

Comparison of algorithms for the calculation of molecular vibrational level densities

K. Hansen^{a)}

Department of Physics, University of Gothenburg, SE-412 96 Göteborg, Sweden

(Received 25 February 2008; accepted 15 April 2008; published online 19 May 2008)

Level densities of vibrational degrees of freedom are calculated numerically with formulas based on the inversion of the canonical vibrational partition function. The calculated level densities are compared with other approximate equations from literature and with the exact Beyer–Swinehart values, for which a simplified but equivalent version is given. All approximate equations agree at high excitation energies, but our results are vastly superior at low energies for large molecules. The results presented here are therefore of particular relevance for thermal processes of very large molecules, e.g., of biological nature, for which the exact state counting can be prohibitively slow. Furthermore, it is valid for situations where anharmonic motion significantly influences the thermal properties. © 2008 American Institute of Physics. [DOI: 10.1063/1.2920483]

INTRODUCTION

In the statistical description of equilibrium and quasi-equilibrium properties of molecules and small particles, the motion of the nuclei is of overwhelming importance. This is due to the fact that in equilibrium distributions most of the excitation energy resides in these degrees of freedom. The rate with which activated processes proceed crucially depends on both the excitation energy and the total number of vibrational states for the systems available at that energy, and a quantitatively correct description of the energy dependence of thermal and quasithermal rate constants hinges on the ability to correctly count the number of states of the system. The function representing this enumeration is the exponential of the Boltzmann entropy or the level density, $\rho(E_m)$, which counts the number of quantum states of the system between the energies E_m and $E_m + dE_m$ (m denotes the microcanonical energy). It can be formally defined as the derivative with respect to the energy of the number of states, W , of the system up to E_m as follows:

$$\rho_{\text{tot}}(E_m) = \frac{dW}{dE_m}. \quad (1)$$

$W(E_m)$ is a step function, which increases by unity every time the energy increases so much that a new state is at or below E_m and ρ will therefore, in principle, be a sum of delta functions, or a sum of Lorentzians if the finite lifetime broadening of the levels is important. The discrete nature of ρ will be apparent at low energies when the total energy is comparable to a typical vibrational quantum $\hbar\omega$, $E_m \approx \hbar\omega$. At higher energies, the spacing between the delta functions will rapidly tend to zero and the function can then conveniently be described as continuous. One can define a fluctuating and a smooth part of the function as

$$\rho_{\text{tot}}(E_m) = \rho(E_m) + \rho_{\text{fluc}}(E_m), \quad (2)$$

where the first term on the right hand side is a smooth function of energy. The fluctuating part of the level density, ρ_{fluc} , is zero when averaged over an energy interval, δE_m , which fulfills the inequalities $\delta E_m \gg 1/\rho_{\text{tot}}(E_m)$ and $\delta E_m d\rho_{\text{tot}}/dE_m \ll \rho_{\text{tot}}$. Within these limits, the definition of the subject of this paper, $\rho(E_m)$, is unambiguous.

In the remainder of the paper, some background and a simple derivation of the equations in Ref. 1 will be given, including an estimate of the leading order correction, before the numerical comparison is presented. To compare with literature formulas, the systems considered here will be restricted to collections of harmonic oscillators, although our equations are applicable to any system for which the canonical partition function can be explicitly written down. Other important examples of such systems are sets of uncoupled anharmonic oscillators, for which the total canonical partition function is the product of the partition functions for the individual degrees of freedom. As will be clear from the derivation of Eq. (16) presented below, the formulas for level densities calculated here are only valid for level densities that increase with energy, e.g., for at least two harmonic oscillators in the high energy limit. This is a minor limitation.

BACKGROUND

At high energies, when all oscillators are excited with a probability approaching unity, ρ can be calculated with an expansion in the reciprocal excitation energy. The expansion is based on a comparison of the thermal excitation energy and the entropy in a canonical ensemble with an *Ansatz* for the level density of the form

$$\rho = c(E_m + E_0)^{s-1}, \quad (3)$$

where c , E_m , and s are constants. The canonical partition function, Z , is generally calculated as

^{a)}Electronic mail: klaws@physics.gu.se.

$$Z = \int_0^\infty \rho(E_m) e^{-E_m/k_B T} dE_m, \quad (4)$$

and the high temperature limit of Z for the level density in Eq. (3) therefore becomes

$$Z \approx c(s-1)!(k_B T)^s e^{E_0/k_B T}, \quad (5)$$

where T is the canonical temperature.

The canonical partition function for $3N-6$ harmonic oscillators with frequencies $\omega_i, i=1, \dots, 3N-6$ is also

$$Z = \prod_i (1 - e^{-\hbar\omega_i/k_B T})^{-1}. \quad (6)$$

Calculation of the Helmholtz free energy, $-k_B T \ln(Z)$, with Eq. (5) and with an expansion of Eq. (6) in the inverse temperature permits a term-by-term comparison, which confirms the *Ansatz* in Eq. (3) and gives the values

$$s = 3N - 6,$$

$$c = \frac{1}{(s-1)! \prod_i \hbar\omega_i}, \quad (7)$$

$$E_0 = \frac{1}{2} \sum_i \hbar\omega_i.$$

This result is well known. With E_0 set to zero, it is the Kassel level density.

The expansion of Eq. (6) and the approximation made in the calculation in Eq. (5) hinge on the smallness of the quantities $\hbar\omega_i/k_B T$ and $\sum_i \hbar\omega_i/sk_B T$. The first of these presents a serious problem when attempting to amend the equations to expand the range of validity to low canonical temperatures, because convergence of the expansion requires inclusion of an increasingly large number of terms.

This problem has been addressed previously by several authors. Two formulas that have found widespread use are the ones given by Haarhoff² and by Whitten and Rabinovitch.³ Haarhoff's approach basically consists of an expansion of the inverse Laplace transform of the partition function in $1/E_m$. The formula given by Whitten and Rabinovitch is derived with a modification of the high energy limit of the level density, and the effective number of oscillators was adjusted as a function of excitation energy. The parameters for the adjustment were derived from a fit to several numerical calculations of a number of rather small molecules. It is limited to energies $0.05\sum_i \hbar\omega_i \leq E_m \leq 4\sum_i \hbar\omega_i$, where the most important restriction is the low energy limit. The formula has two different forms for the two energy intervals and these are pasted together at $E_m = \sum_i \hbar\omega_i/2$.³ Both of these methods involve a small number of parameters derived from the set of vibrational frequencies considered and a few numerical constants.

Other methods have been suggested,⁴⁻⁶ based on inversion of the Laplace transform with systematic expansions. Although some of the methods have the potential to reach good numerical accuracy or even exact solutions, they re-

quire varying amounts of numerical work. This seriously reduces the transparency of the methods and they have found limited use for analytical calculations.

For the purpose of purely numerical evaluations, the recursive Beyer-Swinehart algorithm⁷ seems superior. With this algorithm, one calculates the level density for a collection of oscillators as a convolution of the level densities of the individual oscillators. If a set of $n-1$ harmonic oscillators has a total level density denoted by $\rho(E, n-1)$, the total level density of the same system with one more oscillator added to the system is

$$\rho(E_m, n) = \sum_{k=0}^{[E_m/\hbar\omega_n]} \rho(E_m - k\hbar\omega_n, n-1), \quad (8)$$

where ω_n is the frequency of the n th oscillator, and $[x]$ is the integer part of x . The level density of the physical system is obtained by iterating this equation until all the frequencies have been included. The method is recursive, fairly efficient, and only limited in accuracy by the energy discretization. Although originally applied to strictly harmonic motion, its main idea is to convolute level densities of independent degrees of freedom, and it can therefore be used for anharmonic motion with only minor modifications. The scaling of the computing time can be estimated to be $t \propto s(E_m/\delta E_m) \times \langle E_m/\hbar\omega \rangle \cong s^3(E_m/s\langle\hbar\omega\rangle)^2 \langle\hbar\omega\rangle/\delta E_m$, where s is the number of oscillators, δE_m is the resolution, and the square brackets denote an average over all frequencies. The resolution in units of the average quantum energy, $\delta E_m/\langle\hbar\omega\rangle$, is usually required to be on the order of 10^{-2} or less, and for $s=10^4$ the calculation of the level density at $E_m=s\langle\hbar\omega\rangle$ requires on the order of 10^{14} elementary operations. The memory requirement scales with $E_m/\delta E_m = s(\langle\hbar\omega\rangle/\delta E_m) \times \langle E_m/s\langle\hbar\omega\rangle \rangle$, i.e., 10^6 for this example.

When anharmonicities are not an issue, Eq. (8) can be rewritten in a form that allows a considerable saving in computer resources for systems with many degrees of freedom. To the best of our knowledge, the simplification has not yet been reported in literature. If one writes Eq. (8) for the energy $E - \hbar\omega_n$, it reads

$$\begin{aligned} \rho(E_m - \hbar\omega_n, n) &= \sum_{k=0}^{[(E_m - \hbar\omega_n)/\hbar\omega]} \rho(E_m - (k+1)\hbar\omega, n-1) \\ &= \sum_{k=1}^{[E_m/\hbar\omega]} \rho(E_m - k\hbar\omega, n-1). \end{aligned} \quad (9)$$

Forming the difference between this equation and Eq. (8) at energy E gives

$$\rho(E_m, n) = \rho(E_m - \hbar\omega_n, n) + \rho(E_m, n-1). \quad (10)$$

The reduction in computing time using this equation, which is mathematically equivalent to Eq. (8), is a factor that is on the order of the energy in units of a typical frequency. With the numbers used in the example above, we have a computing time that scales as $t \propto s E_m/\delta E_m = s^2(E_m/s\langle\hbar\omega\rangle) \times \langle\hbar\omega\rangle/\delta E_m = 10^{10}$. The memory requirements are unchanged. The level densities calculated with the algorithm in

Eq. (10) will be used as the common reference for the comparisons with the approximate formulas.

Although the Beyer–Swinehart algorithm is efficient, one may desire alternative procedures. One reason is the severe demand on computer memory for systems with a large number of degrees of freedom and, when Eq. (10) cannot be used for physical reasons, also on computer time. For large molecules, e.g., of biological interest, these limitations will often prevent a numerical study. However, the most important reason is, in our view, that it is an algorithm and not an analytical expression and hence it is difficult to develop any physical intuition for the outcome of the method.

THE RELATION BETWEEN LEVEL DENSITIES AND CANONICAL QUANTITIES

The starting point of the present numerical investigation is the level density given in Ref. 1. Here, we will derive it with simpler means and calculate corrections.

$$Z(T) \approx \rho(E_m) \int dE e^{-E/k_B T + (d \ln(\rho(E)))/dE(E-E_m) + (1/2)(d^2 \ln(\rho(E))/dE^2)(E-E_m)^2}. \quad (12)$$

The first two terms in the argument of the exponential cancel and the Gaussian integral is easily calculated to give

$$Z(T) \approx \rho(E_m) e^{-E_m/k_B T} \left(\frac{2\pi}{-d^2 \ln(\rho(E'))/dE^2|_{E=E_m}} \right)^{1/2}. \quad (13)$$

To find the value of the second derivative appearing in the square root, we consider an excitation energy, E' , and an inverse temperature, $1/T'$, which are close to the actual values E_m , $1/T$. They will be linearly related

$$1/k_B T' = 1/k_B T + \frac{d^2 \ln(\rho(E))}{dE^2} (E' - E_m) + \dots, \quad (14)$$

and hence

$$\frac{1}{d^2 \ln(\rho(E))/dE^2} = \frac{dE_m}{d(1/k_B T)} = -C k_B T^2, \quad (15)$$

where C is the heat capacity. This result enters the equation in a pre-exponential factor and we can therefore approximate the quantities on the right hand side with the canonical values without serious loss of accuracy. Introducing Eq. (15) into Eq. (13), we arrive at

$$\rho(E_m) \approx Z(T) e^{E_m/k_B T} \frac{1}{\sqrt{2\pi k_B T^2 C}}, \quad (16)$$

which is the desired expression for ρ as a function of the microcanonical temperature.

The value of $1/T$, which conforms to the expansion point in Eq. (11), must now be determined. We could use the canonical value in Eq. (17)

Level densities can be rigorously calculated as the inverse Laplace transform of the canonical partition function, Z , when this is known for all temperatures. For finite quantum mechanical systems, the inversion will by necessity produce a sum of delta functions, which seriously complicates the exact analysis. However, since we are interested in the smooth component of the total level density, it is possible to use a saddle point expansion of the level density in the integrand in Eq. (4). The expansion point is conveniently chosen as E_m , i.e.,

$$\left. \frac{d \ln(\rho(E))}{dE} \right|_{E=E_m} = 1/k_B T. \quad (11)$$

Other choices are possible and will influence the value of T , which is determined below. With the choice in Eq. (11), we can write the partition function to second order as^{1,8}

$$\bar{E}(T) = E_m, \quad (17)$$

with the motivation that the average energy in a canonical ensemble is close to the peak value of the integrand in the partition function in Eq. (4). This is the strategy used in Ref. 8. A more precise estimate is obtained if we require consistency and use Eq. (13) to calculate canonical excitation energy

$$\bar{E} = -\frac{\partial \ln(Z)}{\partial(1/k_B T)} = -\frac{d \ln(\rho(E_m))}{d(1/k_B T)} + E_m + \frac{1}{k_B T} \frac{dE_m}{d(1/k_B T)} + \frac{1}{2} \frac{d \ln(1/(k_B T)^2)}{d(1/k_B T)}. \quad (18)$$

The first term is

$$-\frac{d \ln(\rho(E_m))}{d(1/k_B T)} = -\frac{d \ln(\rho(E_m))}{dE_m} \frac{dE_m}{d(1/k_B T)} = -\frac{1}{k_B T} \frac{dE_m}{d(1/k_B T)}. \quad (19)$$

This term appears elsewhere in Eq. (18) with opposite sign and after cancellation we are left with

$$\bar{E}(T) = E_m + k_B T. \quad (20)$$

This says that the product of level density and Boltzmann factor peaks at an energy which is $k_B T$ less than the average thermal energy in a canonical system with the same temperature. This is the result that was derived in Ref. 1 in a somewhat less transparent way. The difference in the results obtained with and without the last term on the right hand side of Eq. (20) is usually minor, on the order of $\exp(-1/C)$. It is

relevant at low excitation energies, and was used in the calculation of level densities of helium droplets.⁹

The saddle point approximation used to calculate Eq. (16) will have corrections from third and higher order derivatives of the logarithm of $\rho \exp(-E/k_B T)$. We will now estimate the leading order correction from an expansion in the inverse of the width of the canonical energy distribution. For simplicity, we will not distinguish between canonical and microcanonical temperatures, energies, and heat capacities here. The expansion of the integrand contains the factors, left out in Eq. (12),

$$\exp\left(\frac{1}{3!} \frac{d^3 \ln(\rho)}{dE^3} (E - \bar{E})^3 + \frac{1}{4!} \frac{d^4 \ln(\rho)}{dE^4} (E - \bar{E})^4 + \dots\right). \quad (21)$$

The corrections are calculated to leading order by expanding the exponential and integrating the terms one by one with the Gaussian function. This renders the third order term zero and the lowest order correction is then the fourth order term. After integration, one gets a relative correction of

$$-\frac{1}{8} \left(\frac{6k_B}{C} + \frac{6k_B T}{C^2} \frac{dC}{dT} + \frac{3k_B T^2}{C^3} \left(\frac{dC}{dT} \right)^2 - \frac{k_B T^2}{C^2} \frac{d^2 C}{dT^2} \right). \quad (22)$$

This scales with $1/C$ and the worst case is therefore seen for two degenerate harmonic oscillators [a single oscillator is excluded, in general, by the conditions used to derive Eq. (16), as mentioned in the Introduction]. A numerical calculation of two degenerate oscillators with quantum energies $\hbar\omega$ gives an error, which is -0.5 for $k_B T = \hbar\omega$ and which approaches zero monotonically when temperature and/or heat capacity increases.

The application of Eqs. (16) and (20) is straightforward if an explicit expression for $Z(T)$ can be written down. For harmonic and uncoupled anharmonic oscillators, the total partition function factors into those of the individual oscillators and the canonical excitation energy are calculated as $\bar{E}(T) = k_B T^2 \partial \ln Z / \partial T$, which is then a sum of the individual canonical excitation energies. The only complication arises for anharmonicities, where one may need a high energy truncation or regularization of the spectrum because the leading order anharmonicities usually cause the level spacings to shrink when the vibrational quantum numbers increase, producing a formally divergent result for a calculation of the partition function. After the partition functions have been established, Eq. (20) is inverted to give $T(E_m)$. Finally, the partition function and the heat capacity are calculated for this temperature after which the level density directly follows from Eq. (16).

The computational effort lies in the inversion of Eq. (20). For convenience, one may calculate the canonical caloric curve $\bar{E}(T)$ at reasonably densely spaced points in the temperature and use this to locate the energy at which the function $\bar{E}(T) - E_m - k_B T$ changes sign. If better precision is required, the method of bisection intervals can be used. The calculation of the caloric curve requires a number of operations, which scale with s and the number of points on the curve. The iterative bisection procedure scales approximately with s . When results for more energies are desired, the ca-

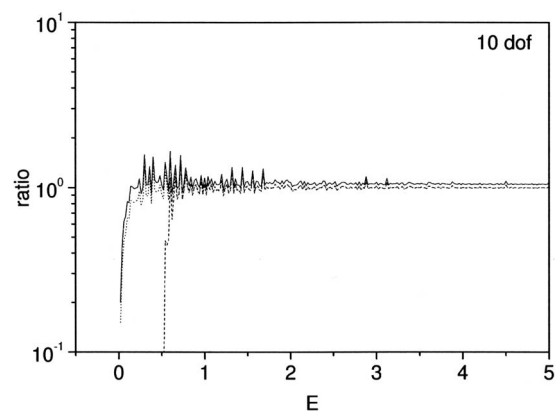


FIG. 1. Comparisons of numerical calculations of level densities for $s=10$ harmonic oscillators. Here and in the rest of the figures the full line is the result from Eq. (16), the dotted line is Haarhoff's result from Ref. 2, and the dashed line that of Whitten and Rabinovitch in Ref. 3. In this and all other figures, the excitation energies are given in units of the average vibrational frequency, $\hbar\langle\omega\rangle$. Here and in Figs. 2–4, the lowest calculated energies are equal to $0.01\hbar\langle\omega\rangle$. For more details, see text.

loric curve needs not be recalculated, and the number of operations needed to calculate $\rho(E_m)$ for N_E different energies thus scales as $N_E s^2$. It should be pointed out that one of the computational advantages of the present method compared to the Beyer–Swinehart algorithm is that, for the latter, the level density for all energies below E_m need to be calculated at a fairly small energy spacing to avoid discretization errors. This adds a factor of $E_m / \delta E_m$ in computing time, which is avoided with our method.

COMPARISON WITH NUMERICAL RESULTS

Figure 1–4 show the Beyer–Swinehart level densities relative to those calculated with the approximate level densities mentioned above, for particle sizes containing $N=10, 100, 1000$, and $10\,000$ oscillators. A deviation to low values indicates an approximate value which is too high and vice versa. This representation is dictated by the fact that the approximate values are finite whereas the Beyer–Swinehart values are occasionally zero for low energies. The curves shown are average values of the ratios of $10\,000$ different vibrational spectra (for $N=10$), 1000 spectra ($N=100$), and 50 spectra ($N=1000$). For $N=10\,000$, all ten calculated

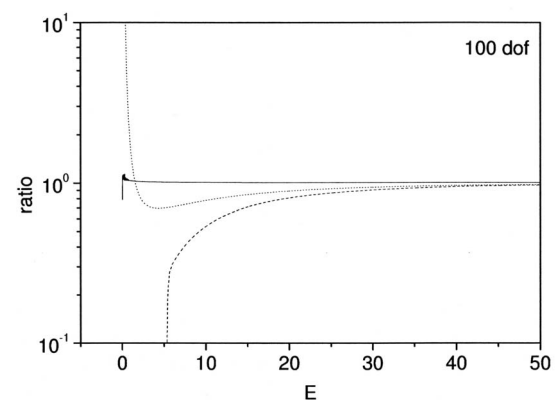


FIG. 2. Comparisons of numerical calculations of level densities with $s=100$ oscillators.

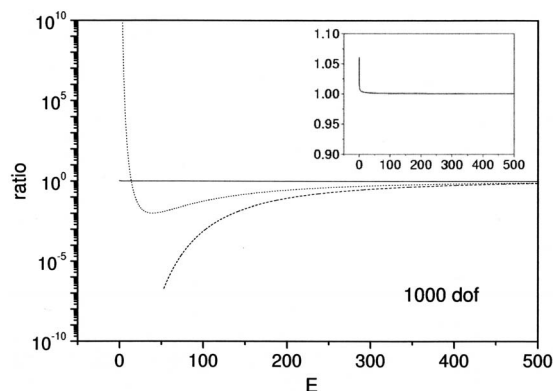


FIG. 3. Comparisons of numerical calculations of level densities with $s = 1000$ oscillators. The inset shows the result from Eq. (16) in an expanded scale.

curves are shown. Similar curves for $N = 20, 50, 200, 500, 2000, 5000$ were also calculated with results that interpolate between those shown. The vibrational frequencies in each simulation were generated randomly with a flat distribution between 0 and 2, after which the spectrum was normalized to give an average of 1.

The energies and frequencies were discretized with an energy unit of 0.01 and individual frequencies are thus multiples of 0.01. Simulations with random frequencies without any discretization were tried but the convergence of the Beyer–Swinehart calculations as a function of the energy resolution proved to be prohibitively slow for systems with a large number of degrees of freedom.

For most sizes, all energies up to one per average oscillator were calculated. The limiting factor relative to computing resources is the Beyer–Swinehart calculation for large sizes, where only a fraction of this upper limit was calculated. The results show a clear tendency for both the Haarhoff and the Witten–Rabinovitch algorithms to fail at low energy. The latter has a built-in low energy cutoff and cannot be used below that, but fails seriously even before this limit is reached. The Haarhoff algorithm fares better but overestimates the level density at very low energies, for $N=1000$

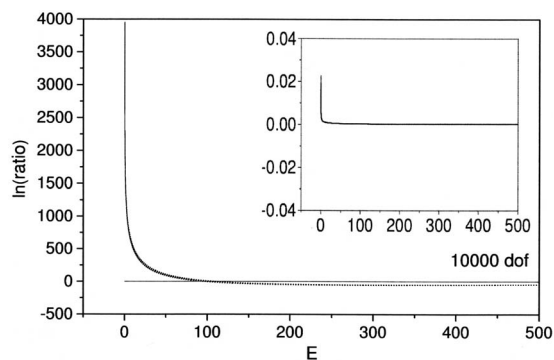


FIG. 4. Comparisons of numerical calculations of level densities with different equations for 10 000 oscillators. The curve calculated from Ref. 3 is absent because the method is undefined for the energy region, shown. For numerical reasons this figure shows the logarithm of the exact-to-approximate ratio, not the ratio itself. The inset shows the result from Eq. (16) in an expanded scale. All ten curves calculated are shown.

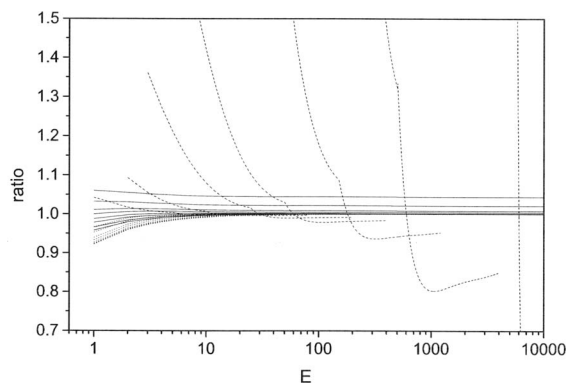


FIG. 5. Comparison of the approximate values for an Einstein spectrum. The curves are for $s = 10, 20, 50, 100, 300, 1000, 10\,000, 10\,0000$. The ratios calculated with Eq. (16) (full lines) decrease with size for a fixed energy, while Haarhoff's result (dotted line) increases slightly. Both converge to unity with size and with excitation energy. The Whitten–Rabinovitch formula (dashed line) shows an increasingly poor performance when size increases (the last curve, for $N=100\,000$, is beyond the limit of the abscissa for this algorithm).

with as much as a factor of 10^{90} . In contrast, the calculations based on Eq. (16) show a surprisingly good agreement with the exact level density.

It is relevant to consider other spectra than those used to generate the curves in Figs. 1–4. Most likely, the hardest test of Eq. (16) is the level density of an Einstein crystal, i.e., one where all vibrational frequencies are degenerate, $\omega_i = \omega$. The level density of this system is discrete with spacing $\hbar\omega$ and $\rho(n\hbar\omega)$ is therefore defined as the number of states with a specific energy $n\hbar\omega$. This level density can be calculated by combinatorics and has been known for a long time. Defining the scaled energy $\varepsilon \equiv E/\hbar\omega$, we have for s oscillators

$$\rho(n) = \frac{(\varepsilon + s - 1)!}{\varepsilon!(s - 1)!}. \quad (23)$$

The canonical partition function is a special case of Eq. (6), from which all relevant quantities required for the application of Eq. (16) can be calculated by standard relations.

Figure 5 shows the comparison between Eqs. (16) and (23). Both of the methods given in Refs. 1 and 2 reproduce the exact result fairly well. There is a minor but clearly noticeable constant deviation of the curves produced with Eq. (16) at high energy, in particular when the number of oscillators is small. This is due to corrections of the type calculated in Eq. (22). Most of the error can be ascribed to the constant term in this calculation.

CONCLUSION AND OUTLOOK

The level densities calculated with a saddle point inversion of the canonical partition function have been shown to be superior to both the Haarhoff and the Whitten–Rabinovitch level densities, except for the case of degenerate oscillators where the Haarhoff formula gives marginally better results. For large molecules, the equations presented here and in Ref. 1 give results that can be used for reliable and fairly efficient calculations of data with even very good experimental precision, whereas the results of Refs. 2 and 3 can be in error by truly astonishing factors.

One of the main difficulties of establishing a relation between energy and density of state in a manageably simple closed form is the inversion of the canonical energy-temperature relation. This problem is not solved here where the relation is found numerically. In an attempt to establish a relation, one may expand the canonical relationship involving sums of Boltzmann factors in powers of $1/T$ in the form $E = sT + \sum_j c_j T^{-j}$. At low energies, this expansion encounters problems due to the fact that the frequencies of all the oscillators enter the coefficients, even those which are effectively quenched because the temperature is too low. In fact, the highest frequencies have the highest weight in such expansion when at the same time they are the least important. This counterproductive procedure can be avoided, as demonstrated in Ref. 9, where the inversion of the infinite spectrum was found to converge at energies as low as that correspond-

ing to the lowest excitation. This venue remains to be explored for the general case of a finite spectrum with an irregular level spacing.

ACKNOWLEDGMENTS

This work has been supported by the Swedish Science Research Council (VR).

- ¹J. U. Andersen, E. Bonderup, and K. Hansen, *J. Chem. Phys.* **114**, 6518 (2001).
- ²P. C. Haarhoff, *Mol. Phys.* **7**, 101 (1963).
- ³G. Z. Whitten and B. S. Rabinovitch, *J. Chem. Phys.* **38**, 2466 (1963).
- ⁴E. Thiele, *J. Chem. Phys.* **39**, 3258 (1963).
- ⁵W. Forst, *Chem. Rev. (Washington, D.C.)* **71**, 339 (1971).
- ⁶H. Wadi and E. Pollak, *J. Chem. Phys.* **110**, 8246 (1999).
- ⁷T. Beyer and D. F. Swinehart, *Commun. ACM* **16**, 379 (1973).
- ⁸M. R. Hoare and Th. W. Ruijgrok, *J. Chem. Phys.* **52**, 113 (1970).
- ⁹K. Hansen, M. D. Johnson, and V. V. Kresin, *Phys. Rev. B* **76**, 235424 (2007).