A 'non-classical' information theory of spectral line shape

1970 J. Phys. A: Gen. Phys. 3 335

(http://iopscience.iop.org/0022-3689/3/4/003)

J. Phys. A: Gen. Phys., 1970, Vol. 3. Printed in Great Britain

A 'non-classical' information theory of spectral line shape

J. G. POWLES† and B. CARAZZA‡§

†The Physics Laboratories, University of Kent, Canterbury, England
‡Centro di Spettroscopia Moleculare, Cagliari, Sardinia

MS. received 11*th March* 1970

Abstract. Information theory is used to obtain the most probable spectral line shape given only a knowledge of a finite number of moments of the line. The analysis of Powles and Carazza is extended to a consideration of non-classical information theory. This gives the possibility of a Lorentzian line shape which was not previously available. It is shown, in particular, that the common observation of a line shape narrowing from Gaussian to Lorentzian is simply explained and can be related to a change in physical parameters. The method is applied to problems in magnetic resonance but is of wide validity.

1. Introduction

The use of information theory in statistical mechanics is well known (Jaynes 1957, Katz 1967) but it has not been used, as far as we are aware, in the analysis of spectral line shape. We have recently applied this method to the problem of the absorption line shape in nuclear magnetic resonance (Powles and Carazza 1970). This analysis was quite successful in predicting actual line shapes given only a few of the moments of the line, but did not include the commonly observed Lorentzian line shape which arises, for instance, in magnetic resonance for motional or exchange narrowing. In this paper we report an information theory which uses non-classical statistics which enables us to include the Lorentzian shaped line, in particular, and very considerably widens the scope for the analysis of spectral distribution by the information theory method. This and the previous analysis (Powles and Carazza 1970) is quite general and may be applied to almost any situation where the results can be described by a spectrum for which moments of the distribution can be calculated. Further applications will be described in a later publication.

2. Classical information theory line shapes

We define the nth moment M_n of a spectral distribution $f(\omega)$ by

$$M_n \equiv \frac{\sum \omega_i{}^n f(\omega_i)}{\sum f(\omega)}. \tag{1}$$

In most cases these moments can be calculated, at least in principle, and are usually the only *exact* information about a spectral line shape which can be calculated. In practice only a few lower moments are actually available owing to computational and/or algebraic difficulties for complex systems. However, the moments do not actually give the line shape and we consider the problem of finding the most probable line shape given a finite number of moments. This is achieved by minimizing the information

$$\mathscr{I} \equiv \sum p(\omega_i) \ln \{p(\omega_i)\} \tag{2}$$

§ Present address: Istituto di Fisica Università, Parma, Italy.

where $p(\omega_i)$ is the probability of finding a component of the spectrum at ω_i and $p(\omega) \propto f(\omega)$. Earlier we showed for instance that, given a finite total intensity, and a knowledge of the second moment only, the most probable line shape is Gaussian,

$$f(\omega) \propto \exp(-\lambda\omega^2) \tag{3}$$

where $\lambda = \frac{1}{2}M_2$.

If only the second and fourth moments are known the most probable line shape is

$$f(\omega) \propto \exp(-\lambda\omega^2 - \mu\omega^4) \tag{4}$$

where λ and μ are obtained from M_2 and M_4, and so on.

Although such expressions include a surprising variety of line shapes (Powles and Carazza 1970) they do not include a Lorentzian and more particularly they do not include 'narrowed' lines since the ratio of M_4/M_2^2 does not exceed three, except in some very special cases.

An alternative analysis to that given by Powles and Carazza (1970) who used equation (2), will prove to be more useful in this paper. We consider complexions more directly, as one does in statistical mechanics, although we must emphasize that we are not necessarily concerned with statistical mechanics problems as such.

Suppose we have n_i contributions to the intensity of the spectrum at ω_i. The total number of contributions is

$$\mathcal{N} = \sum n_i. \tag{5}$$

The number of configurations W_{B} is

$$W_{\mathrm{B}} = \frac{\mathcal{N}!}{\prod(n_i!)}. \tag{6}$$

The second moment M_2 of the line is

$$M_2 = \mathcal{N}^{-1} \sum n_i \omega_i^2. \tag{7}$$

Maximizing W_{B}, i.e. minimizing $\mathscr{I}$, subject to (5) and (7) we have,

$$\sum_i \left(\frac{\partial \ln W_{\mathrm{B}}}{\partial n_i} + \alpha + \lambda\omega_i^2\right) \delta n_i = 0 \tag{8}$$

where α and λ are Lagrange undetermined multipliers.

Hence

$$\frac{\partial \ln W_{\mathrm{B}}}{\partial n_i} + \alpha + \lambda\omega_i^2 = 0 \qquad \text{for all } i. \tag{8$'$}$$

Hence using (6) and Stirling's relation we have

$$\frac{\partial W_{\mathrm{B}}}{\partial n_i} = -\ln n_i \tag{9}$$

so that

$$n_i = \exp(-\alpha - \lambda\omega_i^2) \tag{10}$$

i.e.

$$f(\omega) = A_{\mathrm{B}} \exp(-\lambda\omega^2) \tag{3}$$

as before, for M_2 only known. Corresponding generalizations can be written down immediately.

3. Non-classical information theory line shapes

The analysis just given as well as that of Powles and Carazza (1970) and the definition of information in equation (2) correspond to classical, or Boltzmann statistics. This corresponds to the assumption that each contribution to a spectral line is 'distinguishable'. We now use Bose–Einstein statistics to calculate W, i.e. we assume that the individual contributions to $f(\omega)$ at a given ω are not distinguishable. As in statistical mechanics (e.g. Wilks 1961) we group the contributions in ranges k with n_k contributions at or near ω_k with 'degeneracy' g_k so that

$$W_{\mathrm{E-B}} = \prod_k \frac{(n_k+g_k-1)!}{n_k!(g_k-1)!} \tag{11}$$

so that

$$\frac{\partial W_{\mathrm{E-B}}}{\partial n_k} = \ln\left(\frac{n_k+g_k-1}{n_k}\right). \tag{12}$$

We maximize $W_{\mathrm{E-B}}$ subject to fixed $\mathcal{N}$ and known M_2 which yields

$$n_k = \frac{g_k}{\exp(\beta+\lambda\omega_i^2)-1} \tag{13}$$

i.e.

$$f(\omega) = \frac{A}{\exp(\beta+\lambda\omega^2)-1}. \tag{14}$$

Clearly we must have $\exp\beta > 1$ and $\lambda > 0$.

Corresponding generalizations for any number of known moments can be written down immediately. A correspnds to a 'density of states' function. We take it to be a constant since this is the minimum information assumption.

If we had assumed we knew only the first moment we would have found instead of (14)

$$f(\omega) \propto \frac{1}{\exp(\beta+\alpha\omega)-1} \tag{15}$$

which in statistical mechanics is of course the Einstein–Bose distribution.

For simplicity we shall consider only equation (14) and leave more elaborate line shapes for consideration in a later paper.

Clearly, when β is large (14) goes over to (3), i.e. a Gaussian line shape

$$f(\omega) \simeq A\exp(-\beta)\exp(-\lambda\omega^2). \tag{3'}$$

If β and λ are small the exponential in (14) may be expanded to give

$$f(\omega) \simeq \frac{A}{\beta+\lambda\omega^2} \tag{14'}$$

which is a Lorentzian line shape, provided ω is 'not too large'. In fact it is a cut-off Lorentzian, as we show below. The situation β and λ small corresponds in statistical mechanics to an approach to Einstein–Bose condensation at low temperatures. Thus, as the parameter β varies from large to small, our line shape (14) goes over from Gaussian to Lorentzian and we shall show that for the special case of fixed second moment this corresponds to line narrowing, as is observed for instance for motional or exchange narrowing in magnetic resonance and many other fields.

In our original Boltzmann analysis the line shape, for M_2 given, was entirely determined by M_2 and the total intensity. However, in the Einstein–Bose case we have, for M_2 given, three parameters A, β and λ since in this case A and β appear separately and not as the combination $A\exp(-\beta)$. We propose to circumvent this difficulty by assuming first we know only the total intensity and M_2, which gives (14), and then that we have a disposable parameter β which fixes the value of M_4 and so fixes the values of λ and A. We show later (equation 24) that, at least in one application, β can be related to a physical variable.

We can show that

$$N \equiv \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{\exp(\beta+\lambda\omega^2)-1} = \left(\frac{\pi}{\lambda}\right)^{1/2} \sum_{n=1}^{\infty} e^{-n\beta} n^{-1/2} \tag{16}$$

$$M_2 = \tfrac{1}{2} N^{-1} \pi^{1/2} \lambda^{-3/2} \sum_{n=1}^{\infty} e^{-n\beta} n^{-3/2} \tag{17}$$

and

$$M_4 = \tfrac{3}{4} N^{-1} \pi^{1/2} \lambda^{-5/2} \sum_{n=1}^{\infty} e^{-n\beta} n^{-5/2}. \tag{18}$$

Using (17) and (18) we have

$$\frac{M_4}{M_2^2} = \frac{3\left(\sum_{n=1}^{\infty} \mathrm{e}^{-n} n^{-5/2}\right)\left(\sum_{n=1}^{\infty} \mathrm{e}^{-n} n^{-1/2}\right)}{\left(\sum_{n=1}^{\infty} \mathrm{e}^{-n\beta} n^{-3/2}\right)^2}. \tag{19}$$

Clearly if we choose $\beta \to \infty$ then $M_4/M_2^2 \to 3$ and $\lambda \to \frac{1}{2} M_2$ as for a Gaussian. If we choose $\beta \to 0$ but M_2 finite, then $M_4/M_2^2 \to \infty$, and $\lambda \to 0$ as for a Lorentzian, thus confirming (14′). In fitting a curve $f(\omega)$, given M_2 and M_4, it is convenient to find β using (19) and then λ using (17) and (16), and to find A from (20) using the known total intensity I where

$$A = IN^{-1}. \tag{20}$$

If

$$S_i \equiv \sum_{n=1}^{\infty} \exp\{-(n-1)\beta\} n^{(1/2-i)} \tag{21}$$

we have

$$\frac{M_4}{M_2^2} = 3\frac{S_1 S_3}{S_2^2} \qquad \text{for } \beta \tag{19′}$$

$$\lambda = (S_2/S_1)(2M_2)^{-1} \tag{17′}$$

and

$$A = I(2\pi M_2)^{-1/2} S_2^{1/2} S_1^{-3/2} \mathrm{e}^{\beta}. \tag{20′}$$

It will also be useful to consider the full width at half height $\Delta\omega_{1/2}$ of $f(\omega)$ in (14) which is given by

$$\Delta\omega_{1/2} = 2\left(\frac{2M_2 S_1}{S_2}\right)^{1/2} \{\ln(2-\mathrm{e}^{-\beta})\}^{1/2}. \tag{22}$$

For $\beta \to 0$

$$\Delta\omega_{1/2} \to 2\pi^{1/4}\{\xi(\tfrac{3}{2})\}^{-1/2}\beta^{1/4}(2M_2)^{1/2} \tag{22′}$$

where $\xi(\frac{3}{2}) = 2{\cdot}612\ldots$.

In magnetic resonance I and M_2 are constant (e.g. Abragam 1961) independent of the motion and so we illustrate the line shape (14) in figure 1, for $I = 1$ and for fixed M_2 for various values of β. The parameters are given in table 1.

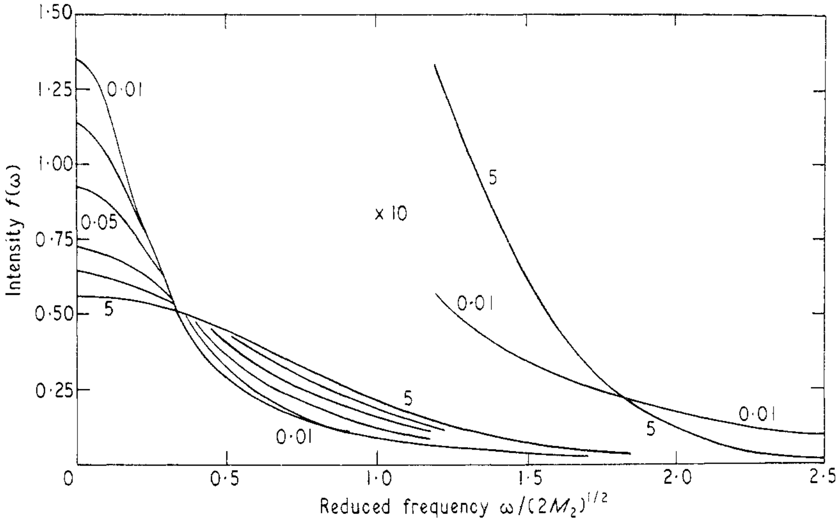

Figure 1. Plots of (half) the line shape $f(\omega)$, equation (14), for $I = 1$ and fixed M_2, for β equal to 5, 0·5, 0·2, 0·05, 0·02 and 0·01. Expanded ordinate scales are used to show the behaviour at large ω.

Table 1. Parameters for line shape for M_2 known

β	S_1	S_2	S_3	M_4/M_2^2	$\lambda \times 2M_2$	A/I	$\Delta\omega_{1/2}/(2M_2)^{1/2}$
∞	1	1	1	3	1	0·564	1·665
5	1·005	1·002	1·001	3·006	0·997	0·561	1·664
0·5	1·891	1·336	1·142	3·63	0·707	0·251	1·371
0·2	3·107	1·606	1·223	4·42	0·517	0·131	1·136
0·05	6·81	1·989	1·299	6·67	0·294	0·0452	0·805
0·02	11·30	2·183	1·322	9·29	0·196	0·0225	0·633
0·01	16·43	2·294	1·331	12·12	0·144	0·0134	0·525
0	∞	2·612... $(=\xi(\frac{3}{2}))$	1·341... $(=\xi(\frac{5}{2}))$	∞	0	∞	0

Figure 1 shows, for $I = 1$, how the Gaussian narrows to a Lorentzian and the intensity near $\omega = 0$ increases as β gets smaller. However, for large enough $\omega/(2M_2)^{1/2}$, the curve of $f(\omega)$ cuts off so that the second moment remains finite however small the value of β. This is illustrated in figure 2 where we compare $f(\omega)$ for $\beta = 0{\cdot}01$ with a Lorentzian having the same intensity and the same width at half height. Even for this relatively small narrowing (by the factor 3·2) the curve is almost Lorentzian and an experimental fit would probably be made for a Lorentzian with maximum intensity 1·34 which is not very different from $f(\omega)$, except for large ω, as indicated on figure 2.

For 'narrowed' lines, i.e. β small, it is usual to express the Lorentzian in terms of a correlation time τ (e.g. Kubo 1962) where

$$\Delta\omega_{1/2} = 2M_2\tau. \tag{23}$$

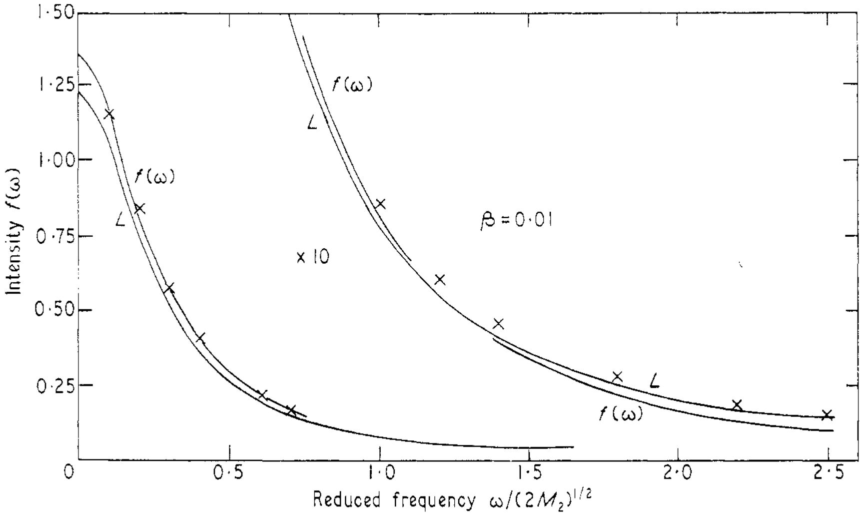

Figure 2. A comparison of $f(\omega)$, equation (14), for $\beta = 0{\cdot}01$ and a Lorentzian L of the same width at half height and total intensity. Also shown is a Lorentzian $\times$ of the same width at half height and the same intensity at $\omega = 0$.

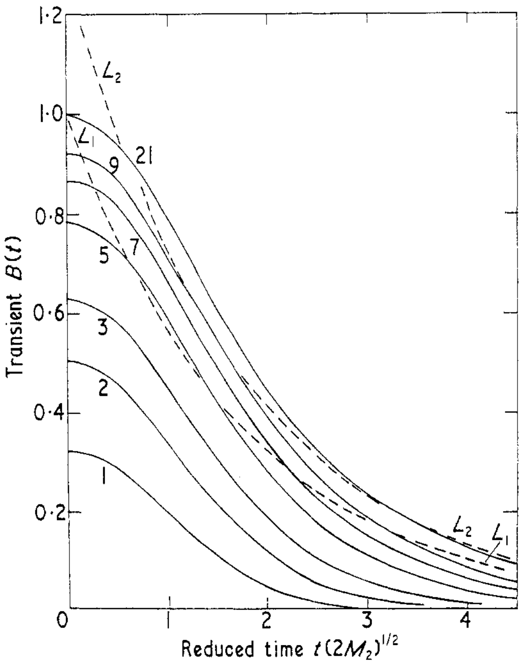

Figure 3. The transient decay $B(t)$, equation (25), corresponding to $f(\omega)$ for $\beta = 0{\cdot}2$, showing how $B(t)$ is a sum of Gaussians for 1, 2, 3, 5, 7, 9 and 21 terms. Also shown is an exponential decay, L_1, for the same width at half height and total intensity as $f(\omega)$. The exponential L_2 corresponds to the same width at half height but is a 'best fit' to $B(t)$.

Consequently, using (22′) we can identify β in terms of τ, i.e.

$$\tau \propto \beta^{1/4} \qquad \text{for } any \text{ value of } \beta. \tag{24}$$

This gives, for instance, a very plausible method of interpreting motional or exchange line narrowing in magnetic resonance. The curves $f(\omega)$ which we give in figure 1 are very similar to those given by the Kubo–Anderson (1954) theory of spectral line narrowing which also applies to narrowing from a Gaussian to, effectively, a Lorentzian.

4. Transients

Spectral information may often be obtained by the transient decay method. The transient is simply the Fourier transform of $f(\omega)$. The Fourier transform $B(t)$ of (14) (normalized to unity at $t = 0$) is

$$B(t) = S_1^{-1} \sum_{n=1}^{\infty} \exp\{-(n-1)\beta\} n^{-1/2} \exp(-t^2)/4n\lambda. \tag{25}$$

For $\beta \to \infty$ this approaches the Gaussian $\exp(-t^2/4\lambda)$, and for $\beta \to 0$ it approaches the exponential $\exp(-2|t|/\Delta\omega_{1/2})$, where $\Delta\omega_{1/2}$ is given by (22′).

It is of interest that $B(t)$ has no cusp at $t = 0$ and always cuts off faster than the Lorentzian (for β small) for large enough times. It shows how the transient (called the Bloch decay in magnetic resonance), for finite β, is made up of a sum of Gaussians. This is illustrated explicitly for $\beta = 0{\cdot}2$ in figure 3. In this figure we also show a Lorentzian with the same $\Delta\omega_{1/2}/(2M_2)^{1/2}$ and total intensity and also the 'best fit' Lorentzian. Even for such modest narrowing (by the factor 1·5) as represented by $\beta = 0{\cdot}2$ a large number of terms in (23) are required for an adequate representation of $B(t)$. We show the result for 21 terms which gives 99·5% of $B(t)$.

References

Abragam, A., 1961, *Principles of Nuclear Magnetism* (Oxford: Clarendon).

Anderson, P. W., 1954, *J. Phys. Soc. Japan*, **9**, 316–39.

Jaynes, E. T., 1957, *Phys. Rev.*, **106**, 620–30; **108**, 171–90.

Katz, A., 1967, *Principles of Statistical Mechanics* (San Francisco and London: Freeman).

Kubo, R., 1962, *Fluctuation, Relaxation and Resonance in Magnetic Systems*, Ed. D. Ter Haar (Edinburgh: Oliver and Boyd).

Powles, J. G., and Carazza, B., 1970, *Proc. Conf. on Magnetic Resonance, Monash*, 1969 (New York: Plenum) to be published.

Wilks, J., 1961, *The Third Law of Thermodynamics* (London: Oxford University).