallest value. One approach is to use these approximation methods to compute a single cutoff intensity for the spectrum, then apply the test in detail for the spectral lines that are more intense than this cutoff.

6.2.3.2. Distance from line center required. The spectral lines that survive the intensity cutoff of Eq. (36) need to be evaluated in terms of how far from line center the Voigt profile calculation should be truncated. Generally, this truncation will occur in the far wings of the spectral line. Unlike the center of the line, the convolution in this portion of the spectral line profile strengthens the absorption. This means that one needs to calculate farther into the wings than would be calculated using either Eq. (31) or Eq. (35).

If the far wings in the region of the cutoff are dominated by the Doppler profile, the character of the wings is that the absorption is falling very rapidly. To use Eq. (35) will not lead to significant error since the points are falling so rapidly that at most only a point or two after the computed cutoff is likely to be above the cutoff absorption. Almost always the far wings are dominated by the Lorentz profile. In the far wings of the Lorentz profile the change in absorption is very small from one wavenumber sample to the next. The convolution of a nearly constant function with the Doppler function yields almost the original function. Any error in placing the cutoff will only allow points with absorptions very slightly above the cutoff criterion to slip by. Since the whole concept is an approximation technique, a slight change in the cutoff criterion will not make much difference. Thus, for lines that are truncated in the far wings, the use of the larger value from Eq. (31) or Eq. (35) will be sufficient.

The larger problem is with lines that barely make the cutoff criterion of Eq. (36). These lines are not truncated in the far wings and the Voigt convolution can change the shape of the spectral line considerably. If the spectral line is close to the Doppler profile limit, the Voigt profile does not differ much from the Doppler profile and Eq. (35) is quite accurate. If the spectral line is near the extreme of the Lorentz profile, the Voigt profile does not differ much from the Lorentz profile and Eq. (31) is quite accurate. Thus, using the larger distance from line center from Eqs. (31) and (35) always works unless the Lorentz width and Doppler widths are close to one another and the cutoff is not in the far wings. If the cutoff is near line center, the convolution will actually make the spectral line weaker in the region of the cutoff and the criteria of Eq. (31) and Eq. (35) will harmlessly go too far into the wings by a small amount.

A further simplification can be of benefit in the one remaining problem case. If the second term under the square root is dropped in Eq. (31), the computation is forced to go farther into the wings of the spectral line. The maximum distance to be calculated from line center in the Lorentz case (Eq. (31)) is then simplified.

$$\nu - \nu_0 \cong \sqrt{\frac{Su\alpha_L}{\pi A_m}}. \tag{37}$$

Near line center this simplification forces calculation further from line center by at most one Lorentz width. Since typical sampling is a few times per Voigt width, this means that the number of extra points calculated is small. As the truncation moves further into the wings of the spectral line, the number of extra evaluations quickly tends toward zero. For any of the problem cases, the Lorentz width will be close enough to the Doppler width that this will increase the number of points calculated sufficiently. Thus, the maximum of Eqs. (35) and (37) is sufficiently far from line center for the truncation of the wings of the spectral line and is the value to be used.

6.3. Conclusions

The decision where to truncate spectral line calculations reduces to two problems. The first is to eliminate any spectral line whose center does not appreciably contribute to the spectrum. This is done in steps. The spectral line intensity is compared to a critical intensity defined by Eq. (36) and minimum values of α_L and α_D. For those spectral lines that pass this test, the actual Doppler and Lorentz widths may be calculated and the

results substituted into Eq. (36). Most spectral lines to be excluded will be excluded by the first criterion. Almost all spectral lines remaining will pass the second test and the calculation of α_L and α_D are required anyway.

The spectral lines that pass the minimum intensity criterion must be tested for how far from line center they are to be truncated. The tests of Eqs. (35) and (37) are to be applied. The larger value represents the distance from line center to which the line needs to be calculated. The time required to calculate the truncation distance is negligible compared to the time required to do the Voigt computations. If the spectral line is outside the spectral region calculated, a test must be applied to decide whether the line is farther from the spectral region than the distance from line center to be calculated. If so, the line need not be calculated at all.

The remaining question has to do with the best value to use for A_m. This is a very difficult problem in general, but usually its value must change a lot to have much effect on the final results. If the goal is to simulate observed spectra, one must pick a value for which the discontinuities in the wings of the lines are small enough to be inconsequential. If many lines are very close to one another, the fact that several of these discontinuities may fall very close to each other must be considered. Thus, we find that laboratory spectra with a signal to noise on the order of 10^3 are almost always well fit with A_m of 10^{-4}. If only an average transmission over a spectral region many times wider than the half-widths of the lines is required or the spectrum will be convolved with an instrument function much wider than the half-widths of the spectral lines, A_m need only be small enough to include the required fraction of the area of the line on average. A general rule is that the value should be a few times smaller than the inverse of the signal to noise of the spectrum simulated or the inverse of the transmission accuracy required. In practice, though, the best means of determination of A_m is to simply do a calculation on one or more test cases with several values of A_m separated by about half an order of magnitude. As A_m is made smaller the calculation approaches the correct calculation asymptotically.

7. Conclusions

The complex Voigt function is used in the calculations of several popular spectral line shapes in spectroscopy. The most prominent of these is the Voigt profile, which is very close to the real part of this function. A large number of algorithms have been written in the past to calculate this function with various tradeoffs between speed and accuracy. In this paper a new algorithm has been described which maintains a relative accuracy of one part in a million while running a few times faster than some of the more popular algorithms which have accuracies at least two orders of magnitude worse. The cost compared to other algorithms is rather modest on modern computers, namely, 1.49 megabytes of RAM at most. This memory is required for the Taylor series approximation tables of the function and its derivatives and for the Lagrange interpolation tables.

At least as important for the speed of calculations using the Voigt profile is the ability to intelligently determine at what point in the wings of the lines it is no longer worthwhile to calculate the spectral line due to the smallness of the absorption or emission, or even when to ignore the calculation of the line altogether. This criterion can be applied independent of the Voigt approximation used. A calculation of where this cutoff should take place must not itself produce significant additional computation. A rapid, simple calculation is presented which characterizes this cutoff location well. Additionally, this technique is adjustable by means of a choice of the minimum absorption or emission to which calculation is necessary. This minimum will be smaller for higher signal to noise simulations and for simulations in which there are many overlapping spectral lines.

A copy of the Fortran 90 subroutines described here may be obtained from either of the authors.

Acknowledgments

Some of the computations in this paper were performed by using Maple™. Maple™ is a trademark of Waterloo Maple, Inc. This paper is based upon work supported by the National Science Foundation under Grant No. ATM-0338475.

References

[1] Thompson WJ. Numerous neat algorithms for the Voigt profile function. Comput Phys 1993;7:627–31.
[2] Pine AS, Looney JP. N_2 and air broadening in the fundamental bands of HF and HCl. J Mol Spectrosc 1987;122:41–55.
[3] Levy A, Lacome N, Chackerian C. Collisional line mixing. In: Rao KN, Weber A, editors. Spectroscopy of the earth's atmosphere and interstellar medium. Boston: Academic Press; 1992. p. 261–337.
[4] Benner DC, Rinsland CP, Devi VM, Smith MAH, Atkins D. A multispectrum nonlinear least squares fitting technique. JQSRT 1995;53:705–21.
[5] Schreier F. The Voigt and complex error function: a comparison of computational methods. JQSRT 1992;48:743–62.
[6] Twitty JT, Rarig PL, Thompson RE. A comparison of fast codes for the evaluation of the Voigt profile function. JQSRT 1980;24:529–32.
[7] Klim A. A comparison of methods for the calculation of Voigt profiles. JQSRT 1981;26:537–45.
[8] Brault JW, Smith WH. Determination of the H_2 4-0 S(1) quadrupole line strength and pressure shift. Astrophys J 1980;235:L177–8.
[9] Washenfelder RA, Toon GC, Blavier J-F, Yang Z, Allen NT, Wennberg PO, et al. Carbon dioxide column abundances at the Wisconsin Tall Tower site. J Geophys Res 2006;111:D22305.
[10] Drayson SR. Rapid calculation of the Voigt profile. JQSRT 1976;16:611–4.
[11] Humlíček J. Optimized computation of the Voigt and complex probability functions. JQSRT 1982;27:437–44.
[12] Gordley LL, Russell JM, Mickley LJ, Frederick JE, Park JH, Stone KA, et al. Validation of nitric oxide and nitrogen dioxide measurements made by HALOE for the UARS platform. J Geophys Res D 1996;101:10241–66.
[13] Pierluissi JH, Vanderwood PC, Gomez RB. Fast calculational algorithm for the Voigt profile. JQSRT 1977;18:555–8.
[14] Wells RJ. Rapid approximation to the Voigt/Faddeeva function and its derivatives. JQSRT 1999;62:29–48.
[15] Abramowitz M, Stegen IA. Handbook of mathematical functions. New York: National Bureau of Standards; 1964.
[16] Gentry B, Strow LL. Line mixing in a N_2-broadened CO_2 Q branch observed with a tunable diode laser. J Chem Phys 1987;86:5722–30.
[17] Armstrong BH. Spectrum line profiles: the Voigt function. JQSRT 1967;7:61–88.
[18] Drummond JR, Steckner M. Voigt-function evaluation using a two-dimensional interpolation scheme. JQSRT 1985;34:517–21.
[19] Norton RH, Rinsland CP. ATMOS data processing and science analysis methods. Appl Opt 1991;30:389–400.
[20] de Boor C. On writing an automatic integration algorithm. In: Rice JR., editor. Mathematical software New York: Academic Press; 1971 p. 201–9, 430–449.
[21] Maple 9. Waterloo, Ontario: Maplesoft, a Division of Waterloo Maple Inc., 2003.
[22] Stewart J. Multivariable calculus, 5th ed. Belmont: Brooks/Cole-Thomson Learning; 2003 [Chapter 15].