SATURATED ABSORPTION LINE SHAPE

J. L. Hall*
Joint Institute for Laboratory Astrophysics
University of Colorado, Boulder, Colorado

There are a number of relevant physical phenomena which will have to be taken into account in a proper theory of the saturated absorption line shape. These include: (1) collisional phase-shift effects [1] and velocity/frequency shifting small-angle collisions [2], (2) power broadening associated with the saturation process [3], (3) z-axis modulation of the saturation parameter [4], (4) the distribution of transverse velocities for a given axial velocity group [4], and, most importantly, (5) coherence-limiting effects due to the finite spatial extent of the light beam [2] and deviations from planar wavefronts [5].

Such a total proper theory does not yet exist to the author's knowledge. Indeed we know of no real saturation theory whatever that is applicable in the low-pressure regime of interest in the standards context (low pressure reduces the pressure-induced frequency offset). Thus we have found it useful to parameterize our observed resonances in terms of a very simple physical "hole-burning" model. It is found that the assumed Lorentzian line shape gives a good representation of the data. Within certain limits, the dependence of the linewidth upon pressure and laser intensity can be well represented by a three-parameter formula of the form suggested by a saturation model [2]. The fact that the three resonance parameters of the model

*Laboratory Astrophysics Division, National Bureau of Standards and Department of Physics and Astrophysics, University of Colorado.
take on systematically different values as one changes the laser spot size gives rise to a very satisfying and essentially quantitative corroboration of our understanding of the relevant physical scaling laws. On the other hand, our values of the phenomenological parameters have no transparent numerical relationship with the phase memory time and transition dipole moment, for example, of a particular molecular velocity group. Our data were obtained to teach us the scaling laws that might lead to more interesting optical-frequency-standard results. The parameterized results are presented here as an aid to communication between experimentalists and theoreticians, and to provide a quasiexperimental testing ground for their calculations.

In the low-pressure ($10^{-2}$ to $10^{-4}$ Torr) regime of interest in the frequency-standards context, the importance of a realistic treatment of collisional effects is greatly reduced. It is sufficient to use a single pressure-broadening parameter to represent collision-induced transitions out of the resonant velocity interval. Basically the resonance line has become so sharp that there are thousands of resolvable velocity intervals within the Doppler linewidth. Thus a negligible fraction of excited absorbers are returned to the resonant velocity groups by a second collision.

On the other hand, many molecules—perhaps a majority—which originally satisfy the resonance condition in $v_z$ will experience a coherent interaction with the radiation field during their entire transit of the laser beam spot. Thus it is absolutely essential to account for the finite duration of the excitation pulse as the transverse dimension of the spatially bounded laser beam is crossed by the absorbers at their (weighted) thermal velocity.

Since we are interested in the longest phase-coherent interaction time to minimize the resonance linewidth, we use carefully prepared single-spatial-mode excitation beams. The phase variation observed along a straight transverse molecular trajectory will then be sufficiently small if the wavefront radius of curvature is sufficiently great. A suitable phase criterion ($\omega \pm n/\lambda$ rad) gives

$$R > b = (2w_0^2/\lambda)^{1/2}$$

where $b$ is the equivalent confocal parameter of the laser beam characterized by Gaussian mode radius $w_0$. Thus a maximum absorption cell length $< 2b$ is implied for the configuration symmetric around the plane containing the beam waist. One then anticipates a "time-of-flight" linewidth $\Delta \nu = k(v <v>/w_0)$ where $<v>$ is a suitably weighted thermal average velocity and $k$ is a dimensionless constant of order unity.

We may estimate the value of $k$ as follows. For the Gaussian beam profile, about 40% of the energy lies inside an aperture of diameter $= w_0$. Two degrees of freedom contribute to the transverse velocity, giving
\[ \nu_{\text{rms}} = (2kT/M)^{1/2} = 5.5 \times 10^4 \text{ cm/sec}. \] The typical "pulse duration" is therefore \( \tau \approx \nu_0 / \nu_{\text{rms}} \) leading to a Lorentz half-width at half maximum, \( \Delta \nu_{\text{HWM}} = 1/2\pi \tau \), finally giving \( k = 1/2\pi \). Thus in this crude model one would expect \( \nu_0 \Delta \nu_{\text{HWM}} = \nu_{\text{rms}} / 2\pi = 88 \text{ kHz mm} \).

From the experimental results with single-mode excitation at room temperature (see text below and Table I) we find \( \nu_0 \Delta \nu_{\text{HWM}} = (70 \pm 5) \text{ kHz mm} \), corresponding to \( k = 1/8 \).

Now the Gaussian time-domain impulse is a self-apodized function, and so the line wings of the frequency-domain response function are free of the oscillations which are usually characteristic of pulse-interrogation resonance problems. Also the nonlinear character of the saturation resonance weights the slow velocity end of the velocity distribution. Thus it is almost surprising that the two values for \( k \) are in such close accord.

Naturally if the spatial quality of the beam is poor, a coherent interaction will not be obtained over the entire aperture and \( k \) can be more like 1/2 or unity. In fact if there are spatially distinguishable standing-wave fields, one can observe "Ramsey two-cavity fringes". It seems appropriate first to study the pure single-mode case.

The line shapes which we wish carefully to investigate can be as narrow as 23 kHz HWHM, a fractional width of \( 2.6 \times 10^{-10} \). It is obvious that mere mechanical stability will not suffice very long in such a domain! A very powerful kind of spectrometer can be built based on the concept of Frequency-Offset-Locking [6]. We acquire good frequency stability in a laser servo controlled to the apparent top of the saturated absorption peak in CH₄. This excellent stability is then transferred to the powerful laser which illuminates the absorption cell of interest. The reference laser need not be understandable in the same fundamental terms we hope to achieve in the external absorption case--it need only be (empirically) established to provide a frequency reference of good stability. In Fig. 1 we show the present version of this reference device. The high signal-to-noise ratio, 109 dB for a 1-sec average, should clearly provide some impressive frequency stability if optimally utilized. At present it is found that short-term frequency noise of the basic laser near the modulation frequency is degrading the performance severely [7]. To investigate this point, a nonoptimal device with about the same free-running frequency stability but a much smaller signal-to-noise ratio was heterodyned against one of our "standard CH₄ laser frequency references". [Actually both were beat against an offset laser (No. 2) to eliminate the sign ambiguity near zero beat.] In Fig. 2 we show the time-sequence beat frequency data for three averaging times. The 5-MHz basic value has been suppressed. It is clear from the figure that the fractional frequency fluctuation is still decreasing with averaging time near 1 sec. These data and many more have been processed to yield a plot of the Allan Variance as a function of averaging time. See Fig. 3. [The Allan Variance is a fractional measure of the (rms) average frequency difference between
adjacent samples of the input test frequency. It is widely used in the
frequency standards game because it is well behaved in the presence of noise
with long time correlations, such as 1/f noise.] In this case the test
frequency was the (offset) second beat (audio) between the two (5 MHz) laser
beat frequencies. The frequency noise of the common heterodyne laser No. 2
drops out of this second beat, leaving just the frequency noise of laser No. 1
plus that of laser No. 3.

From Fig. 3 it may be seen that the $\tau^{-1/2}$ averaging law expected for a
white-noise process is followed up to averaging times of about 100 sec,
where the demonstrated (in-)stability is a matter of 3 or 4 Hz! Ultimately
beyond about $10^4$ sec, the stability is degraded by changes in the same
systematic errors which limit the present reproducibility to about
$1/2 \times 10^{-11}$ Hz. In principle, the short-term frequency stability of the device

Figure 2. Time-domain frequency stability data. By taking the audio beat
between the beats of the two self-stabilized devices with a common local
oscillator, we observe directly the difference of the two stabilized optical
frequencies. The stability is improving with averaging time near 1 sec. The
averaging time in the lower photo was 45 sec.

Figure 1. Laser-saturated methane frequency reference. This version of
the device has a 1-mm mode waist in the absorption cell. The bore of the
gain cell is about 3.2 mm, with $P(\text{Ne}^{20}) = 0.3 \text{ Torr}$, $P(\text{He}^3) = 5.0 \text{ Torr}$.
The connections in the reservoir are for dc-fired Ba getter.
Figure 3. Allan Variance, calculated from data like that of Fig. 2. The frequency-lock loop operates for times shorter than 0.1 sec, the line center lock dominates at larger times.

shown in Fig. 1 should be about one decade better than the nonoptimal device tested in these experiments.

As a fundamental standard, the device of Fig. 1 still has important resatbility defects due to problems with definition of the base line. However as a frequency reference for laser spectroscopy [8], for geophysical strain seismometry [9], and for precision interferometry [10], the CH$_4$ frequency reference allows experiments of unprecedented sensitivity, stability, and precision.

For example, in the Frequency-Offset-Locked Laser Spectrometer, illustrated in Fig. 4, we are able to study the saturation resonance itself in delicious detail. The two frequency-offset-lock loops function very well indeed, transferring essentially the entire available stability of the reference laser (No.1) to the powerful (7 mW) controlled laser (No.3). The use of a local oscillator laser (No.2) offset 5 MHz red eliminates the troublesome
Figure 4. The Frequency-Offset-Locked Laser Spectrometer. Laser No. 1 servo stabilized to its own saturation peak provides excellent frequency stability which is transferred—-with a 5-MHz red frequency offset—to laser No. 2. Laser No. 3 is controlled relative to No. 2 by the synthesizer/frequency offset lock loop. The optical output power of No. 3 is also stabilized. The cavity containing the CH₄ absorber, servo controlled to track the input wavelength, gives good spatial-mode purity.

region near zero beat. Using the external cavity saturation signal as a frequency discriminator, we find the variations in the absolute frequency of laser No. 3 to be about 5 kHz peak to peak, with periods typically in the 10- to 100-msec range. For 1-sec averages, the rms frequency excursion is about 30 Hz. The fact that the frequency fluctuations can be measured in this way already proves that a more suitable reference device could provide better stability if it is needed.

The laser spectrometer employs an automatic laser intensity control to provide a frequency-independent base line for the absorption signals. At present we are interested in the effects of mode size and purity, making it
attractive to put the absorbing gas inside an external resonant cavity. The
minor bother of having to servo stabilize the cavity length to track the
applied optical input frequency/wavelength is compensated by the "contrast
amplification" implicit in a properly designed cavity setup. (Basically for a
low-loss cavity, say of finesse = 40 and resonance transmission \( \simeq 1/3 \), one
can parlay a 1% resonant absorption signal into a 10% fractional change of a
10 times lower basic dc power level. This gambit is obviously helpful in
the frequency reproducibility business.)

In the laser spectrometer of Fig. 4, the digitized resonance line shapes
are obtained directly without frequency modulation and associated derivative
detection. Thus it is appropriate to least-squares fit the data to the Lorentzian
function itself. An antisymmetric term proportional to the Lorentz
derivative is useful in lowering the residuals, although its contribution
seldom exceed 1% of the basic signal [11].

A fit typical of all the data to be presented is shown as Fig. 5. The
Lorentz fit becomes even more precise at higher pressure. By now several

![Graph](image)

Figure 5. Lorentz fit to typical data from "First FBTK" run in external
120-cm cell. The saturation parameter is 0.38 and the observed linewidth
(HWHM) is 33.6 kHz. The signal as plotted is inverted.
hundred resonance curves have been recorded. Our "eyeball" estimates of the linewidths and heights have been checked out in a few dozen interesting cases using the computer least-squares-fit program. As suggested previously, we do not ascribe any fundamental significance to these resonance parameters; rather they form a convenient media for discussion of the data.

(As an interesting side topic, in Fig. 6 we show a line shape characteristic of low-pressure data: The smooth reference curve is the pure Lorentz function. Basically, at low pressures, fluctuations in the collision rate from its average value occasionally allow some slow absorber molecules to interact for quite a long time. At low powers these particles give a disproportionate and very sharp contribution at line center. It is found that this line shape can be well reproduced by "peaking" the Lorentzian through subtraction

![Graph showing resonance curves](image)

Figure 6. Lorentz fit to 0.1 m Torr data from "First FBTK" run. The saturation parameter is 0.25. The measured linewidth is 25.6 kHz (HWHM). Note the characteristic sharpening in the line wings at low power and low pressure. Although the whole signal represents only about 1.5% of the 0.7% unsaturated double-pass absorption, the optical isolation is sufficient to prevent any important curvature of the experimental baseline due to residual interference effects.
of a second-derivative term, although presumably a Gaussian would be more appropriate at the lowest pressure [12].

This body of data is made vastly more useful by considering the ensemble of resonance widths to depend parametrically upon aperture, laser power, and absorber pressure. As a matter of convenience we will model the resonances by pressure-broadened and power-broadened homogeneous holes "burned" into a Gaussian velocity distribution. Thus we define \( P_\perp \) as the linewidth intercept at zero pressure and zero power—ultimately to be related by a proper theory to the mode size [13]. In view of the nonlinear nature of the saturation resonance, we will consider only molecules which transit the entire laser spot without important collisions, neglecting those particles of shorter interaction time whose coherence is terminated by collisions. Thus it is reasonable to identify \( P_\perp \) as the energy "relaxation" rate \( R_\perp \) energy being carried out of the laser beam as molecular excited-state occupation probability. The phase interruption rate, \( \gamma \), is increased beyond \( R_\perp \) by very-long-range collisions, typically of impact parameter greater than \( \sim 15 \text{ Å} \). More robust collisions deviate the molecular trajectory, transferring enough \( z \)-axis momentum to destroy the saturation resonance condition [2]. These molecules do not contribute very much to the signal due to the intrinsically nonlinear character of the saturated absorption process.

We take the phase interruption rate

\[
\gamma (\text{Pressure}) = P_\perp + P_\parallel \times (\text{Pressure})
\]

If \( \mu \) is the transition dipole moment, the maximum radiative interaction rate may be represented by \( \mu E_0 / \hbar \), where \( E_0 \) is the optical electric field on the laser beam axis due to one running-wave component.

Thus we could define the saturation parameter as \( S_0 \equiv (\mu E_0 / \hbar)^2 / \gamma R_\perp \). However, for the present purposes it is more useful to define an effective saturation parameter, \( S \), by the relation

\[
S = \frac{\mu^2}{\hbar} \frac{1}{\gamma R} \langle E^2 \rangle
\]

where \( \langle E^2 \rangle \) is an "effective" electric field (squared) resulting from a suitable three-dimensional average of the laser intensity distribution. For two-level quantum systems interacting with a coherent radiation field, it is known that saturation increases the linewidth by the factor \((1+S)^{1/2} \) [2, 3, 14]. Thus we can model

\[
\Gamma (\text{Pressure, Power}) = \gamma (\text{Pressure}) (1+S)^{1/2}
\]

giving

\[
\Gamma^2 (\text{Pressure, Power}) = \gamma^2 (\text{Pressure}) + \gamma (\text{Pressure}) \frac{\mu^2}{\hbar} \langle E^2 \rangle / R.
\]
The fitting function used was therefore

\[
\left[ \Delta \nu_{\text{HWHM}}(\text{Pressure, Power}) \right]^2 = \left[ P_1 + P_2 \times (\text{Pressure}) \right]^2 + \left[ P_1 + P_2 \times (\text{Pressure}) \right] \times P_3 \times (\text{Power})/P_1 \quad (5)
\]

The great utility of this gambit may be surmised from Figs. 7 and 8 which are essentially two perpendicular slices through the three-dimensional surface of linewidth versus pressure and power. In both figures the "c" character indicates the experimental linewidth, the "0" is the data corrected for the finite value of the orthogonal abscissa. Thus in Fig. 7 we have a set of power-broadening experiments at five values of the pressure. Note all results at fixed power are projected onto essentially the same point by the

![Graph](image)

Figure 7. Projection onto pressure = 0 plane of experimental linewidth surface. The measured points are plotted "c". For the corrected data, plotted as "0", Eq. (5) has been used to remove the effect of nonzero pressure in the experiment. Equation (5) is plotted as the smooth line through the "0"'s. See text regarding points at low powers.
Figure 8. Experimental line width surface projected onto power = 0 plane. Equation (5) was used to correct for finite measurement intensity. See caption of Fig. 7.

assumed pressure correction. Similarly in Fig. 8 we have experimental linewidth at several power levels versus methane pressure. The saturation correction is sufficient to essentially eliminate the power broadening [12].

A pressure-dependent T1 model gave, within the statistical limits, the same "best-value" parameters as our assumed model, with a fairly negligible improvement in the residuals. Several other intensity dependences were investigated, but none gave residuals as small as the assumed form.

The output parameters for data taken at five apertures are presented in Table I. The half-linewidth, P1, varies from 550 to 23 kHz for a larger-than-corresponding change in the laser spot size. The data are presented in historical order to allow the reader to assess for himself the significance of spatial mode purity. For example the first two experiments, EUZ and First FRTK, differ only subtly in the focusing and recollimation adjustments. The EUZ data are presented only to illustrate the significance of the spatial
TABLE I. (All entries are ordinary frequency, not radian frequency.)

<table>
<thead>
<tr>
<th>Run ID</th>
<th>Beam&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P&lt;sub&gt;1&lt;/sub&gt; (kHz)</th>
<th>P&lt;sub&gt;2&lt;/sub&gt; (kHz/mTorr)</th>
<th>P&lt;sub&gt;3&lt;/sub&gt;/P&lt;sub&gt;1&lt;/sub&gt; (kHz/mW)</th>
<th>P&lt;sub&gt;2&lt;/sub&gt;/P&lt;sub&gt;3&lt;/sub&gt; (mW)</th>
<th>ω&lt;sub&gt;P&lt;/sub&gt;&lt;sup&gt;1&lt;/sup&gt; (kHz mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUZ&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5 mm</td>
<td>29.3</td>
<td>12.66</td>
<td>8.9</td>
<td>3.3</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>± 2.3 ± 0.17</td>
<td>± 2.0 ± 1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First FBTK&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5 mm</td>
<td>23.28</td>
<td>9.92</td>
<td>9.10</td>
<td>2.56</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>± 0.48 ± 0.21</td>
<td>± 0.85 ± 0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small spot&lt;sup&gt;d&lt;/sup&gt;</td>
<td>85 μ</td>
<td>846</td>
<td>10.27</td>
<td>635</td>
<td>1.33</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>± 45 ± 0.91</td>
<td>± 162 ± 0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First cavity&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.706 mm</td>
<td>104.0</td>
<td>15.7</td>
<td>59.4</td>
<td>1.75</td>
<td>73</td>
</tr>
<tr>
<td>total data</td>
<td>± 5.6 ± 0.4</td>
<td>± 6.0 ± 0.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second cavity&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.490 mm</td>
<td>141.3</td>
<td>13.26</td>
<td>148.0</td>
<td>0.95</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>± 2.3 ± 0.09</td>
<td>± 5.5 ± 0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>For 1/e<sup>2</sup> intensity.
<sup>b</sup>Passive beam expansion.
<sup>c</sup>Same as footnote a but better alignment.
<sup>d</sup>Focused spot.
<sup>e</sup>Length-stabilized resonant cavity, F ~ 40.

aspects of these experiments. Because of this sensitivity it is difficult to assess the systematic error effects in the entries in Table I: The errors listed are standard-deviation random errors only. Even so, we think that there is a meaningful peak in the pressure-broadening coefficient, P<sub>2</sub>, near apertures of 1/2 mm. A detailed interpretation of this effect in terms of a collision pseudopotential would be facilitated if we could obtain reliable pressure-shift data as well, but the data at the larger apertures are just too uncertain. (The shift is only ~ 75 Hz/mTorr.)

In Table I we report P<sub>3</sub>/P<sub>1</sub> (rather than P<sub>3</sub>) as it is probably a more useful number. The modest cross correlation implicit in the least-squares procedure affects the error estimate.

In the last column of Table I, the deviation upward from ω<sub>0</sub>P<sub>1</sub>=70 kHz mm is perfectly correlated with other kinds of knowledge regarding the spatial quality of the laser beam (tolerance of the auto collimation condition, size of the interference fringes with reduced optical decoupling, etc).

In the next-to-last column of Table I, one may see that, as expected, the
concept of a "saturation intensity" in lifetime-limited (atomic) cases maps into an analogous "saturation power" for long-lived absorbing particles whose radiative interaction time is dominated by the transit-time effect. Corresponding data on the signal size, rather than width, are available from the experiments described. However a number of interesting new factors are operative in this case, and we will defer discussion of the mode-size scaling results to a later publication.

In this paper we show that large amount of saturation data can be represented by a simple model. We hope that the availability of these results and demonstration of the time-of-flight scaling effect will encourage someone to develop a real theory applicable in this interesting limit.

ACKNOWLEDGMENTS

In the course of this work a number of colleagues have made important contributions. Throughout the last several years the productive and pleasant collaboration with R. L. Barger has been appreciated and enjoyed. Some of the early data were taken in collaboration with E. E. Uzgiris. As usual, P. L. Bender has made a number of useful suggestions and instructive criticisms. We thank J. Levine for applying some of his computer expertise to our Lorentz curve-fitting problem. The cheerful and competent computing help of Mrs. P. Kunasz is appreciated.

REFERENCES

7. The author is indebted to G. Kramer of PTB for discussion of this idea.
11. This small asymmetry is found to be strongly saturation-dependent and is thought to be a manifestation of the "recoil momentum" effect. See A. P. Kol’chenko, S. G. Rautian, and R. I. Sokolovskii, Zh. Eksp. Teor. Fiz. 55, 1864 (1968) [Sov. Phys. JETP 28, 986 (1968)].
12. In all of the data presented here we have substantially eliminated the regime of anomalously narrow lines near zero power and zero pressure: different physical processes are operative there.
13. As already noted, at very low powers and pressures the line is formed mostly by very slow particles. However $P_1$ here is to be associated with the same average transverse velocity that is important for the line formation in the regime of moderate pressure and power broadening (say > 20%). Thus we give an intercept definition for $P_1$ and for the present modelling purpose do not take the fit very seriously just near the origin.