Springer Series on Atomic, Optical, and Plasma Physics 73

Klavs Hansen

Statistical Physics of Nanoparticles in the Gas Phase

Second Edition

Springer

Springer Series on Atomic, Optical, and Plasma Physics

Volume 73

The Springer Series on Atomic, Optical, and Plasma Physics covers in a comprehensive manner theory and experiment in the entire field of atoms and molecules and their interaction with electromagnetic radiation. Books in the series provide a rich source of new ideas and techniques with wide applications in fields such as chemistry, materials science, astrophysics, surface science, plasma technology, advanced optics, aeronomy, and engineering. Laser physics is a particular connecting theme that has provided much of the continuing impetus for new developments in the field, such as quantum computation and Bose-Einstein condensation. The purpose of the series is to cover the gap between standard undergraduate textbooks and the research literature with emphasis on the fundamental ideas, methods, techniques, and results in the field.

More information about this series at http://www.springer.com/series/411

Klavs Hansen

Statistical Physics of Nanoparticles in the Gas Phase

Second Edition

Springer

Klavs Hansen
School of Science
Tianjin University
Tianjin, China

ISSN 1615-5653 ISSN 2197-6791 (electronic)
Springer Series on Atomic, Optical, and Plasma Physics
ISBN 978-3-319-90061-2 ISBN 978-3-319-90062-9 (eBook)
https://doi.org/10.1007/978-3-319-90062-9

Library of Congress Control Number: 2018944353

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Stefan, Simon, Rasmus and Jakob

Preface to the Second Edition

Several changes have been made to the text in this second edition relative to the first, and no chapter has been left untouched in the process. The changes are of two types. One is an expansion of the scope of the material covered. The other is the introduction of new developments of the field. The revisions have resulted in two new chapters and several new sections. In addition to these changes, uncountably many minor changes have been made in figures and in the text to improve readability. A part of the preparation of this edition was the correction of the errors and (mostly) misprints spotted in the first edition. During this process, it became clear to me that while two errors may cancel each other, mistakes never do. Happily, most mistakes were of minor importance.

The idea behind the book remains unchanged: To provide a guided tour of a number of interesting phenomena in the field. A reader who compares this edition with the first will realize that occasionally the presentation has changed radically. But although the flavor may have changed here and there, hopefully the nutritional value still makes it worth for the reader to consume the dish.

I want to thank Takeshi Furukawa, Bernd von Issendorff, Vitali Zhaunerchyk, and Mathias Weber for providing illustrative experimental data; Petr Slavíček for an educational conversation on holy water at an Erice Workshop; and Hanna Vehkamäki for enlightening me on the first nucleation theorem.

Corrections, suggestions, and error reports are welcome and can be mailed to me at klavshansen@tju.edu.cn.

Tianjin, China — Klavs Hansen

Preface to the First Edition

This book is a developed version of lecture notes that were prepared for graduate students at the University of Gothenburg and Chalmers University of Technology. It aims to fulfill two needs. First of all, it should summarize some of the relevant literature and provide a collection of results for anybody who works with statistical aspects of nanoparticles. Thermal processes are ubiquitous and a proper understanding of the field is necessary for a complete description of the physics of nanoparticles. It is therefore important to have collected a number of results that refer specifically to small particles and the special features they exhibit.

But a mere collection of literature results would limit the usefulness of the book. A textbook which only reviewed literature results would run a serious risk of degenerating into a supermarket for shopping for the equation which seemed to fit some particular data set. The other purpose is therefore to derive results from scratch in a manner that allows the reader to follow the important steps in the derivation and to gain an understanding of the applicability and the limitations of the equations, both from the literature and of the homegrown variety. The term 'homegrown' does not carry any derogatory meaning. As with food, homemade may be worse or better than dishes sold ready-made. And as with food, one usually wants to know the ingredients before the dish is consumed. One important Leitmotif of the derivations presented here is to make it clear to the reader what are physical assumptions and what are mathematical approximations. A calculation of some physical quantity may at the same time be extremely precise, less accurate and completely unreliable. One should not confuse precision in the numerical estimate of the consequences of a model with the accuracy with which the model has been derived. And one should under no circumstances confuse the accuracy of a model calculation with its reliability.

With these two purposes in mind, the material in this book is often presented in more than one way and some of it is redundant from a strictly logical point of view. Occasionally models are elaborated beyond the applicability to a real physical system. The purpose of this is mainly to explore the limits of the approximations and demonstrate the power of the methods. On the other hand, some problems are presented and solved with a degree of simplification which is not on par with the

best available experimental data of specific systems. Readers should be equipped with the necessary tools to improve on the text and to provide their own solutions for specific systems. The text also contains subjects and subsections that are intended to provide examples and illustrations. I trust that the reader will be able to distinguish between fundamental results, applications and illustrative examples.

The book has a strong emphasis on microcanonical physics. Supported particles are very relevant for applications of nanotechnology but fundamental properties of nanosystems are best studied free of the disturbing and frequently uncontrolled influence of a substrate. A large number of experiments have been and more will continue to be performed in molecular beams, ion traps and storage rings, for which the microcanonical description is the appropriate one. But since microcanonical properties can be converted to canonical by a simple integration, also workers who equilibrate their particles to a external heat bath will find useful results here.

Books on statistical physics are full of equations. Remarkably, one can get away with very little advanced mathematics and yet describe a very wide range of observable phenomena. The number of equations is therefore not a measure of the level of mathematical sophistication the student is required to master. They are simply there to show how one gets from point a to point b, and to show what point b looks like. After all, equations are economically expressed figures.

This book also contains material which is covered in most basic courses on statistical physics. Experience has shown that these skills often need to be brushed up and that some recapitulation of subjects is necessary in practice. When the present text falls short of this target, the reader can consult the additional reading list in Appendix A.

I have received a number of suggestions for the contents and corrections to the text that have helped immensely, from O. Echt (Chapters 1, 3, 7), V. V. Kresin (Chapter 12), and from students who have taken my course. But as the sole author of this manuscript I have no other to blame for its shortcomings.

Perfectionism is only the desire to spend time admiring your (almost) completed work, and it is time to publish this volume. Readers will hopefully report suggestions, misprints, miscalculations and plain errors to me at klavs.hansen@-physics.gu.se.

Gothenburg, Sweden Klavs Hansen

Contents

Chapter 1
Introduction

Statistical physics is basically counting. In the words of E. Schrödinger, 'There is, essentially, only one problem in statistical thermodynamics, the distribution of a given amount of energy E over N identical systems. Or perhaps better, to determine the distribution of an assembly of N identical systems over the possible states in which the system can find itself, given that the energy of the assembly is a constant E'.[1] The calculation of the microscopic energy levels per se is not the problem of statistical mechanics. That problem is of course very important but is left to other branches of physics, primarily applied quantum mechanics. This text will therefore not include much material on details of quantum mechanical calculations of level structures. It should be kept in mind, though, that one of the purposes of the application of statistical physics to nanoparticles is to provide experimentally testable models for the behavior of the systems at finite temperatures in order to assess important quantities of the system, such as energy splittings etc.

After having accepted that the word 'Statistical' is relevant, one may still ask why it is necessary to have a special treatment of nanoparticles. Why not just reread the old textbooks in statistical physics? The reason is that the finite number of constituents (atoms, molecules) of the nanometer-sized particles we will be dealing with gives rise to special phenomena that do not appear in macroscopic systems.[2]

Finite temperature is a fact of life for all particle sizes, but it plays a very different role for large and small particles. Generally, the larger the number of degrees of freedom of a particle, the harder it is to prepare in the ground state. In condensed matter, which is often concerned with a practically infinite number of particles, this problem takes the extreme form that it is not possible to reach a temperature of zero kelvin. For small molecules, on the other hand, excitations usually only play a role for the

[1] E. Schrödinger, *Statistical Thermodynamics*, The Syndics of the Cambridge University Press, 1952.

[2] 1 nanometer = 1 nm = 10^{-9} m.

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_1

translational and rotational degrees of freedom because the vibrational frequencies give excitation energies that are well beyond ambient temperatures. Likewise, electronic excitations of molecules are generally not possible at ambient temperature, except for the cases where ground state degeneracies have been lifted and the molecules can be excited around in these few, almost degenerate states. But even for lifted ground state degeneracies are the level splittings frequently much larger than the temperature and strongly inhibit inter-level excitation.

We can make a simple estimate of the particle size that divides these two regimes. The quantum energy of the lowest vibrational frequency of a particle with radius r is on the order of[3]

$$\hbar\omega \sim \hbar\frac{v_s}{r}, \tag{1.1}$$

where v_s is the speed of sound in the particle. Equating this energy to the temperature will give us the typical radius for which the vibrational quantum energies are low enough to be thermally excited. Solving for r we have

$$r \sim \frac{\hbar v_s}{k_\mathrm{B}T} \sim 1 \text{ nm}, \tag{1.2}$$

if we use $v_s = 10^4$ m/s and $T = 100$ K. This radius is indeed not very different from the size of smallish molecules. Thus, for the vibrational degrees of freedom the transition from an effectively zero kelvin atom at ambient conditions to large systems where zero kelvin is practically unattainable occurs in the nanoparticle size range. A similar transition arises for electronic excitations in metallic particles, where the size of the particle determines the spacing between electronic levels. For large particles the spacing is reduced to the extent that thermal electronic excitations become possible.

Another reason small systems need a special treatment is fluctuations. When thermal properties of large particles or large collections of particles are observed, the observed quantities can be identified with the thermal average values. When the number of constituents in the system is small, this connection is not automatic because fluctuations are relatively large and need to be considered explicitly.

The third reason for giving nanoparticles special consideration is that we can do things with these that we cannot do with macroscopic particles, like measure the masses in mass spectrometers and select particles according to the number of net charges they have, down to the ultimate resolution of one unit of charge. In a few cases we can even prepare gas phase samples of all particles composed of exactly the same number of atoms and having exactly the same geometrical and electronic structure, e.g. the all-carbon molecules called fullerenes.

The fourth reason is that particles composed of a small number of monomers (usually an atom but it can also be a small molecule) have properties that depend on

[3]The term 'on the order of', indicated by the symbol $\sim$, is used to describe an approximate relationship between quantities and indicates the leading order term without any attempt to determine multiplicative dimensionless constants. As an example, the volume of a sphere is on the order of its radius cubed.

the precise number of constituents in a non-trivial fashion. This goes for structure and energy alike. There are different ways particles can have size dependent properties. One example is the smooth dependence due to the ratio of surface and interior ('bulk') atoms, which can be extremely different for small and large particles. Another less trivial example is that the binding energies of metal clusters depend strongly on the number of valence electrons in the particle. Yet another is how binding energies and geometric structures of rare gas clusters depend on the number of atoms in the cluster. For these systems, also the thermal properties have a non-trivial size dependence. Quantities that are proportional to the system size in the macroscopic limit are not so for systems that exhibit such size dependencies. For thermal excitation energies, $\overline{E}$, and entropies, S, for example, one in general has that for a given temperature

$$\begin{aligned} \frac{\overline{E}(T)_N}{N} &\neq \frac{\overline{E}(T)_{N+1}}{N+1} \\ \frac{S(T)_N}{N} &\neq \frac{S(T)_{N+1}}{N+1}, \end{aligned} \tag{1.3}$$

for small particles.

The non-equalities in (1.3) violates the principle of corresponding states, which states that the equations of state for a range of different systems, under some conditions which we need not go into detail with here, can be expressed in terms of a universal function, u, as

$$\frac{P}{P_c} = u\left(\frac{T}{T_c}, \frac{v}{v_c}\right), \tag{1.4}$$

where P is the pressure, T the temperature, v the volume per particle, and the same symbols with subscript c are material-specific quantities.

Small is really different.

1.1 Basic Thermodynamic Concepts

Thermodynamics is traditionally concerned with properties of equilibrium systems. The word 'equilibrium' means that macroscopic properties do not change with time. In general, an observable quantity of a thermodynamic system will have an erratic behavior if measured over time, deviating up and down in a random or seemingly random fashion from one instance to the next. If, for example, the pressure of a gas at constant volume and temperature is measured, a sufficiently fine instrument will be able to detect variations with time corresponding to the random time molecules with random momenta hit the measurement instrument.

Determining a mean value consists of disregarding these fluctuations. In practice this is done by averaging a measurement over a time interval which is long compared to the time scale of the fluctuations. The result of this averaging is the mean, or

average, value which, for a generic quantity A, is written as $\langle A\rangle$ or $\overline{A}$. The formal definition of an average of x is

$$\langle A\rangle \equiv \frac{\int f(x)A(x)\mathrm{d}x}{\int f(x)\mathrm{d}x}, \tag{1.5}$$

where $f(x)$ is the distribution over which A is averaged. As an example x may be the momentum, $f(x)$ the distribution of momenta, and A the kinetic energy of a particle. The distribution here is one-dimensional but may be multidimensional with an obvious generalization of the integrals in the definition.

From the definition we see that averaging is a linear operator:

$$\langle \alpha A + \beta B\rangle = \alpha\langle A\rangle + \beta\langle B\rangle. \tag{1.6}$$

By the definition of time averages we have

$$A(t) = \langle A\rangle + \delta A(t), \tag{1.7}$$

where δA is the fluctuating part of A which averages to zero:

$$\langle \delta A(t)\rangle = 0. \tag{1.8}$$

Equilibrium implies that the mean does not change with time:

$$\frac{\mathrm{d}\langle A\rangle}{\mathrm{d}t} = 0 \quad \text{(equilibrium)}. \tag{1.9}$$

You may not be at ease with the separation of time scales implied in the above definition of mean values, i.e. one time scale over which the average is performed and a longer time scale over which potential changes in this mean is measured. One can then instead consider a collection of systems with the same macroscopic characteristics and use the mean over systems in this *ensemble* to obtain a mean value. This is the origin of the word 'ensemble' which we will use below to label systems with certain macroscopic parameters fixed. These macroscopically identical systems will generally be in different microstates, i.e. be in different quantum states. If the system is in equilibrium and is ergodic, averages of this ensemble will be identical to equilibrium time averages. Ergodic systems are those for which every microstate of the system consistent with the external parameters (energy, pressure etc.) is reached in the course of the time development of the system. This very important assumption of equal a priori probabilities will be made implicitly throughout the calculations presented here.

Like averages, fluctuations do not change with time once equilibrium is reached, and by their nature they average to zero. But even if the average of the fluctuating part of A is zero, this is not the case for the average of its square,

$$\langle(\delta A)^2\rangle > 0. \tag{1.10}$$

The standard deviation, or root-mean-square standard deviation, is used as a measure of the fluctuations;

$$\sigma_A = \sqrt{\langle A^2\rangle - \langle A\rangle^2} = \sqrt{\langle(\delta A)^2\rangle}. \tag{1.11}$$

The argument of the square root is known as the variance of A for the distribution. The square root ensures that the dimension of this is the same as of the mean and that these two quantities can be compared directly.

The variance is an example of a *moment of a distribution* (or, to be precise, of the difference between a moment and the square of another). Generally, the n'th moment of A in a distribution is defined as $\langle A^n\rangle$. For normalized distributions, the zeroth moment is unity, the first moment is the average, or mean value, the second is the variance plus the average squared etc. A specific state has a certain distribution of values of A and hence a specific set of moments, some of which may be infinite. If one or more of the moments of two distributions are different, the distributions are also different.

Relative fluctuations depend on the size of the system, and usually the smaller the systems, the more important the fluctuations. If the contribution of a small part of a system to the average of the whole system is independent of the contributions from the rest of the system, averages for the whole system will be the sum of averages from the subsystems:

$$\langle A\rangle = \sum_i \langle A_i\rangle \propto N. \tag{1.12}$$

The fluctuations will also be independent under the same conditions. They add in square:

$$\sigma_A^2 = \langle \delta A^2\rangle = \sum_i \langle \delta A_i^2\rangle \propto N. \tag{1.13}$$

This gives the relative standard deviation

$$\frac{\sigma_A}{\langle A\rangle} \propto \frac{1}{\sqrt{N}}. \tag{1.14}$$

For macroscopic systems one can ignore any corrections from terms proportional to $1/\sqrt{N}$, but obviously this is not so for a small system. A system composed of a large number of small systems (atoms, molecules, nanoparticles) will, however, still have averages and fluctuations that obey this systematics of large numbers.

As an illustration of these concepts, consider an ideal gas. A macroscopic amount of gas at some externally imposed and time independent conditions have vanishingly small fluctuations in observable quantities in most situations. This holds also for non-ideal gases and explains why thermodynamic tables contain measured values of energy and entropy of this and that gas, but not their fluctuations. For a sufficiently small amount of gas where $\sqrt{N} \sim 1$, this picture changes. The ultimate smallness

for a gas is a single molecule. If one measures the momenta of the center of mass motion of this lonely molecule, one will find values that change over time with fairly large relative variations. If enough of these values are measured, the eventual result will be a Maxwell-Boltzmann distribution. This important distribution will be treated in Sect. 2.2. Here we just mention that the fluctuation in the kinetic energy of this single particle distribution is on the order of the mean value, in good agreement with a naive application of the $1/\sqrt{N}$ rule.

Another example of the importance of fluctuations is Brownian motion. This type of motion is observed in microscopes for small particles suspended in a liquid or a gas. Collisions of the molecules in the fluid with the suspended particle cause an incessant, erratic motion of the particle, due to the fluctuating magnitude and direction of the momentum transfer in the particle-molecule collisions. The displacement of the particle with time averages out to zero if many trajectories are followed because there is an equal probability to go left and right, up and down, and back and forth. But there is a non-zero average *square* displacement from the starting point. This non-zero value indicates that the particle diffuses. The degree of diffusion will, everything else equal, increase with decreasing particle size. The quantitative description of this phenomenon was made by Einstein who used measured values of diffusion constants for particles with a well-defined size to calculate the value of Avogadro's number. Avogadro's number, usually denoted by N_A, has the value $6.0 \cdot 10^{23}$. It is the number of molecules in one mole and connects the atomic mass scale with the easier to measure macroscopic mass scale. It also connects the gas constant R and Boltzmann's constant, $R = N_A k_B$. The value of k_B is $1.38 \cdot 10^{-23}$ J/K.[4]

Thermodynamics describes systems in thermal equilibrium in terms of macroscopic and therefore observable variables. These variables can be grouped according to different criteria. One criterion is whether or not they are intensive or extensive. Pressure, chemical potential (which we will meet below) and temperature are examples of intensive quantities: you get the same, not twice the temperature if you double the size of your system with respect to volume, particle number and excitation energy. Energy and entropy are examples of extensive variables; double the size of (a macroscopic) system and you get twice the amount of energy and entropy. It has been suggested that a system is defined as small if these quantities are not extensive. That definition makes self-gravitating systems such as stars small. The long range of the gravitational potential means that two stars of similar sizes that are brought together will have a potential energy which is different from twice the value of a single star. We will use the word 'small' in a more restricted sense, viz. for systems that posses a number of constituents which does not exceed unity by very many orders of magnitude. In most cases this also means that the spatial extension of these particles is small. Given the typical size of an atom of a few Ångström,[5] this gives particle dimensions in the nanometer range.

Instead of a classification of variables as intensive or extensive, they can be classified according to whether or not they are of mechanical origin. Thermal energy and

[4] Some frequently used constants of nature are given in Appendix B.

[5] $1\text{Å} = 10^{-10}$ m.

pressure are mechanical in the sense that they can be understood in terms of Newtonian mechanics. Pressure is the sum of all momentum changes perpendicular to a surface per unit area, divided by the time interval used to measure them, and thermal energy is the sum of internal excitation energies of all the parts of the system, be they one-body energies as in the ideal gas, or described by interactions that involve more particles.

In contrast, temperature and entropy are quantities that are specific to thermodynamics and statistical physics. This is also true for quantities that involve the entropy, such as the chemical potential. They cannot be defined with a simple extension of mechanical concepts. Gibbs' formula for entropy clearly shows its non-Newtonian nature. According to this expression, entropy is a number associated with a probability distribution. If i is a state of the system and P_i the probability that the system is in this state, Gibbs' entropy is

$$S = -k_{\mathrm{B}} \sum_i P_i \ln(P_i), \tag{1.15}$$

where the summation runs over all possible states. Qualitatively, (1.15) says that the entropy is a measure of the lack of knowledge about the system. If the system is in a specific state with a probability approaching unity, the entropy is close to zero which indeed is a result one must require of a reasonable definition of entropy. If the system has Ω equally probably states, $P_i = \Omega^{-1}$ and the entropy becomes

$$S = k_{\mathrm{B}} \ln(\Omega), \tag{1.16}$$

which is Boltzmann's formula and his epitaph.

It can be shown that the entropy in (1.15) is consistent with the definition in terms of energy transfer into the system with $1/T$ as the integrating factor. One can use the definition to derive the Maxwell-Boltzmann distribution for an ideal gas, by requiring that the equilibrium distribution is the one that corresponds to the maximum entropy of the system, consistent with the values of the macroscopic parameters.

There is a limit to how many external parameters one can impose on a system in equilibrium. It is for example not possible to simultaneously impose values of temperature, volume and pressure for a system of, say, $1.602 \cdot 10^{19}$ argon atoms. Only two of those quantities can be imposed externally. The third will be decided by the system itself. On the other hand, there is also a *minimum* number of parameters needed to specify the complete macroscopic state of the system. The maximum and minimum numbers are identical and we can reason our way to this number. Consider for simplicity a system with only one element or chemical compound. If we specify the type and number of particles as well as the volume, we can make a 'Gedanken calculation' of the energy levels of the system, and we have thus implicitly specified the energy levels in the quantum mechanical description of the system. Not that we necessarily know them, but we know they are there. Still, knowing the individual quantum states is not enough to determine the state of the system. The population of the levels is so far undecided, and only become fixed if we specify, say, the

temperature in the canonical ensemble. This is what the temperature does for a living, as we will also see later. Hence we need to specify three external parameters.

The three external parameters need not be volume, particle number and temperature. Other sets can be chosen, but they have to be three in number,[6] irrespective of which ones are chosen. This can be understood as follows: If we specify e.g. volume, particle number and temperature, all other quantities are in principle fluctuating variables, i.e. their variances are non-zero. For a macroscopic system the fluctuations are minute and measured values are in practice identical to the mean values. This means that instead of specifying the particle number, say, in the example above, we may just instead specify the chemical potential. The description is virtually the same. Hence we can replace the specification of one external parameter by another, properly chosen; The number of parameters does not change in the process.

The number 3 for the degrees of freedom for a macroscopic system holds for a single component and single phase system, for example a gas/liquid/solid made up of identical molecules. In general a system may be composed of different molecules which can be in different phases, liquid or gas phase or one of the many different solid phases some compounds are blessed with. Gibbs derived an expression, known as Gibbs' phase rule, which gives the number of macroscopic degrees of freedom, f, for a chemical system with c components and p phases as

$$f = c - p + 2. \tag{1.17}$$

For the simplest case with one phase and one component, this rule states that $f = 2$, contradicting the above statement that $f = 3$. The reason is that the number of molecules is not considered a degree of freedom in (1.17). Equation (1.17) accounts for a system like a test tube with some fixed amount of material and states how many external parameters a chemist can change independently without adding or extracting material from the test tube. Alternatively, one may consider the concentration a single variable, and thus reducing two variable, volume and particle number, to a single parameter. This conforms perhaps better to an intuitive understanding of the problem; If the density is given for bulk material, only the temperature can be varied. This holds to the same extent that the system's properties are independent of its size, which is manifestly *not* the case for nanoparticles, and this is also the reason phase diagrams for pure substances are two-dimensional, usually with pressure and temperature as the variables. If the number of atoms is not given, Gibbs' phase rule should therefore be modified to

$$f = c - p + 3, \tag{1.18}$$

which correctly predicts the number three for our one component-one phase system.

Once we have worked out the interpretation of Gibbs' phase rule for bulk matter, we need to change numbers again for gas phase nanoparticles. Not because of the

[6] 'Then, shalt thou count to three, no more, no less. Three shalt be the number thou shalt count, and the number of the counting shalt be three.' Monty Python and the Holy Grail.

nano part but because of the gas phase part of their nature. As we will see in Chap. 8, the volume into which a particle is embedded affects only its translational degrees of freedom. These decouple exactly from all other degrees of freedom. The decoupling requires that the volume is so large it does not influence the internal state of the particles, which is what is assumed when the adjective 'gas phase' is used here. The number of variables determining the phase of a freely floating particle is therefore reduced to two. One is usually size. The other can be temperature, energy or chemical potential.

There is not unlimited freedom to use any sets of variables as externally imposed parameters. One cannot specify, for example, both the temperature and the energy of a system to simultaneously have fixed, non-fluctuating values. These two quantities are called conjugate variables. Pressure and volume is another pair. We will not go into detail with the application of these rules. You will recognize pairs of so-called conjugate variables when confronted with the recipe for calculating quantities in specific ensembles.

The external parameters needed to specify the thermodynamic system are called its degrees of freedom (d.o.f.). The term refers to *macroscopic* quantities that can be changed independently. It is also used in statistical physics where the term instead refers to *microscopic* quantities that can be varied independently, i.e. coordinates in a generalized meaning, including e.g. electron spin, and sometimes also as the number of these coordinates. Examples are the number of rotational degrees of freedom for a diatomic molecule (d.o.f. $= 2$), the rotational degrees of freedom for a non-linear molecule (d.o.f. $= 3$), the translational degrees of freedom for a free electron (d.o.f. $= 3$), and the spin of an electron (d.o.f. $= 1$). With this use of the word, the number becomes temperature dependent. For example, the temperature of a gas may be so low that molecular vibrations cannot be excited whereas at elevated temperatures vibrationally excited states will have a non-zero population.

1.2 Ensembles

In thermal equilibrium the link between the microscopic and macroscopic properties of a system is provided by the partition function ('Zustandssumme' in German). The partition function, usually denoted by Z East and by Q West of the Atlantic, is a mathematical function of the three macroscopic degrees of freedom. The thermodynamic quantities of a system can be derived from its partition function by simple operations, such as taking derivative etc. The freedom in the choice of the macroscopic degrees of freedom means that there exists a number of different ensembles and therefore different types of partition functions.

The three most frequently used ensembles are the microcanonical ensemble, where particle number, N, volume, V, and energy, E, are specified as external parameters; the canonical ensemble where N, V, T are specified; and the grand canonical

ensemble where V, T, and μ, the chemical potential, are specified. We will mainly be concerned with the microcanonical and the canonical ensembles. The grand canonical ensemble is useful for quantum statistics for reasons that will become clear when we calculate an example.

1.3 The Microcanonical Ensemble

The microcanonical partition function[7] is the number of states of a system with a specified energy, volume and number of constituents (atoms, molecules):

$$Z_{mc}(N, V, E) = \rho(E) \quad \text{(microcanonical)}. \tag{1.19}$$

Although the quantities N and V do not appear on the right hand side of this equation, they are there implicitly. The need to specify the volume in the partition function is not intuitively clear if you work with a free particle, as discussed in connection with Gibbs' phase rule. One usually does not associate a nucleus, an atom or a cluster with a volume the size of a vacuum chamber in a lab. But it is in fact meaningful. If the particle is thermally equilibrated with respect to its translational degrees of freedom, the volume of the vacuum chamber will enter in the translational part of the partition function, although it will have no effect on the intrinsic properties of the particle, which do not change if the size of the vacuum chamber is changed from a lot of empty space to even more empty space.

1.4 The Level Density

The quantity $\rho(E)$ is known as the level density at energy E. In the form $\rho(E)\mathrm{d}E$ it gives the number of states between the energies E and $E + \mathrm{d}E$. It is a very important quantity in statistical physics so some of its general properties will be presented in detail. It is related to the average difference between two levels E_n, E_{n+1}

$$\rho(E) = \frac{1}{\langle E_{i+1} - E_i \rangle}. \tag{1.20}$$

The average is taken over an energy interval broad enough to wash out the rapid state-to-state fluctuations in $E_{i+1} - E_i$, but still small enough for the average to be considered as pertaining to a precise energy. Equation (1.20) is meaningful if the states are closely spaced. At low energies and for small systems, this may not be the situation. Then the level density is better described as a sum of δ functions,

[7] It is tempting but probably wasted effort to suggest to rename this 'the nanocanonical partition function' and 'the nanocanonical ensemble'.

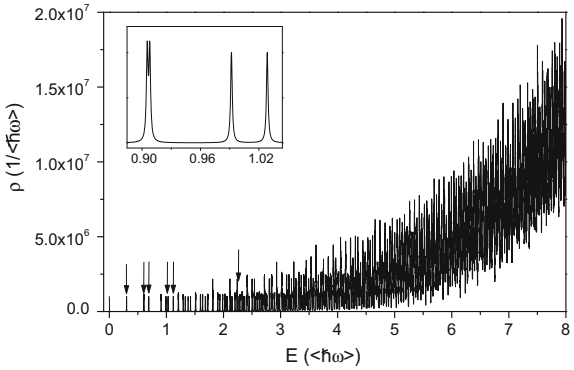


Fig. 1.1 The level density for six harmonic oscillators with an average frequency $\langle\hbar\omega\rangle = 1$. The six fundamental excitations are indicated with arrows. All other peaks are combinations of one or more of these, containing at least two energy quanta. The two lowest frequencies are close to the ratio 1:2, which causes a near degeneracy of overtones. The Lorentzian broadening causes these near degenerate peaks to be slightly higher, as illustrated in the inset. The choice of unit for ρ, $(\hbar\langle\omega\rangle)^{-1}$, is arbitrary. Another choice will give a simple scaling of the ordinate values

$$\rho(E) = \sum_i \delta(E - E_i), \tag{1.21}$$

where E_i is the energy of the state i and the sum runs over all i's. (See Appendix C for some properties of the δ function.)

There is nothing fundamentally different between these two descriptions. They look different just as an airplane seen from the front looks different from one seen from the side. Figure 1.1 shows an example of a level density for a system composed of six harmonic oscillators with randomly selected quantum energies, $\hbar\omega_k$, $k = 1, .., 6$, with an average value of 1. For display purposes the δ functions of infinitely long lived quantum states have been broadened with the normalized Lorentzian shapes,

$$\frac{1}{\pi}\frac{\Gamma}{(\hbar\omega - \hbar\omega_k)^2 + (\Gamma/2)^2} \tag{1.22}$$

with a width of $\Gamma = 0.002\langle\hbar\omega\rangle$. In the limit $\Gamma \to 0$, this line-shape has the property of the δ function. Even for a system this small and for total energies below ten average vibrational quanta is the level spacing very small and the spectrum is well on the way to fuse into a continuum at the high energy end of the figure.

The assumption of equilibrium means that all states of the system with the same total energy E are populated with the same probability. More precisely: if the system's

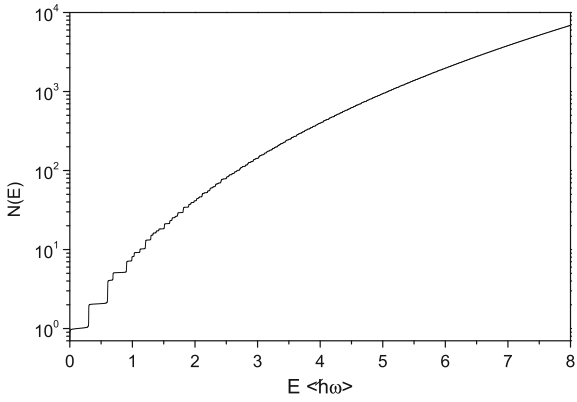


Fig. 1.2 The integrated level density for the six harmonic oscillators in Fig. 1.1

energy is within the small energy interval δE, the probability that the system is measured to be in some state, i, within that interval is

$$P_i = \frac{1}{\rho(E)\delta E}, \tag{1.23}$$

independent of i.

In order to get better acquainted with the concept of level density we will show some important applications. Consider a microcanonical system, i.e. an isolated system with a certain energy and a definite set of state energies. We label the states with the index i and can then calculate the total number of states, N_t, which have energies below some specified value of E as

$$N_t(E) = \sum_i \Theta(E - E_i), \tag{1.24}$$

where the step function $\Theta(x)$, the integral of the δ function, is 1 if $x > 0$, 0 if $x < 0$, and 1/2 for $x = 0$. Figure 1.2 shows the level density of Fig. 1.1 integrated this way. The transition from a step-like function to a function which is effectively smooth and for which the continuum approximation can be used happens already around $E \simeq 2\hbar\langle\omega\rangle$ in this representation. This is by definition also the energy at which the levels are sufficiently dense to convert the summation into an integration over a continuous function. Above this energy we can without any further ado differentiate the result with respect to energy to get the level density $\rho(E)$:

$$\rho(E) = \frac{\mathrm{d}N_t(E)}{\mathrm{d}E}. \tag{1.25}$$

From this somewhat formal definition of ρ we see that the number of states in the energy interval δE around E is equal to $\rho(E)\delta E$, as already mentioned. The level

density has dimension 1/energy, and the value of $\rho(E)$ therefore depends on the units you use.

Occasionally, one uses the term density of states (DOS) synonymously with the level density we have defined here. At other times, however, DOS denotes the density of excitations of a single degree of freedom, as used for the single particle excitation spectrum of electrons in a Fermi gas, for example. The level density is the number of all combinations of singly and multi-electron excitations that give a specific energy. We will make the distinction here and use the expression level density for the density of states of an entire system.

For harmonic oscillators, 'the fruit fly of physics', level densities can be calculated approximately with simple means. The harmonic oscillator is not only simple, it is also a good approximation to the motion of the nuclei in condensed matter (phonons) and molecules (vibrations). The level density of a single harmonic oscillator with quantum energy $\hbar\omega$ is

$$\rho(E) = \sum_{i=0}^{\infty} \delta(E - i\hbar\omega). \tag{1.26}$$

We have set the energy of the quantum ground state to zero so the usual $\hbar\omega/2$ does not appear. It is possible to get rid of the δ functions if one looks at the problem on the scale of $\hbar\omega$. There is precisely one state in this energy interval, and we can then write the smooth version of this level density as $\rho = 1/\hbar\omega$.

Now consider two harmonic oscillators (h.o.) with the same frequency. For an energy which is a multiple of $\hbar\omega$, the total level density, ρ_2, is a convolution of the single h.o. level densities, denoted by ρ_1;

$$\rho_2(E = n\hbar\omega) = \sum_{i=0}^{n} \rho_1(n-i)\rho_1(i) = \sum_{i=0}^{n} 1 \cdot 1 = n + 1. \tag{1.27}$$

Let's repeat the calculation for three oscillators:

$$\rho_3(E = n\hbar\omega) = \sum_{i=0}^{n} \rho_2(n-i)\rho_1(i) = \sum_{i=0}^{n} (n+1-i) \cdot 1 \approx \int_0^n (n+1-x)\mathrm{d}x \approx n^2/2, \tag{1.28}$$

where the sum was approximated with an integral. The situation is depicted in Fig. 1.3. For ρ_4 one gets $\rho_4 \approx n^3/3!$ This can be repeated ad nauseam to give $\rho_f = \rho_{f-1}/(f-1)$ or, for f oscillators and with the energy scale $\hbar\omega$ reintroduced;

$$\rho_f(E) \approx \frac{1}{(f-1)!\,\hbar\omega} \left(\frac{E}{\hbar\omega}\right)^{f-1}. \tag{1.29}$$

This is a crude approximation in absolute terms but it captures the essential features of a better calculation at not too low energies. Equation (1.29) is a poor approximation at low energies, because we approximated a discrete sum with an integral in the

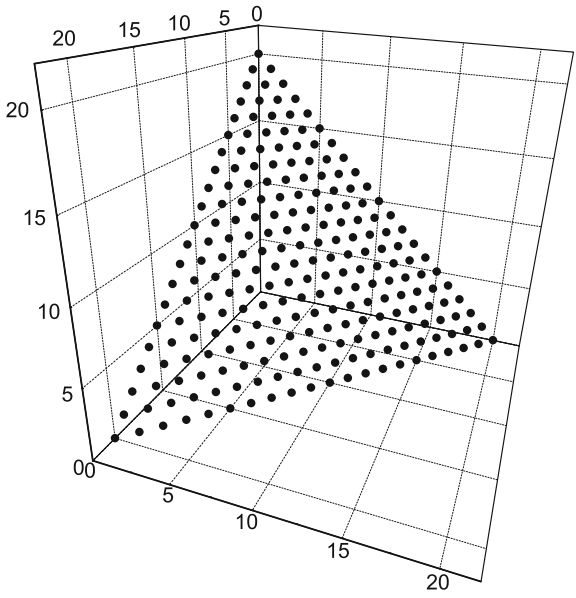


Fig. 1.3 The states for three degenerate harmonic oscillators on the $E = 20\,\hbar\omega$ surface. The axis are the quantum numbers and states are restricted to reside on integer points on a two dimensional surface in the three-dimensional space

derivation, which is a better approximation the higher the energy is. We will return to this problem later and show that one gets a better approximation in the high energy limit, $E/\hbar\omega \gg 1$, by adding $f/2$ to the scaled energy, n,

$$\rho_f(E) \approx \frac{1}{(f-1)!\,(\hbar\omega)^f}\left(E + \frac{f}{2}\,\hbar\omega\right)^{f-1}. \tag{1.30}$$

One important feature of (1.29, 1.30) which survives all improvements of the equation is that level densities grow rapidly with the excitation energy, and the more rapidly the larger the system is.

1.5 Temperature and Boltzmann Factor

Let's now couple two otherwise isolated systems, one of which consists of h.o. and one which we can leave unspecified. By assumption, the total energy is conserved, and we imagine that the two systems can exchange energy freely without destroying the level structure of either system. This is of course an approximation; if the systems couple, there must be coupling matrix elements and they will have the unavoidable consequence that energy levels change. We will ignore these changes because they can be made arbitrarily small if we are prepared to wait long enough for an equilibrium to establish. This is a Gedanken experiment, so we will gladly accept this. Still

assuming ergodicity, every state of the combined system, which is microcanonical, is then occupied with the same probability.

We will first consider the situation when the unspecified system has only two states, like the spin of an electron in a magnetic field, with the energies E_1, $E_1 + \Delta E$. Without loss of generality we can assume that E_1 is equal to zero. The probabilities of the small system being in the two states are

$$P_1 \propto \rho(E), \qquad P_2 \propto \rho(E - \Delta E), \tag{1.31}$$

where ρ is the level density of the large system. The normalized probabilities are

$$P_1 = \frac{\rho(E)}{\rho(E) + \rho(E - \Delta E)}, \qquad P_2 = \frac{\rho(E - \Delta E)}{\rho(E) + \rho(E - \Delta E)}. \tag{1.32}$$

If the energy E of the big system is large enough, in a sense which will be discussed below, we can rewrite the two probabilities with a Taylor-expansion of the logarithm of $\rho(E)$ to first order in ΔE:

$$\ln(\rho(E - \Delta E)) \approx \ln(\rho(E)) - \Delta E \frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E}. \tag{1.33}$$

Let's consider the nature of this logarithmic derivative. The level density is a quantity with dimension, and at first glance the above expansion may seem ill-defined. Logarithms of quantities with dimensions are usually not encouraged. But logarithmic derivatives of dimensionful[8] quantities are actually well-defined, whenever the logarithm itself is well-defined. Formally it can be written as

$$\frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E} \equiv \frac{1}{\rho(E)} \frac{\mathrm{d}\rho(E)}{\mathrm{d}E}. \tag{1.34}$$

Any multiplicative dimension on ρ cancels from this expression. Another way of understanding this is to calculate the number of states in a small interval of width δE around E to get $\rho(E)\delta E$. Then one can take the derivative of the logarithm of this pure number. Since δE is a constant it will drop out from the derivative of $\ln(\rho(E)\delta E)$.

We can express the probabilities in (1.32) in a simpler form if we pretend that the expansion in (1.33) is exact with the terms given:

$$P_1 = \frac{\rho(E)}{\rho(E) + \rho(E - \Delta E)} = \frac{1}{1 + \exp\left(-\Delta E \frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E}\right)} = \frac{1}{1 + \mathrm{e}^{-\beta \Delta E}}, \tag{1.35}$$

[8] Yes, it is a word.

and

$$P_2 = \frac{\exp\left(-\Delta E \frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E}\right)}{1+\exp\left(-\Delta E \frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E}\right)} = \frac{\mathrm{e}^{-\beta\Delta E}}{1+\mathrm{e}^{-\beta\Delta E}}, \tag{1.36}$$

where we have used the notation

$$\frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E} \equiv \beta. \tag{1.37}$$

At this point it is just that, notation. Later it will be clear that the choice of symbol on the right hand side is not accidental. These expressions are beginning to look familiar if you have solved some problems in statistical physics.

We can now generalize the two-level system to a small system with any kind of level density, ρ_s, coupled to a large system. For the sake of argument we will describe both systems with a harmonic oscillator level density, similar to the leading order term in (1.29), although we will take the energy to be a continuous variable. The number of degrees of freedom for the large and the small system are denoted f and s, respectively, and we have $s \ll f$. The probability that the energy is partitioned with $E-\varepsilon$ to the large system and ε to the small system is the product of the numbers of states in the two systems at these energies, $\rho_s(\varepsilon)\delta\varepsilon$ and $\rho_f(E-\varepsilon)\delta\varepsilon$. The total number of states for a given ε is therefore

$$P(E,\varepsilon) \propto \rho_s(\varepsilon)\delta\varepsilon\, \rho_f(E-\varepsilon)\delta\varepsilon. \tag{1.38}$$

The value of $\delta\varepsilon$ can be absorbed into a normalization constant, c, to give

$$P(E,\varepsilon) = c\, \rho_s(\varepsilon)\, \rho_f(E-\varepsilon). \tag{1.39}$$

The distribution of ε depends on E, s and f. When the number of oscillators increases, with the average energy per d.o.f., E/f, kept fixed, it converges to a certain limiting distribution which we will now find.

Let's first show that in that limit the most probable value, ε_m, of ε only depends on E and f through the ratio E/f. The peak of the distribution of ε is found as the derivative of the logarithm of the distribution (the logarithm is a monotonically increasing function of its argument):

$$\frac{\mathrm{d}\ln\left[\rho_s(\varepsilon)\rho_f(E-\varepsilon)\right]}{\mathrm{d}\varepsilon} = 0. \tag{1.40}$$

If we introduce the explicit form of the h.o. level density from (1.29), this gives the peak energy ε_m

$$\varepsilon_m = E\frac{s-1}{f+s-2}. \tag{1.41}$$

Let now f go to infinity, keeping E/f constant. Then the right hand side of (1.41) approaches a constant value which is independent of ε_m. With the specific choice of ρ_s it is equal to $(s-1)E/f$. In the more general case, beyond the harmonic oscillator description, it will still only depend on E/f and s.

Because ε_m/E is small we can expand the distribution (1.39) in ε. One may naively try to perform the expansion directly in the level density of the large system:

$$\rho_f(E-\varepsilon) = \rho_f(E) - \frac{\mathrm{d}\rho_f(E)}{\mathrm{d}E}\varepsilon + \frac{1}{2}\frac{\mathrm{d}^2\rho_f(E)}{\mathrm{d}E^2}\varepsilon^2 + \cdots \tag{1.42}$$

This turns out to be a bad idea. The ratio of the second and first order terms in this expansion is $-f\varepsilon/2E$. With the estimate of $\varepsilon \sim \varepsilon_m$, the ratio of the second to first order term becomes

$$\frac{1}{2}\frac{\mathrm{d}^2\rho_f(E)}{\mathrm{d}E^2}\varepsilon^2 \Big/ \frac{\mathrm{d}\rho_f(E)}{\mathrm{d}E}\varepsilon \sim \frac{s}{2}. \tag{1.43}$$

The second order term is as least as important as the first order term in this expansion. Higher order terms in ε in the expansion will also contribute significantly for similar reasons. Hence it is impossible to calculate the whole ε distribution without including a large number of terms derived from the large system. This is not convenient. Neither is it necessary.

The cause of these problems is the fact that the level density is a rapidly varying function of energy. Instead of an expansion of the function itself, one then expands the logarithm of the function. After expansion and re-exponentiating one has

$$\rho_f(E-\varepsilon) = \rho_f(E)\ \exp\left(-\varepsilon\frac{\mathrm{d}\ln(\rho_f(E))}{\mathrm{d}E} + \frac{\varepsilon^2}{2}\frac{\mathrm{d}^2\ln(\rho_f(E))}{\mathrm{d}E^2} + \cdots\right). \tag{1.44}$$

This makes the terms in the exponential with second and higher order derivatives small. For our example the second order term will be of order $\varepsilon^2 f/E^2$ which goes to zero when $f \to \infty$ in the prescribed way, both on an absolute scale and relative to the first derivative term. Higher order terms will contribute even less. This solves our problems with finding a rapid convergence of the expansion series.

We will identify the logarithmic derivative of the large system's level density with the reciprocal temperature of the system,

$$\frac{1}{k_\mathrm{B}T} \equiv \beta \equiv \frac{\mathrm{d}\ln(\rho_f(E))}{\mathrm{d}E}. \tag{1.45}$$

The temperature is a property of an infinitely large system and has by definition no intrinsic fluctuations.[9] The large system is called a heat bath.

[9] Landau and Lifshitz disagree with this statement and derive results for temperature fluctuations. Our derivation of a canonical temperature renders fluctuations impossible. Heated discussions (pun intended) erupt on this subject in the literature from time to time.

In most situations, k_B will be set equal to unity (see Exercise 1.4). This means that the same units must be used for temperature and energy and that heat capacities and entropies are dimensionless. That should cause no problem. Consider it a reminder that units must be right before plugging numbers into any computer.

The distribution of excitation energies, ε, in the system with level density $\rho(\varepsilon)$ can be written in terms of the temperature as

$$P(T, \varepsilon) \propto \rho(\varepsilon) \mathrm{e}^{-\varepsilon/k_B T}. \tag{1.46}$$

This is the Boltzmann distribution, and the exponential is the Boltzmann factor.

As a check of the definition of temperature, it should reproduce one of the numbered laws in thermodynamics (the zeroth, if you're counting), viz. that two systems in equilibrium with a third will also be in equilibrium with each other. We will shortcut that proof and just show that that two systems brought in contact will equilibrate to the same temperature. This can be done easily and we can in fact also understand the origin of this empirical rule. Consider two heat baths with level densities ρ_1 and ρ_2 in thermal contact. The partitioning of energy in equilibrium will be determined by the product $\rho_1(E_1)\rho_2(E - E_1)$. Finding the maximum of this distribution gives the equation

$$\frac{\mathrm{d}\ln(\rho_1(E_1))}{\mathrm{d}E_1} + \frac{\mathrm{d}\ln(\rho_2(E - E_1))}{\mathrm{d}E_1} = \frac{\mathrm{d}\ln(\rho_1(E_1))}{\mathrm{d}E_1} - \frac{\mathrm{d}\ln(\rho_2(E - E_1))}{\mathrm{d}E} = 0, \tag{1.47}$$

or

$$\frac{1}{T_1} = \frac{1}{T_2}, \tag{1.48}$$

which is the required relation, easily derived.

The concepts of temperature and heat bath were derived here for a collection of harmonic oscillators. One may be worried that the discreteness of the energy transfer from the heat bath to the smaller system makes the temperature definition less than universal. It is possible to let the quantum energy decrease to a sufficiently small value, at least for some degrees of freedom. This is not an imaginary procedure. Solids described by the Debye model of vibrations have arbitrarily small frequencies in the limit of infinite system size. But any sufficiently large system can be used, at least as long as the level density does not have a local maximum as a function of energy. In practice any system large enough will have an infinitely small level spacing and the level density can be considered a continuous function.

What all this shows is that one can derive the Boltzmann factor simply from counting states, since this is what the level density does. One often sees the statement that thermodynamics is a balance between energy, i.e. the Boltzmann factor, and entropy which is effectively the logarithm of the level density, or the number of available states. As you have just seen, this view is a little superficial. Both of these competing factors have their origin in level densities; the entropy from that of the system itself and the Boltzmann factor from the level density of the heat bath.

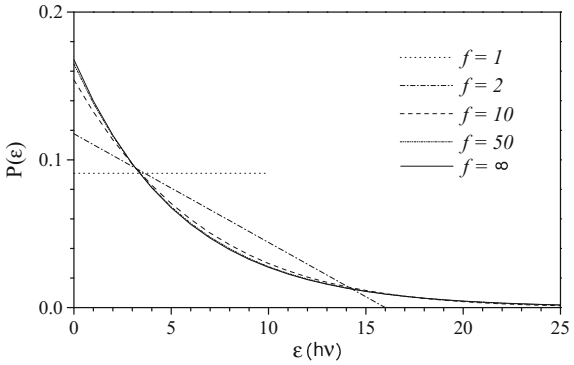


Fig. 1.4 The energy distribution for a single harmonic oscillator in thermal contact with a heat bath of f harmonic oscillators of the same frequency, calculated with (1.39). The total energy per total number of degrees of freedom is taken constant, $E/(f+1) = 5\,h\nu$. The distributions are calculated as $P(\varepsilon) \propto (5(f+1) + f/2 - \varepsilon)^{f-1}$. Something resembling a Boltzmann factor is already beginning to appear for $f = 2$. For $f = 10$ oscillators the curve already bears a strong resemblance to the Boltzmann factor ($f = \infty$), and for $f = 50$ the curves are almost indistinguishable

The way we have derived the Boltzmann factor also explains the paradox of the equal a priori probabilities. The appearance of the Boltzmann factor seems to contradict the assumption that all states are occupied with equal probability. This is surprising since the canonical distribution was derived using this assumption. The inconsistency is only apparent. The explanation is that the Boltzmann factor determines the population of the states in the small system alone, and does not include the heat bath. If one considers the system composed of both the small system and the heat bath, the assumption of equal a priori probabilities still holds.

Figure 1.4 shows how the Boltzmann factor appears when the small system is a single h.o. and the size of the heat bath increases from one h.o. (which is a very small heat bath, indeed) to an infinite number of harmonic oscillators, all with identical level spacings.

1.6 The Canonical Ensemble

After this review of the microcanonical ensemble which included the definition of temperature, we can proceed to define the canonical partition function. It is defined as the sum of states, weighted with the Boltzmann factor:

$$Z_c(N, V, T) = \sum_i \mathrm{e}^{-\beta E_i} \quad \text{(canonical)}, \tag{1.49}$$

where i runs over all distinct quantum states and $\beta \equiv 1/k_B T$. We will often leave out the subscript and arguments that give the type of partition function if these are clear

from the context. The canonical partition function is the normalization constant of the probability density in (1.46);

$$P(T,\varepsilon) = \frac{\rho(\varepsilon)\mathrm{e}^{-\varepsilon/k_\mathrm{B}T}}{\int_0^\infty \rho(\varepsilon)\mathrm{e}^{-\varepsilon/k_\mathrm{B}T}\mathrm{d}\varepsilon} = \frac{\rho(\varepsilon)\mathrm{e}^{-\varepsilon/k_\mathrm{B}T}}{Z_c}. \tag{1.50}$$

In terms of quantum states, the state i is populated with a probability P_i which is proportional to the Boltzmann factor $\exp(-E_i/k_\mathrm{B}T)$ for that state:

$$P_i(T) = \frac{\mathrm{e}^{-\beta E_i}}{Z_c}. \tag{1.51}$$

The sum over all states can be written in a form which may seem formalistic but which will shortly turn out to be useful:

$$\sum_i = \int \left[\sum_i \delta(E - E_i)\right] \mathrm{d}E. \tag{1.52}$$

The integrand $\sum_i \delta(E - E_i)$ is a sum of δ functions and converts the sum over states into an integral over energy. The integrand is nothing but the level density;

$$\sum_i \delta(E - E_i) = \rho(E). \tag{1.53}$$

We can then write the canonical partition function as

$$Z_c(N,V,T) = \int \rho(E)\mathrm{e}^{-\beta E}\mathrm{d}E = \int Z_{mc}(N,V,E)\mathrm{e}^{-\beta E}\mathrm{d}E \tag{1.54}$$

The rewrite of the discrete sum in (1.49) into the continuous integral in (1.54) is valid irrespective of the energy spacing between the quantum states summed over. If the spacing is sufficiently small, the level density is for all purposes a continuous function and the integral in (1.54) should cause no conceptual problem. If the levels have an average spacing which is comparable to or larger than some fraction of the temperature, the level density needs to be described in terms of a discrete spectrum, i.e. a sum of δ functions. In both cases, however, one can use the integral in (1.54).

In mathematical terms (1.54) is the Laplace transform of the level density. The Laplace transform can be inverted, and if you know the partition function for every temperature, you therefore also know the level density. This will be used with an approximate inverse Laplace transform in Chap. 4.

There is an important energy functional associated with the canonical partition function, called Helmholtz' free energy, F. It is defined as

$$Z_c = \sum_i \mathrm{e}^{-\beta E_i} = \mathrm{e}^{-\beta F}. \tag{1.55}$$

Among other things, F is the normalization of the quantum level population probabilities:

$$P_i = \mathrm{e}^{-\beta(E_i - F)}. \tag{1.56}$$

From the definition of F and the partition function as a sum of Boltzmann factors, the word 'energy' in the name 'free energy' seems reasonable: It is an energy which is a function of a certain sum over the energies of all the states of the system. Solving (1.55) for F gives:

$$F = -T \ln\left(\sum_i \mathrm{e}^{-\beta E_i}\right). \tag{1.57}$$

We can understand the role of the summation in more detail if we express it in terms of a characteristic energy of the system. There is only one quantity that qualifies as a characteristic energy, viz. the mean excitation energy, $\overline{E}$, at the specified temperature, and we can write

$$\mathrm{e}^{-\beta F} = \sum_i \mathrm{e}^{-\beta E_i} = g\mathrm{e}^{-\beta \overline{E}}. \tag{1.58}$$

The quantity g is introduced to make the last equality in (1.58) correct. It is the effective number of the thermally populated states that contribute to the sum in the partition function. This number can be understood as the number of states of the system up to an energy which is on the order of $\overline{E}$. The (temperature dependent) number g looks similar to the exponential of the entropy, by Boltzmann's relation between the entropy S and the number of states W

$$S = k_B \ln(W) \tag{1.59}$$

This is exactly what it is. We can see this explicitly by calculating the entropy of the canonical ensemble with Gibbs' entropy formula:

$$\begin{aligned} S = -\sum_i P_i \ln(P_i) &= -\sum_i \frac{\mathrm{e}^{-\beta E_i}}{Z} \ln\left(\frac{\mathrm{e}^{-\beta E_i}}{Z}\right) \\ &= \sum_i \frac{\mathrm{e}^{-\beta E_i}}{Z} (\beta E_i + \ln(Z)) \\ &= \sum_i \frac{\beta E_i \mathrm{e}^{-\beta E_i}}{Z} + \sum_i \frac{\mathrm{e}^{-\beta E_i}}{Z} \ln(Z). \end{aligned} \tag{1.60}$$

The first term is the thermal excitation energy divided by the temperature, $\beta\overline{E}$, and the last term sums up to give $\ln(Z)$. Hence we have

$$S = \beta\overline{E} + \ln(Z) = \beta\overline{E} - \beta F \;\Rightarrow\; F = \overline{E} - TS. \tag{1.61}$$

In other words:

$$Z_c = \mathrm{e}^{-\beta\overline{E}}\mathrm{e}^{S}. \tag{1.62}$$

The value of F depends on the choice of the zero of energy. If the energy of the ground state is set to zero, the free energy is always negative for finite temperature because the partition function is never less than unity. Because $F = \overline{E} - TS < 0$, the entropy term is always numerically greater than the thermal excitation energy,

$$TS > \overline{E}. \tag{1.63}$$

This is consistent with the macroscopic definition of entropy

$$S = \int_0^T \frac{\mathrm{d}E}{T'}, \tag{1.64}$$

which is easy to see when it is rewritten as

$$TS = \int_0^T \frac{T}{T'}\mathrm{d}E > \int_0^T \mathrm{d}E = \overline{E}. \tag{1.65}$$

By the same kind of argument it is possible to show that the free energy varies either quadratically or with a higher power of T for low temperatures (see Exercise 1.10).

In usual terrestrial environments, one does not work in a zero pressure environment, or at least did not when the foundation of thermodynamics was laid down, and this motivated the definition of a variation of Helmholtz free energy, viz. Gibbs free energy;

$$G = F + PV, \tag{1.66}$$

where P is the ambient pressure and V the volume of the system. In parallel, one defines the enthalpy, H, as the energy modified by the pressure-volume term;

$$H = E + PV. \tag{1.67}$$

The difference between results calculated for F and G get their largest contribution from the degrees of freedom associated with the translational motion, and the highest relative difference is seen for a monoatomic gasses. The average internal energy per atom of such a gas is $3T/2$, and the enthalpy is $5T/2$. Disregarding translational motion, however, the difference between enthalpy and energy is usually negligible.

1.7 Mean Values in the Canonical Ensemble

As an example of how to apply the Boltzmann distribution we will calculate an expression for the mean energy, $\overline{E}$:

$$\overline{E} = \frac{\int_0^\infty E\rho(E)\mathrm{e}^{-E/k_\mathrm{B}T}\,\mathrm{d}E}{\int_0^\infty \rho(E)\mathrm{e}^{-E/k_\mathrm{B}T}\,\mathrm{d}E}. \tag{1.68}$$

It is easy to verify that this is equal to

$$\overline{E} = -\frac{1}{Z}\frac{\mathrm{d}}{\mathrm{d}\beta}\left(\int_0^\infty \rho(E)\mathrm{e}^{-E/k_\mathrm{B}T}\,\mathrm{d}E\right) = -\frac{1}{Z}\frac{\mathrm{d}}{\mathrm{d}\beta}Z = -\frac{\mathrm{d}}{\mathrm{d}\beta}\ln(Z). \tag{1.69}$$

The heat capacity can also be calculated in terms of properties of the distribution. We will show most steps here for a full illustration of this type of calculations:

$$\begin{aligned} C_v \equiv \frac{\mathrm{d}\overline{E}}{\mathrm{d}T} = -\beta^2\frac{\mathrm{d}\overline{E}}{\mathrm{d}\beta} = \beta^2\frac{\mathrm{d}}{\mathrm{d}\beta}\left(\frac{1}{Z}\frac{\mathrm{d}Z}{\mathrm{d}\beta}\right) = \beta^2\left(\frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}\beta^2} - \left(\frac{1}{Z}\frac{\mathrm{d}Z}{\mathrm{d}\beta}\right)^2\right) \\ = \beta^2\left(\frac{1}{Z}\int_0^\infty E^2\rho(E)\mathrm{e}^{-\beta E}\mathrm{d}E - \overline{E}^2\right) = \beta^2(\overline{E^2} - \overline{E}^2). \end{aligned} \tag{1.70}$$

Easy to do and we also learn that heat capacities are fluctuations in energy.

Partition functions are effectively summarizing a lot of information about an equilibrium system, but they cannot tell you everything there is to know about it. Quantities beyond energy, entropy and such quantities derived from partition functions by elementary operations may need to be found by calculations of other sums. For the calculation of such sums, one uses the probability distribution in the form of a sum over the quantum states of the system. In principle this proceeds by writing down the value of the observable, call it A, for every state of the system, A_i. Then this is summed over all states with the probability that the system is in the particular state, not unlike the way one calculates expectation values in quantum mechanics;

$$A = \frac{1}{Z}\sum_i A_i\mathrm{e}^{-\beta E_i}. \tag{1.71}$$

An example of the questions one can answer with such a calculation is 'What is the probability that at least half the degrees of freedom are in their quantum mechanical ground state?' Then A_i is 1 if the state i fulfills the criterion and zero otherwise.

Calculations of this type can be performed with an operator known as the density operator. From quantum mechanics we know that all information of a system is contained in its wave function, so if we can write thermal expectation values in terms of wave functions, or more generally in terms of quantum states, we have solved the

problem of characterizing a state in as much detail as possible. This is accomplished with the quantum mechanical expectation values of the operator, summed over every energy eigenstate;

$$A = \frac{1}{Z}\sum_i \langle i|\hat{A}|i\rangle e^{-\beta E_i}, \tag{1.72}$$

where $\hat{A}$ is the operator of the observable A and the brackets indicate the quantum mechanical expectation value of the operator in state i. If we introduce the density operator ρ (not to be confused with the level density), defined as

$$\rho \equiv \sum_i e^{-\beta E_i}|i\rangle\langle i|, \tag{1.73}$$

we can write the partition function as

$$Z = \sum_j \langle j|\rho|j\rangle = \mathrm{Tr}(\rho) = \sum_j e^{-\beta E_j}, \tag{1.74}$$

where Tr is the trace operation which consists of summing the diagonal elements. The thermal average of the expectation values of the operator $\hat{A}$ can be expressed with ρ as

$$\begin{aligned} A &= \frac{1}{Z}\mathrm{Tr}(\hat{A}\rho) = \sum_j \langle j|\hat{A}\rho|j\rangle \\ &= \frac{1}{Z}\sum_j \langle j|\hat{A}\left[\sum_i |i\rangle e^{-\beta E_i}\langle i|\right]|j\rangle \\ &= \frac{1}{Z}\sum_j \langle j|\hat{A}|j\rangle e^{-\beta E_j}, \end{aligned} \tag{1.75}$$

which is (1.72). The choice of orthonormal quantum states, $\langle i|j\rangle = \delta_{i,j}$, was used in the construction of the density operator.

1.8 The Grand Canonical Ensemble

The grand canonical partition function is the sum over canonical partition functions with all possible particle numbers, with the weight of each term in the sum determined by a quantity known as the chemical potential:

$$Z_{gc}(\mu, V, T) = \sum_{N=0}^{\infty} Z_c(N, V, T)e^{\beta\mu N} \quad \text{(grand canonical)}. \tag{1.76}$$

N denotes the number of replicas of the particles/atoms/molecules present in the system and *not* the size of the particle. For some systems these two are identical, for example if N is the number of electrons in a Fermi gas. In another situation they are different, as when N is the number of dimers, X_2, of atom X in a gas.

In (1.76) the chemical potential determines the average particle number, $\overline{N}$. Both the microcanonical and the canonical ensembles have fixed particle numbers by definition, whereas the grand canonical ensemble has a fixed chemical potential. Consequently its conjugate variable, the particle number, is not rigorously conserved. One calculates the mean particle number in the ensemble as

$$\overline{N} = \frac{1}{Z_{gc}} \sum_{N=0}^{\infty} N Z_c \mathrm{e}^{\beta\mu N} = T \frac{\partial}{\partial \mu} \ln\left(Z_{gc}\right). \tag{1.77}$$

Just like temperature is not a property of a (finite) isolated system, but of a heat bath connected to it, so is the chemical potential not a property of the system but of a sufficiently large reservoir of particles (and energy) that can be exchanged with our system. Loosely speaking, the chemical potential is the price one pays to extract a particle from the reservoir into the system. This is analogous to the role of temperature in the canonical ensemble, in which the temperature is the price for the energy transfer from the heat bath to the system. A more negative chemical potential of the reservoir means that a particle will need to be lifted higher to insert it into the system. This will be reflected in a more rapid decrease with N of the exponential factor in (1.76) and consequently a smaller average N.

We can find μ in terms of canonical quantities by considering the value of N for which the sum in (1.76) has an extremum and the second derivative is negative. The value of the second derivative gives the fluctuation of the particle number. The N for which the summand in (1.77) attains its maximum value will be a good approximation to the mean of N if the third and higher derivatives give small contributions over the range of N-values determined by this width of the N-distribution. Under these conditions the mean is determined implicitly as the maximum of the summand of (1.77) or, equivalently, as the maximum of the logarithm of the summand:

$$\beta\mu + \frac{\partial}{\partial N} \ln(Z_c) = 0. \tag{1.78}$$

With the definition of the free energy of the canonical partition function, this translates into

$$\mu = \frac{\partial F_N}{\partial N}. \tag{1.79}$$

where F_N is the free energy of the canonical partition function.

As an example we calculate the chemical potential of a large number of non-interacting, identical and therefore indistinguishable particles. An approximate realization of this situation is a gas of identical molecules (the approximation refers to

the non-interacting aspect). If the canonical partition function for a single particle is denoted by z, the total canonical partition function for N of these systems is

$$Z_c(N) = \frac{z^N}{N!}. \tag{1.80}$$

The power N reflects that each of the particles in the gas contributes the factor z to the total partition function. The factorial denominator is the consequence of the indistinguishability of the particles. The consequences of this indistinguishability will be discussed in more detail in Chap. 2. Evaluating (1.79) with a finite difference instead of a derivative,

$$\frac{\partial \ln Z_c(N)}{\partial N} \approx \ln(Z_c(N)) - \ln(Z_c(N-1)), \tag{1.81}$$

one gets a chemical potential of

$$\mu = -T \ln(z/N). \tag{1.82}$$

As a check that this defines the maximum of the particle number distribution, we calculate the second logarithmic derivative of the partition function and confirm that it is negative;

$$\frac{\partial^2 \ln Z_c(N)}{\partial N^2} = -\frac{1}{N}. \tag{1.83}$$

Hence the sum in (1.76) has a maximum for the value of N given by (1.82). The free energy, on the other hand, has a positive curvature as a function of N, indicating that the free energy is at a minimum. This is a general result; free energies are minimized in equilibrium. The chemical potential is a property of the particle reservoir and the value of the chemical potential calculated in (1.82) is therefore the value this reservoir must have to ensure chemical equilibrium between the system and the reservoir with N particles in the system on average. Conversely, if N does not match the number corresponding to the chemical potential of the reservoir, particles will flow between the two.

The indeterminate particle number can be a serious disadvantage for the use of the grand canonical ensemble in the description of the thermal properties of small particles if the properties of these depend critically on the number of constituents. Addition of one atom to a 10 atom cluster, for example, will often change the structure and energy of the particle significantly. For this reason the grand canonical ensemble should be used with care when the interest centers on size specific features. On the other hand, a fluctuating particle number may represent the physical situation, for example in cases where chemical reactions take place and there is a reservoir of atoms or molecules which can accept or donate these. In these cases the use of the grand canonical ensemble is warranted.

If you believe that the definitions of the partition functions are meaningful, you can begin to understand Schrödinger's remark quoted in the introduction. The simple microcanonical partition function is the mother of the canonical partition function, the grandmother of the grand canonical partition function and is, as you can see from the definition in (1.19), simply a number of states.

Exercises

1.1 Show that the variance of a quantity and of its fluctuating part is identical when averaging over time, i.e. that $\langle A^2\rangle - \langle A\rangle^2 = \langle(\delta A)^2\rangle$.

1.2 The diffusion equation reads

$$D\nabla^2 c = \frac{\partial c}{\partial t}, \tag{1.84}$$

where D is the so-called diffusion constant and ∇ is the Laplace operator,

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{1.85}$$

The diffusion constant has dimension length squared divided by time, which already tells you that diffusing particles move a typical length $\sqrt{Dt}$ during time t. Solve the equation with all particles initially at the origin. Hint: Make the Ansatz that the solution is of the form

$$c \propto t^{\alpha} \mathrm{e}^{-\gamma r^2/t}, \tag{1.86}$$

with $r^2 = x^2 + y^2 + z^2$.

1.3 Consider a particle of mass equal to that of the electron moving in a linear box of length 1 nm. Use (1.20) to find the level density as a function of energy.

1.4 When we set Boltzmann's constant equal to one, the temperature is measured in the same units as the energy. Why can we do that? Show that a unit Boltzmann's constant makes heat capacities dimensionless. A frequently used energy unit is the electronvolt which has the numerical value 0.1602 aJ $= 1.602 \cdot 10^{-19}$ J. What is the conversion from eV to K (how many kelvin is one electronvolt)? Answer the same question for cm^{-1} and kJ/mol. The energy unit wavenumber is defined as the energy of a photon with wavelength 1 cm. It is often used in spectroscopy. The kJ/mol is standard in chemical thermodynamics.

1.5 What is the unit of entropy?

1.6 Calculate the temperature for the $f = \infty$ case in Fig. 1.4.

1.7 Calculate the mean and width of the energy distribution of one h.o. in contact with another system composed of two h.o., of 10 h.o. and of infinitely many h.o, i.e. the mean and width of the curves shown in Fig. 1.4, but for an arbitrary energy. Assume all frequencies are identical, keep the energy per degree of freedom independent of size and use the discrete level density.

1.8 In the early days of hadron physics when people started discovering all sorts of strange and supposedly elementary particles related to neutrons and protons, it was noted by Hagedorn that the number of particles increases exponentially with the mass of the particle, roughly as $\exp(mc^2/200\ \text{MeV})$. This led him to postulate that there would be an absolute highest temperature. Why would he say that and what is the value of this Hagedorn temperature in kelvin?

1.9 What is the second and fourth moment of a Gaussian distribution with variance σ^2 and mean zero,

$$G = \mathrm{e}^{-x^2/2\sigma^2}. \tag{1.87}$$

Some relevant integrals are found in Appendix C.
Use these results to calculate the fourth moment of a Gaussian with variance σ^2 and first moment x_0.

1.10 In this exercise you will show that the free energy of a large canonical system decreases quadratically or with a higher power of T at low T. To do so, assume that the thermal excitation energy of the system can be written as a Taylor series in the temperature. This assumption requires that the system is large, or more precisely that the temperature is on the order of or larger than the lowest excitation energy.

Set the zero order term in the expansion to zero. This just defines the zero of energy. Use the macroscopic definition of the entropy to show that the first order term is also zero, i.e. that the zero temperature heat capacity is zero.

Conclude that the entropy at $T = 0$ is zero and further that the free energy varies quadratically or with a higher power of temperature for low temperatures. Use the definition of the free energy and the properties of the partition function to verify that the variation is indeed a decrease and not an increase.

Finally, show that the condition of a large system is necessary by attempting an expansion of the free energy of a system with a finite gap between the ground state and the first excited state, i.e. with the spectrum $0, E_1, E_2, \ldots$.

1.11 If the Hamiltonian of a system separates into two parts where each coordinate appears in only one of them,

$$H = H_1 + H_2, \tag{1.88}$$

the energies will also be a sum

$$E(1,2) = E_1 + E_2. \tag{1.89}$$

Show that for a canonical ensemble, the partition function of this type of Hamilton function is the product of the two partition functions calculated separately,

$$Z_{1,2} = Z_1 Z_2. \tag{1.90}$$

The combined level density of two sets of degrees of freedom is the convolution of the two level densities. A convolution of f and g is defined as the integral

$$\int_{-\infty}^{\infty} f(x)g(y-x)\mathrm{d}x. \tag{1.91}$$

You have therefore shown that the level density, i.e. a microcanonical partition function, for such a system is a convolution and that it gives rise to a canonical partition function which is a product.

1.12 Given the Maxwell-Boltzmann distribution for the momentum of an ideal gas in three dimensions,

$$f(p)\mathrm{d}p = cp^2 \mathrm{e}^{-\beta p^2/2m}\mathrm{d}p, \tag{1.92}$$

where $p \equiv \sqrt{p_x^2 + p_y^2 + p_z^2}$, find the mean value, $\langle E_k \rangle$, and variance σ_k of the kinetic energy of a single atom. Then use the relations for the average, $\langle E_{k,tot} \rangle$, and variance, $\sigma_{k,tot}^2$, of the sum of kinetic energies for N atoms with the same mass

$$\langle E_{k,tot} \rangle = \sum_{i=1}^{N} E_k \tag{1.93}$$

$$\sigma_{k,tot}^2 = \sum_{i=1}^{N} \sigma_k^2 \tag{1.94}$$

to find the exact values of the relative standard deviation of the kinetic energy as a function of the number of particles in the gas.

1.13 What is the probability that a (quantum mechanical) canonical system is in its ground state, expressed in terms of Z_c?

1.14 Consider a system that has a value of $-T \ln(z_c/N)$ that is less than the chemical potential of a reservoir to which it is connected. Do particles flow into or out of the system?

1.15 What is the chemical potential of photons?

Chapter 2
The Relation Between Classical and Quantum Statistics

The definition of thermal average values in the canonical ensemble as a sum over all states with a Boltzmann factor as the weight, given in (1.71), is rigorous. It is also useful insofar as the levels and their degeneracies can be calculated quantum mechanically. This is possible for a number of Hamiltonians but certainly not for all interesting ones. A classical treatment provides an alternative. A classical description of a system requires that quantum numbers are large compared to unity and is useful because one does not need to know the quantum spectrum of the system to apply it. It is therefore desirable to know precisely how classical and quantal statistical methods are related.

There are two approaches to explain the classical-quantum statistics connection. One can either start from the truth as we have learned it at school, that the statistics of applied quantum mechanics is fundamental and derive classical statistics in the limit of large quantum numbers; the 'bottom-up' approach. Alternatively one can use an approach closer to the historical development; start with classical dynamics and postulate certain rules which will turn out to give the right result; the 'top-down' approach. We will do both. One advantage of the 'bottom-up' approach is that it provide some training in the ensembles that were introduced in Chap. 1. The main but not exclusive virtue of the 'top-down' approach is conceptual; you will be wiser after having understood the procedure.

An essential step toward application of the classical equations of motion and determining the range of validity of the results, is to establish the rule which relates the classical and the quantum mechanical counting of states. We also need to consider the special quantum mechanical effects of the indistinguishability of particles. We will start with the latter.

Quantum statistical mechanics is influenced in an essential manner by the symmetry of the wave function upon exchange of identical particles. It is well known from quantum mechanics that fermions of the same type (half-integer spin particles, electrons being the prime example), are described by a wave function that must be

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_2

antisymmetric on exchange of any two of the fermions. This has the consequence that two identical fermions cannot be in the same single particle quantum state. This rule is absolutely essential for the description of the types of matter where bound state electrons are involved, be it in atoms, molecules and solids, and irrespective of whether these solids are metals or insulators. Basically the Pauli exclusion principle, as the rule is called, prohibits most of the states one would otherwise include in a counting of states. For identical bosons (integer spin particles) an analogous rule applies, except that the wave function must be symmetric. This means that bosons are social particles and will tend to cluster together. The consequences of this is probably best known for the Bose-Einstein condensates which you can read about in press releases from the Nobel committee. But even everyday objects like lasers(!) work precisely because of this social tendency because photons are bosons.

Quantum mechanically, identical particles are indistinguishable. Even in the classical limit, the partition functions therefore need to be divided by the number of possible permutations of the number of particles in the system. For N identical particles this number is $N!$ Hence the canonical partition function of N indistinguishable particles with the single particle partition function z is

$$Z_N = \frac{z^N}{N!}, \tag{2.1}$$

which we have already used in the Introduction. The factor of $N!$ appears only if the particles really *are* indistinguishable, which means that it will be applicable for a gas but not for atoms in a solid, where the atoms can be identified by their position in the lattice. Molecules in a gas will move around freely apart from infrequent collisions with other molecules and possibly the walls of a container, and after some time it becomes impossible to determine the position of a specific molecule, even classically.

As indicated, indistinguishability may be a time dependent question. We can speculate on the time scales. If we perform a Gedanken experiment where we measure the positions of gas molecules at some specific time such that we can tell them apart at the time of measurement, even if they are otherwise indistinguishable, the entropy will be less than before the measurement. The entropy of the measurement device increases during this process, so no violation of the rule that the total entropy of a closed system cannot decrease occurs. For the sake of argument, we assume that the measurement is done such that the wave functions of individual molecules collapse to Gaussian wave packets, and for simplicity consider the subsequent development as a one-dimensional problem. The wave packets spread out as

$$(\Delta x)^2 = (\Delta x(0))^2 + \left(\frac{\Delta p}{m} t\right)^2. \tag{2.2}$$

For a Gaussian wave packet, $\Delta p = \hbar \Delta x(0)/2$ and we can therefore express Δx in terms of $\Delta x(0)$ and time as

$$(\Delta x)^2 = (\Delta x(0))^2 + \left(\frac{\hbar}{2m\,\Delta x(0)}t\right)^2. \tag{2.3}$$

We can find the time it takes for two identical molecules to have a quantum mechanical overlap and become indistinguishable. This time depends on $\Delta x(0)$. The longest value of this time for two particles a distance d apart corresponds to $\Delta x(0) = d/2\sqrt{2}$ and is equal to $t = md^2/8\hbar \sim m/\hbar\rho^{2/3}$, where m is the mass of the particle and ρ the gas density. For an ideal gas at Standard Pressure and Temperature of $P = 1$ bar $= 1.013 \times 10^5\,\mathrm{N/m^2}$, $T = 293$ K, the time is $t \sim (m/1\text{ u}) \times 2 \cdot 10^{-10}$ s. This time is short but not zero.

The example is a little artificial because the width of the wave packet corresponds to a very low kinetic energy. A more realistic estimate is obtained by using the average root-mean-square thermal speed $\langle p^2\rangle^{1/2}/m = \sqrt{2T/m}$. Under the same conditions as before, this gives $t \sim d\sqrt{m/T} = \sqrt{m/T}\rho^{-1/3} = 2 \cdot 10^{-12}$ s for a 1 u particle (the initial width of the wave packet can be ignored in this calculation).

In practise this time dependence causes few problems for gases. For liquids the situation may be different because the motion of atoms is constrained, but not completely hindered as it is in a solid at low temperature. The calculation of the thermal properties including the proper distinguishability factor may then be non-trivial and depends on the specific molecular properties of the liquid state. We will leave this subject for future studies.

A consequence of quantum statistics is that the counting of states is often much more complicated for a system with a fixed number of particles than for a system with a fixed chemical potential. In a grand canonical ensemble, with its fixed chemical potential and fluctuating particle number, the partition function is calculated with a summation over all particle numbers, with the chemical potential as the constant price you pay for the addition of a single particle. In the canonical and microcanonical ensembles you need to restrict the summation over states to those that have precisely the right number of particles. Implementation of this restriction is in general a non-trivial task, and the description of fermionic and bosonic systems is usually very cumbersome in the canonical ensemble and is best done with the grand canonical ensemble. There are exceptions. Bosons such as photons and phonons are cases for which the canonical partition functions can often be calculated analytically. Disregarding those, the feasibility of summing over states with the correct quantum statistics is the single important property that makes the grand canonical ensemble useful and in practice *the* ensemble of choice, even in cases where particle numbers are conserved. The practical, computational advantages of the ensemble can simply outweigh this inconsistency.

The cases where the grand canonical ensemble is particularly useful are those where the system can be described in the independent particle approximation, i.e. where the energies of the system can be written as sums of single particle state energies, and when at the same time the system is a strongly degenerate fermionic or bosonic system. Strongly degenerate means for fermions that the lowest energy single particle states are occupied with a probability which is close to unity, and for bosons that a significant fraction of the particles are in the lowest energy state.

For fermions, the energy of the highest occupied level at zero temperature is called the Fermi energy and is on the order of usual molecular electronic energies, i.e. several eV. For bosons, the highest occupied level at zero temperature is simply the single particle ground state, which we can assign zero energy. For sufficiently low temperatures we therefore have

$$\mu \approx E_f \quad \text{(fermions, low temperature)} \tag{2.4}$$

$$\mu \approx 0 \quad \text{(bosons, low temperature)}. \tag{2.5}$$

2.1 Fermi and Bose Statistics of Independent Particles

Fermionic and bosonic systems of independent particle states are prototype 'bottom-up' situations and are the natural choices to illustrate a concrete application of the grand canonical ensemble and the transition from quantum to classical statistics.

Irrespective of whether a system is strongly degenerate or not, i.e. whether or not the occupation number is small compared to unity or not, the total energy for a specific many-particle state is given in the single particle picture by

$$E = \sum_j n_j \varepsilon_j, \tag{2.6}$$

and the number of particles by

$$N = \sum_j n_j, \tag{2.7}$$

where n_j is the occupation number of state j, i.e. the (integer) number of particles in that single particle state. The canonical partition function is therefore

$$Z_c(N, V, T) = \sum_{(n_1, n_2, \ldots)} \delta_{N, \sum n_j} \, \mathrm{e}^{-\beta \sum_j n_j \varepsilon_j}, \tag{2.8}$$

where the sum runs over all combinations of the n_j's consistent with the permutation symmetry of the particles. The permutation symmetry requires that

$$n_i = 0, 1 \qquad \text{(fermions)} \tag{2.9}$$

$$n_i = 0, 1, \ldots N-1, N \qquad \text{(bosons)} \tag{2.10}$$

The Kronecker delta $\delta_{n,m}$, which is one if $n = m$ and zero otherwise, picks out the configurations with the right total number of particles in (2.8).

The grand canonical partition function given in (1.76) can, with (2.8), be written as

$$Z_{gc} = \sum_{N=0}^{\infty} \sum_{(n_1,n_2,\ldots)} \delta_{N,\sum n_j} \, \mathrm{e}^{-\beta \sum_j n_j \varepsilon_j} \, \mathrm{e}^{\beta \mu N}. \tag{2.11}$$

The limit Note that the N in (2.10) is replaced by ∞. Summation over all values of N cancels the Kronecker delta because

$$\sum_m \delta_{n,m} = 1, \tag{2.12}$$

for any value of n. We can then sum unrestricted over all sets of occupation numbers, apart from the permutational symmetry constraint in (2.9, 2.10). We have:

$$Z_{gc} = \sum_{(n_1,n_2,\ldots)} \mathrm{e}^{-\beta \sum_j n_j \varepsilon_j + \beta \sum_j n_j \mu}. \tag{2.13}$$

The exponential can be factorized into contributions from each level, and one ends with a product of grand canonical partition functions for individual levels:

$$Z_{gc} = \prod_j \left(\sum_{n_j} \mathrm{e}^{-\beta n_j \varepsilon_j + \beta n_j \mu} \right). \tag{2.14}$$

This expression is valid for both fermions and bosons. The difference between the two shows up when one calculates the sums, taking the permitted particle numbers into account,

$$Z_{gc} = \prod_j \left(1 + \mathrm{e}^{-\beta(\varepsilon_j - \mu)}\right) \quad \text{(fermions, independent particles)}, \tag{2.15}$$

and

$$Z_{gc} = \prod_j \left(1 - \mathrm{e}^{-\beta(\varepsilon_j - \mu)}\right)^{-1} \quad \text{(bosons, independent particles)}. \tag{2.16}$$

From the partition functions one finds the populations of the individual quantum states to be

$$p_j = \frac{\mathrm{e}^{-\beta(\varepsilon_j - \mu)}}{1 \pm \mathrm{e}^{-\beta(\varepsilon_j - \mu)}}, \tag{2.17}$$

(+ for fermions, − for bosons). These populations are illustrated in Fig. 2.1 for both types of systems, both with $\overline{N} = 100$ particles and with single particle states that are equidistant in energy, $\varepsilon_j = j\Delta$, where j is a non-negative integer. Also shown are the chemical potentials for temperatures from $T = 5\Delta$ to 50Δ.

The similarity between the bosonic and fermionic chemical potentials shown in Fig. 2.1 is in fact an identity, as will be shown in Chap. 10. Note the very high

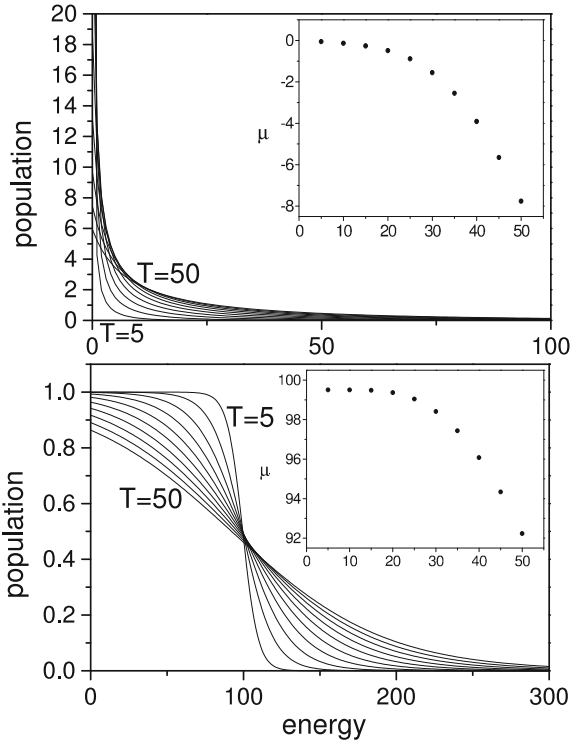


Fig. 2.1 The population in single particle states for $\overline{N} = 100$ bosons (top frame) and fermions (bottom frame) in the grand canonical ensemble for equidistant single particle spectra at temperatures 5–50. All energies and temperatures are in units of the spacing in the single particle spectrum, Δ. The chemical potentials vs. temperature are shown in the insets

temperature needed to cause any change in the chemical potential away from its zero temperature value. This is caused by the large number of particles in the system in combination with the equidistant level spacing. A higher electron number will extend the flat piece even more. See Chap. 10 for details.

The stage is now set to find the classical limit of the thermal properties of the quantum gases. This limit is defined as the situation where the occupation number of each state is much less than unity; $p_j \ll 1$. We can ignore the exponential in the denominator in (2.17) and get

$$\overline{N} = \sum_j \mathrm{e}^{-\beta(\varepsilon_j - \mu)} = \mathrm{e}^{\beta\mu} \sum_j \mathrm{e}^{-\beta\varepsilon_j}. \tag{2.18}$$

The sum is nothing but the canonical partition function for a single particle (identical expressions for a fermion and a boson), $z_{c,1}$, and we find that the chemical potential is

$$\mu = -T \ln\left(z_{c,1}/\overline{N}\right). \tag{2.19}$$

This chemical potential should be compared with the one calculated for a classical gas of indistinguishable molecules, (1.82), in Chap. 1. Apart from the replacement

of a fixed number N of atoms with an average number, $\overline{N}$, the two expressions are identical. Because the chemical potential is the derivative of the canonical partition function, this also mean that the classical limit of the (canonical) partition functions for non-interacting bosons and fermions have the same form as the classical limit written down in Chap. 1, (1.80). It is noteworthy that the indistinguishability factor $N!$ appeared automatically here, without any need to introduce it by hand.[1]

2.2 Classical Phase Space

So let's turn to the second point of this chapter, the 'top-down' approach. In a quantum description of matter you have well-defined energy levels which can be numbered and counted. In classical mechanics this is not the case. An instructive example is (again) the harmonic oscillator. Quantum mechanically it has one more state each time the energy is increased by $\hbar\omega$. Suppose you had not solved the quantum mechanical problem and still had to decide the number of states of the system at a certain energy. What would you do?

The solution is found when we take a closer look at the classical counterpart of the Hilbert space used in quantum mechanics. It is called phase space and is, like the Hilbert space, a multi-dimensional space. Unlike Hilbert space it is not spanned by square integrable functions, but the coordinates and momenta of all the particles in the system. These coordinates and momenta need not be the usual linear quantities you learn about in the first course on mechanics, but can be generalized coordinates and their conjugate momenta, an angle and its associated angular momentum, for example.

A system of N particles will span a $6N$ dimensional space, or more generally $2dN$ dimensions if for some reason the physical space is d dimensional. A state of the system is defined as the point in this $6N$ dimensional space that specifies all the momenta and coordinates. The classical microcanonical partition function is the area of the surface with the prescribed energy, embedded into this space. It does not sound as if it is easy to calculate, and usually it isn't. The canonical partition function is often an easier target. It is calculated as the integral over the whole space with the Boltzmann factor as the weight function:

$$Z \propto \int \mathrm{e}^{-\beta E} \mathrm{d}x_1 \mathrm{d}p_1 \mathrm{d}x_2 \mathrm{d}p_2 \ldots \mathrm{d}x_{3N} \mathrm{d}p_{3N}, \tag{2.20}$$

where the energy E is a function of all coordinates and momenta:

$$E = E(x_1, x_2, \ldots, x_{3N}, p_1, p_2, \ldots, p_{3N}). \tag{2.21}$$

[1] Historical remark: The factor was postulated by Gibbs before quantum mechanics was even an idea, in order to get the correct additivity of the entropy of a classical gas.

As an application of the classical distribution we will calculate a few results for an ideal gas. In an ideal gas the molecules do not interact with each other or anything else, and the Hamiltonian is therefore a sum of the kinetic energies of all molecules, which for simplicity will be assumed to have the same mass m in this example:

$$H = \sum_i \frac{p_i^2}{2m}. \tag{2.22}$$

The Boltzmann factor therefore factorizes, and for every molecule the momentum distribution is

$$P(p_x, p_y, p_z)\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z \propto \mathrm{e}^{-\beta\frac{p_x^2+p_y^2+p_z^2}{2m}}\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z. \tag{2.23}$$

The distribution is spherically symmetric in momentum space. With $p^2 \equiv p_x^2 + p_y^2 + p_z^2$ and integrating out the angular dependence, which just gives a multiplicative constant of 4π, we have

$$P(p)\mathrm{d}p \propto p^2\mathrm{e}^{-\beta\frac{p^2}{2m}}\mathrm{d}p \propto v^2\mathrm{e}^{-\beta m v^2/2}\mathrm{d}v. \tag{2.24}$$

This is the Maxwell-Boltzmann distribution of momenta or speeds of gas molecules in an ideal gas in three dimensions.

The absence of intermolecular interactions is an unnecessary restriction. Any realistic Hamiltonian will contain terms that represent the interaction of the gas molecules with each other. It may even describe a condensed phase where interactions are plentiful. As long as the interaction terms only depend on the positions and not on the momenta/velocities of the molecules, the coordinates can be integrated out independently of the momenta,[2] and consequently the velocity distribution is still given by (2.24) for these situations.

The classical partition function in (2.20) still leaves out the value of the constant of proportionality. A suggestion of what that constant can be is found in the dimensions of Z. If we want to have any correspondence between the classical and the quantum cases, we must at least demand that the dimensions of the two quantities are identical. The quantum partition function is dimensionless, as it is a sum over pure numbers. But the integral in Z in (2.20) has dimension of coordinate times momentum, all to the power $3N$. If you calculate the dimensions of the product of a coordinate and its conjugate momentum, you get the dimension of Planck's constant (recall Heisenberg's indeterminacy relations for position and momentum). This is no accident. The normalization constant is $1/h^{3N}$. In other words: one state of a system with $3N$ sets of conjugate coordinates and momenta has a volume of h^{3N} in phase space. We can define a semiclassical canonical partition function as

[2] In principle. In practice it is not that easy. This is another Gedanken calculation.

$$Z_{semiclass} \equiv \frac{1}{h^{3N}} Z_{class}. \tag{2.25}$$

It should be clear that this prescription only works for systems that actually have a classical description. This rules out the application of (2.25) to spins and similar degrees of freedom.

For many purposes the multiplicative constant difference between the semiclassical expression and the truly classical partition functions is no problem, because any multiplicative constant drops out when calculating observables with logarithmic derivatives or comparing volumes in phase space of similar dimensionality. For calculations of entropies, it does count, however. But even for entropies, the accounting will not be affected if one considers reaction where the d.o.f. is the same before and after the reaction. Only if the net amount of d.o.f changes in the reaction will the pre-integral factor matter. A change in the d.o.f. will occur when one or more excitations is frozen out or thawed during the reaction, which happens by changing the quantum energy of excitations from below to above the temperature or vice versa.

Equation (2.25) will be justified with the calculation of several examples at the end of this chapter, combined with the hope that the perfect agreement with our Ansatz will convince you that all other classically meaningful cases can be treated this way. Before that, we will calculate some general results for the classical limit.

2.3 A Few Elementary and Useful Results from Classical Statistical Mechanics

The classical partition function allows a simple estimate of the high energy/temperature limit of the thermal properties of the systems for which the classical limit exists. A coordinate q_i may appear in quadratic form uncoupled from other coordinates and momenta in the Hamiltonian, H:

$$H = \alpha q_i^2 + H'(q', p), \tag{2.26}$$

where p is the set of all momenta and q' is the set of all coordinates except q_i. The contribution to the canonical partition function from q_i can then be factored out:

$$\begin{aligned} Z &= \frac{1}{h^{3N}} \int \mathrm{e}^{-\beta H} \prod_j \mathrm{d}q_j \mathrm{d}p_j = \frac{1}{h^{3N}} \int \mathrm{e}^{-\beta(\alpha q_i^2 + H')} \mathrm{d}q_i \prod_j{}' \mathrm{d}q_j \mathrm{d}p_j \\ &= \int \frac{1}{h} \mathrm{e}^{-\beta \alpha q_i^2} \mathrm{d}q_i \int \frac{1}{h^{3N-1}} \mathrm{e}^{-\beta H'} \prod_j{}' \mathrm{d}q_j \mathrm{d}p_j, \end{aligned} \tag{2.27}$$

where the primed product of differentials is the one where $\mathrm{d}q_i$ is left out. The last integral is denoted Z' and then

$$Z = Z' \frac{1}{h} \int \mathrm{e}^{-\beta\alpha q_i^2} \mathrm{d}q_i . \tag{2.28}$$

If q_i can take all values, the integration over q_i can be expressed in closed form;

$$Z = \frac{1}{h} \sqrt{\frac{\pi}{\beta\alpha}} Z' . \tag{2.29}$$

It should be clear that this holds for all degrees of freedom that separate the same way q_i did, coordinates and momenta alike. Consequently, the contributions to the partition function from all those degrees of freedom will be proportional to $\sqrt{1/\beta} = \sqrt{T}$, taken to a power which is this number of degrees of freedom. It is equally clear that the factorization does not depend on whether or not the Hamiltonian is quadratic in the degree of freedom or has some other functional form, although the value of the specific integral will.

The average thermal energy of the system in the canonical ensemble is

$$\langle E \rangle = -\frac{\partial \ln(Z)}{\partial \beta} = -\frac{\partial \ln\left(\sqrt{\frac{\pi}{\hbar^2 \beta\alpha}}\right)}{\partial \beta} - \frac{\partial \ln(Z')}{\partial \beta} = \frac{T}{2} + \langle E' \rangle , \tag{2.30}$$

with the above meaning of the primed quantity. Hence the contribution to the thermal energy is $T/2$ from each degree of freedom that enters into the energy quadratically. With this result, the canonical properties of the Maxwell-Boltzmann distribution can practically be read off the Hamiltonian without any further work. The canonical partition function for an ideal gas of N particles is

$$Z \propto \beta^{-3N/2} = T^{3N/2} , \tag{2.31}$$

and the average kinetic energy is

$$\langle E \rangle = -\frac{\partial \ln(Z)}{\partial \beta} = \frac{3N}{2} T . \tag{2.32}$$

The 3 appears because of the number of independent coordinates/momenta in space. The heat capacity from these types of degrees of freedom is also easy to find. For one d.o.f. it is simply $C = {}^1\!/_2$ ($= k_\mathrm{B}/2$), and it is additive, like the energy. Note that the value of α does not appear in either of these quantities.

These rules, known as equipartition, can be very useful because they quickly give values for the classical thermal energies and heat capacities. One use is to judge, from experimental data, whether certain d.o.f.'s are classical or not. Historically, equipartition was a problem for statistical mechanics because the heat capacity of electrons in metals was observed not to obey this simple law. Not knowing the Pauli principle, it was very difficult to explain their anomalously low heat capacity.

Another simple observation can be very useful. If the level density can be written as a power of the excitation energy, the partition function is easily calculated to be:

$$Z \propto \int_0^\infty E^{s-1} \mathrm{e}^{-\beta E} \mathrm{d}E \propto \beta^{-s} = T^s. \tag{2.33}$$

The converse also holds; a partition function of the latter form will only arise from a power law level density, with the powers of the two related as s and $s-1$.

2.4 Semiclassical Calculations of Spectra

In special cases it is possible to calculate the spectrum of a system without solving the Schrödinger equation, and instead make use of the semiclassical spectrum. In this section, you will learn (or relearn) about the method, which has the great advantage that it is easily implementable. The downside is that it is not generally applicable. In fact, we will restrict the treatment to one-dimensional potentials and non-degenerate levels.

Semiclassical levels are determined as the energies of the classical orbits that fulfil the condition that the action integral around a closed classical orbit is an integer multiple of Planck's constant h, i.e.

$$\oint p \mathrm{d}x = nh. \tag{2.34}$$

As an aside we recall that de Broglie's relation between momentum and wavelength gives the identical and well known criterion that the classical trajectory should have a length corresponding to an integer number of wavelengths, λ;

$$n = \oint \frac{p}{h} \mathrm{d}x = \oint \lambda^{-1} \mathrm{d}x. \tag{2.35}$$

The momentum p is a function of the coordinate x, given by the potential and the total energy, E;

$$p = \sqrt{2\mu(E - V(x))}, \tag{2.36}$$

and the task is then to evaluate the action integral in (2.34) with this momentum.

To illustrate the idea, consider the particle in a box. The potential is constant and set to zero between 0 and L, and to infinity outside of this region. The quantum number of the state with energy E is

$$n = \frac{2}{h} \int_0^L (2\mu E)^{1/2} \mathrm{d}x = \frac{\sqrt{8\mu E} L}{h}, \tag{2.37}$$

which gives the energy

$$E = \frac{h^2}{8\mu L^2} n^2, \tag{2.38}$$

in complete agreement with an exact and fully quantum mechanical solution. Problem 2.3 asks you to solve similar problems for a couple of other potentials.

2.5 Quantum Corrections to Interatomic Potentials

Above we discussed how to convert the classical canonical partition function to the semiclassical by division with a factor $2\pi\hbar$ to an appropriate power, and to account for the indistinguishability with the factor $1/N!$ In this section we will go a step closer to the quantum limit and calculate quantum corrections to the equations of motion for particles interacting with a two-body potential.

The classical motion of a particle represents it as a point moving on a trajectory. The correction we will calculate here amounts to treating it as a propagating wave packet. The simplest of these are Gaussian. In $\overline{x}$-space;

$$\phi \propto \mathrm{e}^{\alpha(\overline{x}-\overline{x}_0)^2/2}. \tag{2.39}$$

The probability distribution $|\phi|^2$ for this wave function in coordinate space has the width

$$\langle \Delta\overline{x}^2 \rangle = \frac{1}{2\alpha}. \tag{2.40}$$

For the momentum the width is equal to the thermal width:

$$\langle \Delta\overline{p}^2 \rangle = \langle 2mE_k \rangle = 3mT. \tag{2.41}$$

This relation expresses that the particle is not in a pure momentum eigenstate, but rather in a superposition of kinetic energy eigenstates with a width determined by the temperature. Next we use the Heisenberg indeterminacy relations, better known under the slightly more convenient names Heisenberg uncertainty relations, or uncertainty principle, to relate the width of the wave packet to the temperature;

$$\Delta p \Delta x = {}^1\!/_2\, \hbar. \tag{2.42}$$

(The equality holds for a Gaussian wave packet.) This gives the shape parameter, α, of the wave packet in terms of the temperature:

$$\alpha = \frac{6mT}{\hbar^2}. \tag{2.43}$$

Because we are dealing with two-body interatomic potentials, we can replace the mass m with the reduced mass, $\mu \equiv (1/m + 1/m)^{-1}$, in the following.

The next step is to place the wave packet on a potential energy surface, because this is what we are interested in. This introduces other length scales in the problem that may not be compatible with the one determined by T. Specifically, if the temperature is low, the wave packet in (2.39) will be very broad and reach into regions of very high energy. When the presumed wave packet spreads into regions with potential energies that are much higher than $\Delta p^2/2\mu$, it is a sign that the Ansatz Gaussian wave function is a poor description of the physical situation. For a quantitative calculation of where this limitation sets in, we use the example of a one-dimensional harmonic oscillator. We have the criterion:

$$1/2\,\mu\omega^2\,\Delta x^2 = 1/2\,\mu\omega^2 \left(\frac{\hbar}{2\Delta p}\right)^2 \lesssim \frac{\Delta p^2}{2\mu}. \tag{2.44}$$

This requirement means that the potential energy associated with the width of the wave packet is less than the kinetic energy of the same. Using that the kinetic energy is half the total energy for a harmonic oscillator, $\Delta p^2/2\mu = E/2$, we get the condition for using the Gaussian wave packet

$$E \gtrsim 1/2\,\hbar\omega, \tag{2.45}$$

which looks like a reasonable requirement. It tells us that we are dealing with a high temperature approximation of the quantum contribution to the equations of motion. The condition is illustrated in Fig. 2.2.

After having established the limit of the approximation, we proceed with the calculation of the average potential. It is the average over the wave packet:

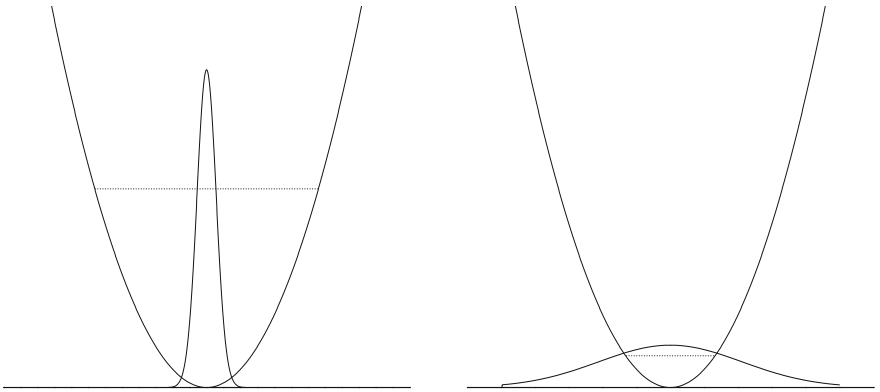


Fig. 2.2 Two situations where the wave packet dynamics can (left) and cannot (right) be used, according to the criterion in (2.44). The potentials are quadratic and the wave packets are Gaussian. The energies calculated from the widths of the wave packets are given as dotted lines

$$V_{qc}(\overline{r}) = \int V(|\overline{r} + \overline{R}|)\,|\phi|^2\,\mathrm{d}\overline{R}, \tag{2.46}$$

where V is the classical potential, and $\overline{r}$ is the relative interatomic coordinates. The integrand is a function of the distance between the two atoms and the integration is over a three-dimensional space. We perform the integral with an expansion in these coordinates to second order in the distance. First $\overline{R}$ is expressed in polar coordinates with the z-axis along the line connecting the atoms,

$$|\overline{r} + \overline{R}| = \left(R^2 \sin^2\theta + (r + R\cos\theta)^2\right)^{1/2} \approx r + \frac{R^2}{2r} + R\cos\theta - \frac{R^2}{r}\cos^2\theta, \tag{2.47}$$

where $r = |\overline{r}|$, $R = |\overline{R}|$, and terms to second order in R relative to the leading order term have been retained. The expansion of the potential along the interatomic axis, retaining terms to the same order, gives;

$$V = V(r) + V'(r)\left(\frac{R^2}{2r} + R\cos\theta - \frac{R^2}{r}\cos^2\theta\right) + {}^1\!/_2\, V'' R^2\cos^2\theta + \cdots \tag{2.48}$$

This gives the effective (quadratic Feynman-Hibbs) potential

$$V_{QFH}(r) \approx \left(\frac{6\mu}{\pi\beta\hbar^2}\right)^{3/2} \int \mathrm{e}^{-\frac{6\mu}{\beta\hbar^2}R^2} \tag{2.49}$$
$$\times\left[V(r) + V'(r)\left(\frac{R^2}{2r} + R\cos\theta - \frac{R^2}{r}\cos^2\theta\right) + {}^1\!/_2\, V'' R^2\cos^2\theta\right]\mathrm{d}\phi\,\mathrm{d}\theta\,\sin\theta\,R^2\,\mathrm{d}\overline{R}.$$

The integrals are standard. We get

$$V_{QFH}(r) = V(r) + \frac{\hbar^2}{24\mu T}\left(V'' + \frac{2V'}{r}\right). \tag{2.50}$$

This is the effective two-body potential in the high energy limit for the classical potential V. Figure 2.3 shows a calculation of the correction to the Lennard-Jones potential at the temperature 14.4 K (= $6\epsilon/50$, using the well depth $\epsilon = 120$ K of argon). The quantum correction term increases when the temperature is lowered. At the already fairly low temperature used in the calculation for the figure (see Problem 2.4), the main effect of the correction is found at small distances where it effectively increases the hard sphere radius by a small amount and leaves the rest of the potential unchanged.

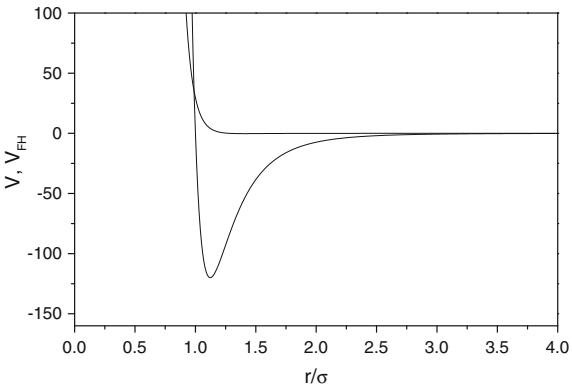


Fig. 2.3 The Lennard-Jones potential and the quadratic Feynman-Hibbs quantum correction for argon. The line that appears to be monotonically decreasing is the correction. It is slightly negative at an interatomic distance just above σ which is 3.4 Å, the characteristic length of the problem. The energy unit is kelvin

2.6 Classical Limits, Example 1: The Harmonic Oscillator

We will now give a few examples of the classical limit of thermal properties of single particle systems with known or traceable quantum mechanical properties. First the harmonic oscillator. The canonical partition function for a quantum harmonic oscillator is one of the simplest to calculate. It is

$$Z_{qm} = \sum_{n=0}^{\infty} \mathrm{e}^{-\beta n\hbar\omega} = \frac{1}{1-\mathrm{e}^{-\beta\hbar\omega}}. \tag{2.51}$$

At high temperatures, $T \gg \hbar\omega$, the average quantum number is large and one approaches the classical limit:

$$Z_{qm} = \frac{1}{1-(1-\beta\hbar\omega+\frac{(\beta\hbar\omega)^2}{2}+\cdots)} = \frac{1}{\beta\hbar\omega(1-\frac{\beta\hbar\omega}{2}+\cdots)} \tag{2.52}$$

$$\approx \frac{T}{\hbar\omega}\left(1+\frac{\hbar\omega}{2T}\right) = \frac{T}{\hbar\omega} + 1/2.$$

The next-to-leading term of $1/2$ is the same $1/2$ that appeared in the improved formula for the level density, (1.30). This will be demonstrated explicitly in Chap. 4.

Let's now calculate the classical partition function. The energy is

$$E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{2.53}$$

Inserting this into (2.20) we have:

$$Z_{class} = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}p\, \mathrm{e}^{-\beta(\frac{p^2}{2m}+1/2m\omega^2x^2)}. \tag{2.54}$$

The exponential factorizes into two parts which depend on x and p alone. We therefore end up with two Gaussian integrals that can be done[3]

$$Z_{class} = \int_{-\infty}^{\infty} \mathrm{e}^{-\beta 1/2 m\omega^2 x^2} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{e}^{-\beta \frac{p^2}{2m}} \mathrm{d}p \tag{2.55}$$

$$= \left[\sqrt{\frac{2}{\beta m}} \frac{1}{\omega} \int_{-\infty}^{\infty} \mathrm{e}^{-x'^2} \mathrm{d}x'\right] \left[\sqrt{\frac{2m}{\beta}} \int_{-\infty}^{\infty} \mathrm{e}^{-p'^2} \mathrm{d}p'\right] = \frac{2\pi T}{\omega}.$$

If you compare this with the leading order in Z_{qm} from (2.52) you see that the missing constant of proportionality is h:

$$Z_{qm} = \frac{1}{h} Z_{class}. \tag{2.56}$$

At least our rule in (2.25) holds for the harmonic oscillator.

We can use semiclassical quantization to understand the origin of this rule. Semiclassically, the energy of a harmonic oscillator is quantized as

$$E = \frac{p^2}{2m} + 1/2\, m\omega^2 x^2 = n\hbar\omega, \tag{2.57}$$

where n is a non-negative integer. This equation defines a curve in phase space which is an ellipse (or a circle, if you prefer, if you redefine the mass or time suitably) on which the semiclassical harmonic oscillator moves. The length of the two axes are determined by solving (2.57) with $p = 0$ and $x = 0$ for the biggest values of the coordinate x and the momentum p with the result

$$x_0 = \left(\frac{2n\hbar}{m\omega}\right)^{1/2}, \quad p_0 = (2nm\hbar\omega)^{1/2}. \tag{2.58}$$

We can find the volume in phase space, V_n, of the lowest $n + 1$ states that have the quantum numbers from 0 to n, as the area of the ellipse:

$$V_n = \pi x_0 p_0 = \pi 2n\hbar\omega = nh. \tag{2.59}$$

The volume in phase space of state n is then $nh - (n - 1)h = h$. It is illustrated in Fig. 2.4.

As the alert reader will have noticed, the ground state is not represented correctly this way. This is a general problem, both with semiclassical quantization and with the

[3] If we are average physicists, that is. As Predrag Cvitanovic remarked, a Gaussian integral is the only integral an average physicist can do (Classics Illustrated: Field Theory, NORDITA Lecture notes, 1983).

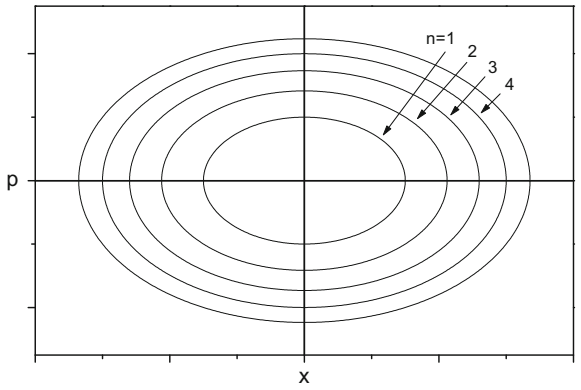


Fig. 2.4 The phase space of a single harmonic oscillator. The states permitted according to semiclassical quantization are drawn as lines. The area between each line is constant and equal to h

classical partition function. At energies comparable to or below the lowest quantum mechanical excitation this is an example of the dangers of using semiclassical results for small quantum numbers.

2.7 Classical Limits, Example 2: A Free Particle

The free particle is particularly important because it is frequently encountered. There are no discrete quantum numbers associated with the translational motion, and we need a trick to calculate the level density (which for a single particle is also the density of states). The trick is to regularize the calculation by using the particle-in-a-box solutions for the energy eigenstates, and then let the size of the box go to infinity.

The energy of a particle in a box is given by

$$E = \frac{\hbar^2 \pi^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2), \tag{2.60}$$

where L is the length of the box and (n_x, n_y, n_z) is a set of positive integers. The equation for the energy looks like the equation for a sphere in the three-dimensional space spanned by n_x, n_y and n_z, with radius

$$r = \frac{\sqrt{E2m}L}{\hbar\pi} = \sqrt{n_x^2 + n_y^2 + n_z^2}. \tag{2.61}$$

The radius and volume of the sphere are dimensionless. To count the number of states with energy up to E, we should count the number of lattice points inside the volume defined by this radius. This is difficult to do exactly and we will use the volume instead, corresponding to a high energy approximation. The relative error

of the result will approach zero when the energy goes to infinity. More correctly, we will use the part of the volume where all coordinates are positive, because negative quantum numbers do not give rise to new states. To avoid such double counting we therefore need to divide the whole volume of the sphere by $2^3 = 8$. Hence the total number of states below E is

$$N_t(E) = \frac{1}{8} V \left(r = \frac{\sqrt{E2m}L}{\hbar\pi} \right) = \frac{1}{8} \frac{4\pi}{3} \left(\frac{\sqrt{E2m}L}{\hbar\pi} \right)^3 . \tag{2.62}$$

The level density for the physical volume of L^3 is therefore

$$\rho_{qm}(E) = \frac{\mathrm{d}N_t(E)}{\mathrm{d}E} = \frac{1}{h^3} m^{3/2} 4\pi E^{1/2} \sqrt{2} L^3. \tag{2.63}$$

Let's now calculate the analogous level density according to the classical-to-quantum recipe. The volume element in phase space is [4]

$$\mathrm{d}V_f = \mathrm{d}\overline{x}\mathrm{d}\overline{p}, \tag{2.64}$$

and the total volume up to energy E is

$$V_f(E) = \int_{V=L^3} \mathrm{d}\overline{x} \int^{p^2/2m \leqslant E} \mathrm{d}\overline{p} = L^3 \int_0^{(2mE)^{1/2}} 4\pi p^2 \mathrm{d}p = L^3 4\pi \frac{(2mE)^{3/2}}{3}. \tag{2.65}$$

The density of classical phase space at energy E is then:

$$\rho_{class}(E) \equiv \frac{\mathrm{d}V_f(E)}{\mathrm{d}E} = L^3 4\sqrt{2}\pi m^{3/2} E^{1/2}. \tag{2.66}$$

We see that for a free particle the quantum mechanical translational level density is the volume in phase space divided by h^3

$$\rho_{qm}(E) = \frac{1}{h^3} \rho_{class}(E). \tag{2.67}$$

Another way of stating this result is to write the density of translational states for a free particle in three dimensions as

$$\mathrm{d}n = \frac{1}{h^3} \mathrm{d}\overline{x}\mathrm{d}\overline{p}. \tag{2.68}$$

[4] A $\overline{quantity}$ is a vector in this section, not an average.

In this example we have calculated the level density and not the partition function. This makes no difference for the conclusions about a multiplicative factor. A constant multiplied on the level density will give the same constant multiplied on the partition function, and vice versa.

2.8 Classical Limits, Example 3: A Particle in the Earth Gravitational Field

The third example is the case of a particle in a constant gravitational field close to earth, bouncing back when it hits the surface;

$$\begin{aligned} V(x) &= mgx, \quad x > 0, \\ V(x) &= \infty, \quad x \leqslant 0. \end{aligned} \tag{2.69}$$

The potential in the other two dimensions is constant, and since these solutions are the well-known plane waves, we can ignore these and concentrate on the vertical direction. The semiclassical partition function for this d.o.f. is

$$Z = \frac{1}{h}\int_0^\infty \mathrm{d}x \int_{-\infty}^\infty \mathrm{e}^{-\beta\left(\frac{p^2}{2m}+mgx\right)}\mathrm{d}p. \tag{2.70}$$

Both integrals in this expression can be done easily, and the result is

$$Z = \frac{1}{h}\sqrt{\frac{2m\pi}{\beta}}\frac{1}{\beta gm}. \tag{2.71}$$

Z is proportional to a power of the temperature and we can therefore apply the results from Sect. 2.3 on the functional form of the level density. Explicitly, if we calculate the partition function as

$$Z = \int_0^\infty \rho(E)\mathrm{e}^{-\beta E}\mathrm{d}E, \tag{2.72}$$

we can set $\rho(E) = cE^{1/2}$, which gives us

$$Z = c\int_0^\infty E^{1/2}\mathrm{e}^{-\beta E}\mathrm{d}E = c\beta^{-3/2}\int_0^\infty x^{1/2}\mathrm{e}^{-x}\mathrm{d}x. \tag{2.73}$$

With the substitution $u = x^{1/2}$ one gets

$$Z = c\beta^{-3/2}2\int_0^\infty u^2\mathrm{e}^{-u^2}\mathrm{d}u. \tag{2.74}$$

The integral can be done easily if we know one trick and how to do a Gaussian integral.[5] The trick will occasionally be useful so here it is:

$$\int_0^\infty x^2 \mathrm{e}^{-x^2} \mathrm{d}x = -\frac{\mathrm{d}}{\mathrm{d}\alpha} \int_0^\infty \mathrm{e}^{-\alpha x^2} \mathrm{d}x \Big|_{\alpha=1} = -\frac{\mathrm{d}}{\mathrm{d}\alpha} 1/2 \sqrt{\frac{\pi}{\alpha}} \Big|_{\alpha=1} = \frac{\sqrt{\pi}}{4}. \tag{2.75}$$

This gives

$$Z = c\beta^{-3/2} \frac{\sqrt{\pi}}{2}. \tag{2.76}$$

We can now identify the two different expressions for Z in (2.71, 2.76) and find that they agree provided

$$c = \frac{2^{3/2}}{hgm^{1/2}}. \tag{2.77}$$

By expressing the partition function in two equivalent ways, we have derived the level density for a particle in a constant gravitational field, complete with numerical constants.

Because the level density is the reciprocal of the average of the spacing between two neighboring quantum mechanical levels (see (1.20)), the calculation therefore allow us to draw conclusions about the quantum energies of the system. This may seem a little surprising at first sight. It should not be, though, because that is what we learn from Niels Bohr's Correspondence Principle. For large quantum numbers the quantum mechanical solution must have observables that approach those of the analogous classical problem.

To show explicitly that this is indeed correct and that the introduction of $1/h$ into (2.70) is justified, we find the quantum eigenstates for the problem by solving the Schrödinger equation. For the gravitational potential it reads

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + gmx\psi = E\psi, \tag{2.78}$$

with $\psi(x) = 0$ for $x \leqslant 0$. Expressing the height in scaled dimensionless units as

$$u = x \left(\frac{2m^2 g}{\hbar^2} \right)^{1/3}, \tag{2.79}$$

one gets the equation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}u^2} - u\psi = \varepsilon\psi, \qquad -\varepsilon \equiv E \frac{2m}{\hbar^2} \left(\frac{\hbar^2}{2m^2 g} \right)^{2/3}. \tag{2.80}$$

[5] Predrag Cvitanovic has remarked...

We have used the same symbol for the wave function although it is a different function of u than of x. The quantization is imposed by the boundary condition $\psi(0) = 0$. Note that the scaled energy is dimensionless like the scaled coordinate. If the zero of the scaled coordinate u is shifted by ε, $v \equiv u - \varepsilon$, the equation reads

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}v^2} - v\psi = 0, \tag{2.81}$$

with the condition that $\psi(-\varepsilon) = 0$. The differential equation is second order and thus has two solutions. These are known as the Airy functions Ai and Bi. Only Ai are relevant solutions here (decreasing for sufficiently large positive values of v). The zeroes are given asymptotically by $(3\pi(4n-1)/8)^{2/3} \approx (3\pi/2)^{2/3} n^{2/3}$, with corrections on the order $1/n$. The energies for large quantum numbers are then

$$E_n = \varepsilon \frac{\hbar^2}{2m} \left(\frac{2m^2 g}{\hbar^2} \right)^{2/3} \approx \left(\frac{3\pi}{2} \right)^{2/3} n^{2/3} \frac{\hbar^2}{2m} \left(\frac{2m^2 g}{\hbar^2} \right)^{2/3} \tag{2.82}$$

The level density is

$$\rho(E) = \frac{1}{\frac{\mathrm{d}E_n}{\mathrm{d}n}} = \frac{(8E)^{1/2}}{ghm^{1/2}}, \tag{2.83}$$

i.e. identical to the above result.

Exercises

2.1 The translational partition function of an ideal gas atom is

$$z_{c,1} = V \frac{\sqrt{8\pi^3}}{(\frac{mT}{\hbar^2})^{3/2}}. \tag{2.84}$$

Use this to calculate the chemical potential for a helium gas at the standard conditions, i.e. for pressure equal to 1 bar and temperature 293 K. Repeat the calculation for the same density but at 1000 K. Pay attention to the sign. Now dilute the helium a factor three. What is the change in chemical potential?

2.2 Calculate and plot the thermal excitation energy of a quantum mechanical harmonic oscillator. Perform a high temperature expansion to get the four highest powers of T. Compare the precision of the expansion numerically to the exact solution as a function of temperature.

2.3 Consider first the harmonic oscillator with a potential $1/2\mu\omega^2 x^2$ with excitation energy E. Use the simplified semiclassical procedure derived to show that this gives the spectrum

$$E = n\hbar\omega, \tag{2.85}$$

or

$$\rho = \frac{\partial n}{\partial E} = \frac{1}{\hbar\omega}. \tag{2.86}$$

Repeat the calculation for the gravitational potential and compare the result with the previously derived level density.

2.4 In the calculation of the Feynman-Hibbs leading order correction to interatomic potentials, the temperature of 14.4 K was used. Is this within the region of validity of the calculation?

2.5 Calculate the Feynman-Hibbs correction for a harmonic oscillator potential.

2.6 Consider the phase space of an N atom particle of which all vibrational degrees of freedom are so highly excited that they can be considered classical. An isolated particle will not sample the whole phase space, because conservation laws restrict the motion. What is the dimension of the subspace on which the particle moves when all these conservation laws are taken into account?

2.7 Which way does the oscillator move in Fig. 2.4.

2.8 Consider the analogue to Fig. 2.4 for a partly hindered rotor. They occur in nature when a group of atoms is bound to the rest of a molecule by a single bond that can be rotated. A typical example is the group CH_3, which can rotate around the bond connecting the carbon atom to wherever it is attached. Draw generic phase space trajectories for this situation for energies above and below the energy needed to pass a high potential energy point. To be specific, consider three identical barriers located equidistantly in the angular coordinate.

2.9 Calculate the volume of phase space in units of h of a gold atom in a 1 cm^3 container with kinetic energy between 0.1 and 1 eV.

2.10 Find the scaling that leads from (2.78) to (2.80). It is not enough to verify that the solution give is correct.

2.11 Consider the level density of a particle in a spherical box. Use dimensional arguments that the level density must have the form

$$\rho(E) = f\left(\frac{E}{\frac{\hbar^2}{mR^2}}\right)\frac{1}{\frac{\hbar^2}{mR^2}}, \tag{2.87}$$

where m is the mass of the particle and R the radius. After this, repeat the determination of the level density in analogy to the calculation for the particle in a gravitational field. Finally, compare with the solution of the full quantum mechanical problem. All relevant mathematical information can be found in Handbook of Mathematical Functions (see Appendix A).

Chapter 3
Microcanonical Temperature

The definition of temperature in Chap. 1 arose out of a need to understand some basic thermodynamics. If a system has the possibility to exchange energy with a heat bath, the Boltzmann factor will be the weight that expresses the decrease in the number of available states for the heat bath when some energy is transferred out of it and into the smaller system. In this chapter we will apply this definition to a small microcanonical system, i.e. without contact with an external heat bath.

It sounds contradictory that one can have a temperature when the system is one of constant energy and when temperatures are only defined for ensembles where energy can be exchanged with a heat bath of practically infinite size. In standard statistical mechanics the problem would be dispensed with by the observation that the choice of ensemble doesn't matter because average values and fixed, externally imposed values will be practically identical due to the extremely small relative fluctuations in a macroscopic system. This is not so for a finite system, as discussed in Chap. 1. Nevertheless, a microcanonical temperature can be defined. It does require some care when handling it, though. Once it is established, however, it will be both a conceptually and practically useful quantity.

3.1 Definition

Consider a system with a level density $\rho(E)$. It is a function of the system's energy, as indicated, and may also depend on other quantities, like angular momentum. For the sake of argument we will ignore these dependencies. This is partly a question of convenience and partly a question of identifying the most relevant quantity of the problem. The functional form of the level density will be left unspecified initially.

The canonical description provides the average excitation energy as a function of the temperature, $\overline{E}(T)$. We will try to establish the inverse relation for a

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_3

microcanonical system, i.e. to derive a temperature from the excitation energy, $T(E)$. This is accomplished with the expansion already encountered in (1.44):

$$\ln(\rho(E-\varepsilon)) = \ln(\rho(E)) - \varepsilon\frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E} + \frac{\varepsilon^2}{2!}\frac{\mathrm{d}^2\ln(\rho(E))}{\mathrm{d}E^2} - \frac{\varepsilon^3}{3!}\frac{\mathrm{d}^3\ln(\rho(E))}{\mathrm{d}E^3} + \cdots \tag{3.1}$$

In analogy with the definition of the canonical temperature, we will define the microcanonical temperature of a system with level density $\rho(E)$ as

$$T \equiv \left(\frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E}\right)^{-1}. \tag{3.2}$$

The meaning of this equation is different, though. The level density now refers to the system itself, and not to a heat bath with which the system is in contact. Here we cannot expect the higher order derivatives of the level density to vanish, because the heat capacity is not arbitrarily large, which is the key to the simplification of the canonical temperature. Expressing the level density with the microcanonical temperature requires at least one higher order term:

$$\rho(E-\varepsilon) = \rho(E)\exp\left(-\varepsilon\frac{1}{T} - \frac{\varepsilon^2}{2!}\frac{1}{C'T^2} - \cdots\right), \tag{3.3}$$

with the microcanonical heat capacity defined as $C'^{-1} \equiv \partial T/\partial E$.

There are things this temperature can do and things it definitely cannot do. It can not be used to calculate thermal fluctuations of the excitation energy and, by the same token, it can also not (definitely not!) be used to calculate thermal excitation energy distributions of the entire system by inserting it into a Boltzmann factor. But it can be used a convenient tool when the ε in (3.1) is small compared with E because then the higher order terms will also be small. Integrals of the type

$$\int_0^E \rho(E-\varepsilon)\mathrm{d}\varepsilon \approx \rho(E)\int_0^\infty \mathrm{e}^{-\varepsilon/T}\mathrm{d}\varepsilon = T\rho(E), \tag{3.4}$$

are then easily calculated.

Figure 3.1 illustrates the use of the microcanonical temperature with an example from nuclear physics. In experiments where nuclei are collided at high energy, the debris can under some conditions be described as originating from a hot soup of nuclear matter. Measuring the right averages experimentally will give you our new friend, the microcanonical temperature. This is what is plotted vs. a scaled collision energy. For the nuclei in the figure, the temperature is on the order of several MeV. At high collision energies the curve clearly deviates from the predicted low energy behavior which is approximately $T \propto \sqrt{E}$ (nucleons are fermions. See Chap. 10 for the explanation of how this dependence arises). The deviation is interpreted as a signature of a nuclear phase transition. Similar curves are found in simulations of

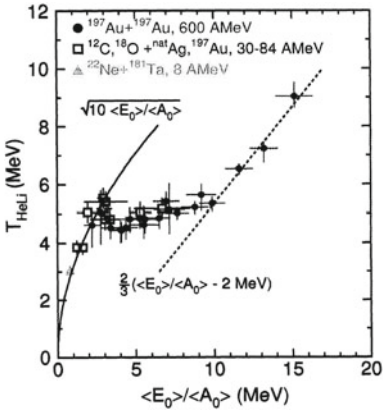


Fig. 3.1 Microcanonical temperatures measured in nuclei as described in the text. A_0 is the number of nucleons in the system. Reprinted from J. Pochodzalla et al., *Phys. Rev. Lett.* **75** (1995) 1040, http://prl.aps.org/abstract/PRL/v75/i6/p1040_1. Copyright (1995) by the American Physical Society

clusters and we will see other examples of this type of curves in Chap. 13 on phase transitions.

To get some quantitative intuition for the concept of a microcanonical temperature, let's calculate it for the approximate level density, which we will have occasion to use later,

$$\rho(E) = a\,(E + E_0)^s\,. \tag{3.5}$$

The three parameters a, E_0, and s represent different physical quantities. s is related to the high temperature canonical heat capacity as discussed in the previous chapter, E_0 is an offset in the relation between temperature and thermal energy, and a is related to the entropy of the system. All three parameters will be discussed in detail in connection with vibrational level densities in Chap. 4. For the present purpose we note that they can be adjusted to reproduce the value, slope and curvature of the level density vs. energy and can therefore be expected to represent a level density over a reasonable broad energy interval.

When s is sufficiently large this level density will act as a normal heat bath. For small s it will be a sort of baby heat bath. From (3.5) and the definition of the microcanonical temperature we find its value;

$$T = \frac{E + E_0}{s}. \tag{3.6}$$

It is worthwhile here to compare with the relation between the corresponding canonical quantities. The mean energy in the canonical ensemble is, in the high temperature limit,

$$E = (s+1)T - E_0 \Rightarrow T = \frac{E + E_0}{s+1}. \tag{3.7}$$

From this result it is clear that $s + 1$ is the canonical heat capacity. Equation (3.7) is the same as the microcanonical value, apart from the fact that the canonical heat

capacity $s+1$ appears in the relation, vs. s in the microcanonical version. The general expansion of the level density in terms of the microcanonical temperature, (3.3), applied to the specific choice of level density in (3.5) becomes

$$\rho(E-\varepsilon) = \rho(E)\exp\left(-\frac{\varepsilon}{T} - \frac{\varepsilon^2}{2sT^2} - \frac{\varepsilon^3}{3s^2T^3} + \cdots\right). \tag{3.8}$$

3.2 Finite Size Heat Bath

When s is not too close to unity one can often use the microcanonical temperature and just keep the first order ε term in the expansion of the level density. This approximation can fail for two reasons. One is that s is small. When $s \sim 1$ the higher order terms in the expansion in (3.1) will not automatically be small, as seen from the example calculation in (3.8). The other possible reason is that ε may not be small. This will be the case if it is an activation energy, representing the energy threshold that needs to be overcome before some specific reaction can happen.

The expression 'activated process' is used for a process that requires so much energy localized on a specific degree of freedom or small set of degrees of freedom that it does not happen immediately, even if the total energy content exceeds the threshold energy for the process. An activated process is characterized by the presence of a ratio of two level densities in the expression for the rate constant that describe the speed with which the reaction proceeds. The activation energy, denoted by E_a, can be very much higher than T, and it is for those situations a special treatment is required. On top of the activation energy, which is a fixed number, an activated process may also consume a small statistical component. The distributions of such energies can also be described with an effective temperature. By necessity, this temperature must be influenced by E_a. In short, we need to repair the equations to answer the questions arising from a large activation energy.

We do this with an expression where the second order term is incorporated as a correction to the temperature:

$$\rho(E-E_a) \approx \rho(E)\exp\left(-\frac{E_a}{T-\frac{E_a}{2C'}}\right). \tag{3.9}$$

The additional term in the denominator, suggested by C. Klots, is the leading order term of a correction to the temperature called the finite heat bath correction, and (3.9) is known as the finite heat bath (FHB) approximation. The physical picture of this is intuitive clear. The exponential is simply the Boltzmann factor, evaluated at a temperature which is the average temperature of the 'before' value, T (no energy extracted), and the 'after' situation, $T - E_a/C'$, where E_a has been removed from the system, leading to a decrease in the microcanonical temperature by the amount $\Delta T = E_a/C'$. What the formula says is that the difference in the logarithm of the

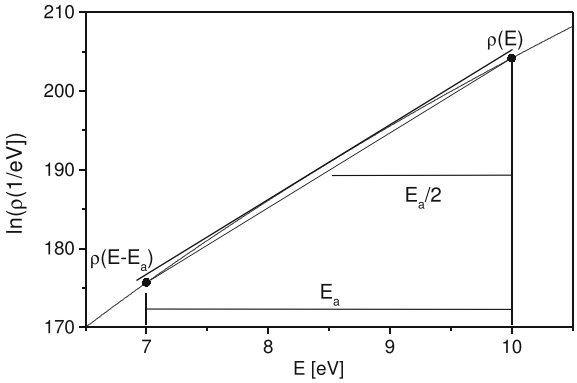


Fig. 3.2 The first order finite heat bath approximation for 100 harmonic oscillators. The vibrational quantum energies are all identical with a value of $\hbar\omega = 0.04$ eV, giving an the energy offset (see (1.30)) of $E_0 = 2$ eV, and $E_a = 3$ eV. The slightly curved line is the logarithm of the level density, calculated in the high energy limit, vs. excitation energy. The FHB approximation is the straight line tangent to the curve at 8.5 eV. The lower straight line connects the ρ's of the energies E and $E - E_a$. These two lines are parallel when the approximation is perfect. From the figure it is clear that the FHB correction improves the estimate significantly compared to the case where the derivatives at either E or $E - E_a$ are used

level densities can be well approximated by the derivative at the midpoint times the interval length E_a.

Figure 3.2 is a sketch of the idea with a numerical example. To appreciate the quality of the approximation, another illustration for the same model system is given in Fig. 3.3. It shows the ratio of the approximate to the exact ratios of level densities (the ratio of the level densities of the cold (energy $E - E_a$) and hot (energy E) systems), evaluated for all possible values of the finite heat bath correction:

$$\exp\left(-\frac{E_a}{T - \frac{E'}{C'}}\right) \frac{1}{\frac{\rho(E-E_a)}{\rho(E)}}, \tag{3.10}$$

vs. E' for $E - E_a < E' < E$. The leading order FHB correction gives a fairly accurate value of the ratio of level densities, even though the correction of the temperature amounts to a significant fraction of the total energy content (compare the ordinate scales of Figs. 3.2 and 3.3).

Up to this point, the finite heat bath approximation has been applied to a single system. This is a somewhat artificial situation because an activated process is bound to tie up some degrees of freedom and the level densities in the numerator and the denominator can not be expected to be identical. A more physical application is therefore to use it to describe the ratio of level densities of two different systems, and the finite heat bath is indeed useful for the manipulation of expressions describing rate constants of activated processes that involve such ratios. We will write a rate

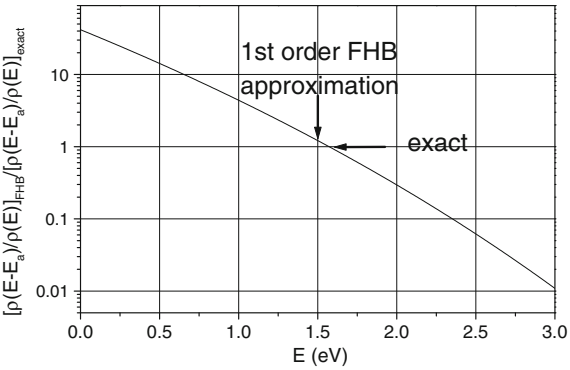


Fig. 3.3 The ratio of the two ratios in (3.10). The data are those used in Fig. 3.2. The ordinate 1 is the perfect approximation. The deviation from unity at the abscissa 1.5 eV which corresponds to the choice $E_a/2$ in (3.9) gives the error in the approximation

constant in a generic form which will be justified in Chap. 5 for unimolecular rate constants:

$$k = f\frac{\rho_p(E - E_a)}{\rho_r(E)}, \tag{3.11}$$

where f is a relatively slowly varying function of energy, E_a is the activation energy, and the subscripts p, r refer to the product particle and the initial state (reactant particle), respectively.

In the canonical ensemble the activation energy E_a enters a Boltzmann factor, $\exp(-E_a/T)$ in the Arrhenius formula

$$k = \omega \mathrm{e}^{-E_a/T}, \tag{3.12}$$

which was discovered by Svante Arrhenius in connection with his work on atmospheric warming by CO_2 more than a century ago. It describes a rate constant in a canonical ensemble. It is a very simple and accurate expression, considering the number of parameters that enter it. It is also very convenient for a number of purposes, because its inverse is easily found. We will therefore try to mimic its behavior by expressing (3.11) in exponential form;

$$k = g\mathrm{e}^{-E_a/T_e}, \tag{3.13}$$

where T_e is a properly chosen temperature.

There are several possible choices for T_e, because a change in the definition can be offset by a simultaneous redefinition of g. Infinitely many, in fact. To see this, we calculate the simultaneous changes in g and T_e required to keep the observable rate constant unchanged. To make the point we will even restrict ourselves to changes in g that can be parametrized as a single dimensionless factor, α, that multiplies k. In other words we require that

$$g\mathrm{e}^{-E_a/T_e} = \alpha g \mathrm{e}^{-E_a/T'_e}. \tag{3.14}$$

This equation is easily solved for T_e' in terms of α and T_e. The solution is

$$T_e' = T_e \frac{1}{1 + \frac{T_e}{E_a} \ln(\alpha)}. \tag{3.15}$$

This gives the T_e' that produces exactly the same rate constant with the pre-exponential factor αg as T_e does with the pre-exponential factor g. With even the best experimental data it is therefore not possible to determine the value of g before T_e has been defined or vice versa, neither from measurements of rate constants nor from an analytical expression for the rate constant. We must therefore consider the definition of the microcanonical temperature for this problem carefully.

One possible definition of T_e was made by C. Klots who defined an equivalent temperature, $T^{\ddagger}$, the isokinetic temperature, as the temperature which gives the same rate constant in a *canonical* ensemble as the *microcanonical* rate constant at the excitation energy E:

$$k(T^{\ddagger}) \equiv k(E). \tag{3.16}$$

This implicitly defines $T^{\ddagger}$ as a function of the energy. The definition is certainly possible. We will use another definition that does not involve canonical quantities.

A fairly accurate expression for rate constants can be obtained with a rewrite of the ratio of level densities in (3.11). We will use the parametrization of the individual level densities of (3.5), with subscripts r and p referring to reactant and product state, respectively. For a reasonable realistic description we must allow that $s_r \neq s_p$, which makes all three constants for both level densities free parameters. With the definitions $\Delta s \equiv s_r - s_p$, $\bar{s} \equiv (s_r + s_p)/2$ and the definition

$$E_a' \equiv E_a + E_r - E_p, \tag{3.17}$$

we get the rate constant

$$\begin{aligned} k(E) &= f \frac{a_p}{a_r} \left[(E + E_p - E_a)(E + E_r)\right]^{-\Delta s/2} \frac{(E + E_p - E_a)^{s_p + \Delta s/2}}{(E + E_r)^{s_r - \Delta s/2}} \\ &\approx f \frac{a_p}{a_r} (E + E_r - E_a'/2)^{-\Delta s} \frac{(E + E_r - E_a')^{\bar{s}}}{(E + E_r)^{\bar{s}}}. \end{aligned} \tag{3.18}$$

Equating the last fraction to a Boltzmann factor yields the effective emission temperature, T_e:

$$\mathrm{e}^{-E_a'/T_e} \equiv \frac{(E + E_r - E_a')^{\bar{s}}}{(E + E_r)^{\bar{s}}}. \tag{3.19}$$

The choice of E_a' as the activation energy is natural in the light of the functional form of the ratio in (3.19). Solving (3.19) for T_e gives

$$T_e = \frac{E + E_r - E_a'/2}{\bar{s}} - \frac{1}{12\bar{s}} \frac{E_a'^2}{E + E_r} + \frac{1}{24\bar{s}} \frac{E_a'^3}{(E + E_r)^2} + \cdots \tag{3.20}$$

The fact that the activation energy appears only in the combination $E_a + E_r - E_p$ means that for all intents and purposes, E_a' is the measured activation energy for the problem. We remind the reader that E_r and E_p (denoted by E_0 in (3.5)) are defined as the offsets in the canonical energy-temperature relation, also known as the caloric curve.

With this temperature the rate constant is calculated to

$$k(E) \approx f \frac{a_p}{a_r} (E + E_r - E_a')^{-\Delta s} \, \mathrm{e}^{-E_a'/T_e}, \tag{3.21}$$

with E_a' defined in (3.17) and T_e defined in (3.20). The factor f, which is only required to be a slowly varying function of energy, can be calculated with results that are derived in Chap. 5.

There are several features here that should be noted in (3.21). One is that the ratio of the a's appear as a pre-exponential, as we expect it to do. The other is that the excitation energy of the product particle, modified by an energy offset, appears in the expression to the small power $s_p - s_r$, in addition to the strong dependence on this energy in the temperature of the exponential. This is also not entirely unsurprising, if for no other reason then because of dimensionality. The last important feature is the fact already mentioned that the apparent activation energy differs from the true, quantum mechanical value by the amount $E_r - E_p$. If the level densities can be calculated from vibrational degrees of freedom in the harmonic approximation, E_r and E_p are the zero point energies of these oscillators, and the apparent, or effective, activation energy E_a' is then the activation energy of the corresponding classical systems. For identical average frequencies in product and reactant, the difference amounts to $E_r - E_p = 3/2\, n\hbar\bar{\omega}$ for loss of an n atom fragment. In general, however, one cannot expect to have such a simple interpretation. The equation remains valid beyond these cases, however.

One of the main lessons of (3.21) is that it provides a microcanonical analogue of the canonical Arrhenius expression. Using the leading order term for T_e gives

$$k \approx \omega \mathrm{e}^{-\frac{E_a'\bar{s}}{E + E_r - E_a'/2}}, \tag{3.22}$$

with the frequency factor from (3.21);

$$\omega \equiv f \frac{a_p}{a_r} (E + E_r - E_a')^{-\Delta s}. \tag{3.23}$$

The fact that the temperature is not usually measured directly and that the denominator of the fraction in the exponential in (3.22) has additional terms makes the interpretation of data in this setting less straightforward than for the canonical

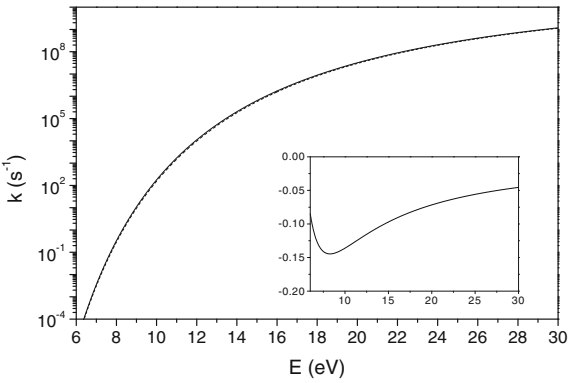


Fig. 3.4 Exact (full line) and approximate (dashed line) rate constants for the system of 100 oscillators, 3 eV activation energy system described in the main text. The inset shows the relative error, $(k_{app} - k_{exact})/k_{exact}$

expression. Setting the terms E_p, E_r to zero in a fitted experimental value will result in an uncontrolled error.

Often the effective temperature and not the energy is the experimentally determined quantity, and the inverse of (3.20) is required. We invert it by successive approximations to next-to-leading order:

$$\begin{aligned} E &\approx T_e\bar{s} - E_r + E'_a/2 + \frac{1}{12}\frac{E'^2_a}{E + E_r} + \cdots \\ &\approx T_e\bar{s} - E_r + E'_a/2 + \frac{1}{12}\frac{E'^2_a}{T_e\bar{s} + E'_a/2}. \end{aligned} \tag{3.24}$$

Equations (3.17, 3.20, 3.21, 3.24) are the main results of this section.

A numerical comparison between exact and approximate rate constants is shown in Fig. 3.4. The system for which the rates are calculated is the same as used in Fig. 3.2 for the precursor. The decay was the loss of a single atom, and the frequency factor f was set to 10^{15} s^{-1}. The two rate constants were thus

$$\begin{aligned} k_{app} &= f\frac{99 \cdot 98 \cdot 97(\hbar\omega)^3}{(E + E_r - E'_a)^3}\mathrm{e}^{-E'_a/T_e} \\ k_{exact} &= f\frac{99 \cdot 98 \cdot 97(\hbar\omega)^3}{(E + E_r)^3}\left(\frac{E + E_p - E_a}{E + E_r}\right)^{97}. \end{aligned} \tag{3.25}$$

The results derived above work in most situations. They do fail for small systems. For those, reactions can take place with excitation energies that are close to threshold. This makes the second and higher order terms in (3.20) important. As will be shown in Chap. 7, decay rates will, under certain conditions, be proportional to the heat capacity of the decaying particles. Measurements of decay rates under these conditions is a commonly applied experimental technique, and it is therefore of importance to have values for the effective heat capacity also for small particles. This is derived in the following.

We will start the derivation by requiring that the heat capacity is constant, if possible; that is, to demand that

$$E = CT_e - E_0, \tag{3.26}$$

and that the Arrhenius expression in (3.21) holds. The definition of ω is retained. In principle the value of the modified E_a' could also be kept free, but it will turn out to be unchanged compared to the previously determined value, so we will make this choice from the start. The heat capacity C and the constant E_0 are both unknown at this point. Once they are determined, the caloric curve in (3.26) determines the temperature, of course.

To avoid a cluttered notation scaled energies will be defined as

$$\begin{aligned} x &\equiv \frac{E + E_r}{E_a'}, \\ x_0 &\equiv \frac{E_0 - E_r}{E_a'}. \end{aligned} \tag{3.27}$$

After taking the $\overline{s}$th root and the logarithm of the Arrhenius expression we have that

$$-\frac{C_v}{\overline{s}}(x + x_0)^{-1} = \ln\left(1 - x^{-1}\right). \tag{3.28}$$

The right hand side is calculated with an expansion to first order in the scaled energy, x;

$$-\frac{\overline{s}}{C}(x + x_0) \approx \frac{1}{\ln\left(1 - x'^{-1}\right)} - \frac{1}{\left(\ln\left(1 - x'^{-1}\right)\right)^2 \left(1 - x'^{-1}\right) x'^2}(x - x'), \tag{3.29}$$

where x' is the expansion point. Choosing an expansion point is equivalent to choosing a rate constant and therefore an experimental time scale. The requirement that the expression is Arrhenius-type means that the left hand side of (3.29) at the expansion point is equal to

$$-\frac{\overline{s}}{C}(x + x_0) = -\frac{\overline{s}}{\ln(\omega/k)} = \frac{1}{\ln\left(1 - x'^{-1}\right)}. \tag{3.30}$$

If we identify k with $1/t$, $\ln(\omega/k)$ is replaced with $\ln(\omega t)$. Due to the large value of the argument of the logarithm, the logarithm can be considered constant over a reasonably large range of times/rate constants. Typically it varies less than 10% if the time varies a factor 10.

From (3.30) we find the terms on the right hand side of (3.29). Reintroducing the scaling parameters and introducing the shorthand $G \equiv \ln(\omega t)$ gives:

$$\frac{E + E_0}{CE_a'} = \frac{1}{G} + \frac{\overline{s}}{G^2} e^{G/\overline{s}} \left(1 - e^{-G/\overline{s}}\right)^2 \left(\frac{E + E_r}{E_a'} - \frac{1}{1 - e^{-G/\overline{s}}}\right). \tag{3.31}$$

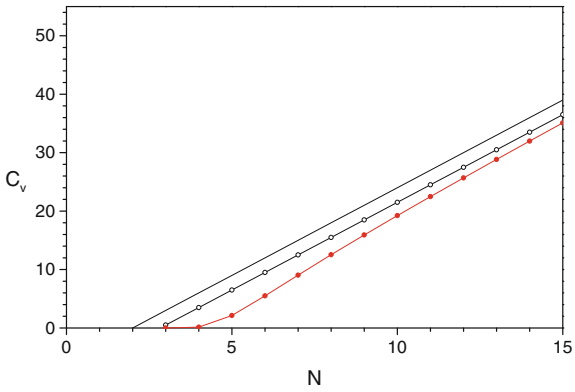


Fig. 3.5 Effective heat capacities vs. particle size for three different approximations; $C = 3N - 6$ (line), $C = 3N - 8.5$ (open circles), (3.32) (red, filled circles)

Identifying the coefficients to different powers of E gives

$$C = \frac{G^2}{\overline{s}} e^{-G/\overline{s}} \left(1 - e^{-G/\overline{s}}\right)^{-2}, \tag{3.32}$$

and

$$\begin{aligned} E_0 &= E_a' \frac{G}{\overline{s}} e^{-G/\overline{s}} \left(1 - e^{-G/\overline{s}}\right)^{-2} - E_a' \left(1 - e^{-G/\overline{s}}\right)^{-1} + E_r \\ &\approx -E_a' \left(\frac{1}{2} + \frac{G}{6\overline{s}}\right) + E_r + \mathcal{O}\left(\left(\frac{G}{\overline{s}}\right)^3\right). \end{aligned} \tag{3.33}$$

The last approximation can be used for $G = 25$ and $\overline{s} > 9$ to an accuracy better than 1%. There is some time dependence in the heat capacity due to the factor $e^{-G/\overline{s}}$. For example, for a dynamic range of a factor of ten in time and the size 6, this implies a reduction of around 20% from that factor. Half of this is compensated by an increase in the factor G.

The effect of the finite size is to reduce the heat capacity, most strongly for small sizes. Figure 3.5 shows the size dependence of the C_v before and after correction. There is a clear cutoff two to three cluster sizes above the one induced by the reduction due to vibrational and rotational motion, i.e. the -6 in $C(\text{uncorrected}) = 3N - 6$. The shift is clearly visible in experimental data.

3.3 Level Densities and Canonical Partition Functions

The microcanonical temperature is useful for easy manipulations of corrections to rate constants of activated processes. It is also useful when converting canonical thermal properties into microcanonical ones. This section will be devoted to extracting

level densities from canonical partition functions with the help of the microcanonical temperature.

The canonical partition function is the Laplace transform of the level density, as expressed in (1.54). Once the level density is known for all energies, the partition function can be calculated by an single integration. It is also possible to invert the transformation with the inverse Laplace transform. This requires a canonical partition function that has been continued analytically into the complex plane. Then the inverse Laplace transform is given by

$$\rho(E) = \frac{1}{2\pi i} \int \mathrm{e}^{\beta E} Z(\beta)\, \mathrm{d}\beta. \tag{3.34}$$

The variable β is complex, and the integral is along a line with constant real part of β in the complex plane to the right of any singularity in $Z(\beta)$. Denoting the imaginary and real part of β as $\mathrm{Im}(\beta)$ and $\mathrm{Re}(\beta)$, respectively we have:

$$\rho(E) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \mathrm{e}^{(\mathrm{Re}(\beta)+\mathrm{iIm}(\beta))E}\, Z[\mathrm{Re}(\beta) + \mathrm{iIm}(\beta)]\, \mathrm{dIm}(\beta). \tag{3.35}$$

This inversion formula can occasionally be useful but is unfortunately often not practically applicable. The effort invested into finding a general and closed expression for vibrational level densities at low energies, even in the simplest harmonic oscillator approximation, testifies to this.

Instead, we will calculate two approximate expressions, by postulating a simple form for the level density and then calculate forward and identifying the parameters in the partition function with those of the level density. The first calculation is done to gain some intuitive understanding of the result, the second to obtain a more accurate expression. The latter actually turns out to be very accurate, as will be demonstrated in applications in Chap. 4.

In the first approximation, we ignore the difference between the canonical and the microcanonical temperature and use the canonical caloric curve $E(T)$ to find $T(E)$ by inversion. The level density is approximated with the exponential of the canonical entropy,

$$\rho(E) \simeq \mathrm{e}^{S(T)}. \tag{3.36}$$

Obviously a parameter with dimension reciprocal energy is missing from this expression. We can repair this deficiency by noting that energy fluctuations in the canonical ensemble is $2\sigma_E = 2\sqrt{C_v}T$. This is the effective range of energies over which states are sampled to give the entropy canonically. We therefore have the approximate relation

$$S \approx \ln\left(\rho(E) 2\sqrt{C_v}T\right), \tag{3.37}$$

or

$$\rho(E) \approx \frac{1}{2\sqrt{C_v}T} \mathrm{e}^{S}. \tag{3.38}$$

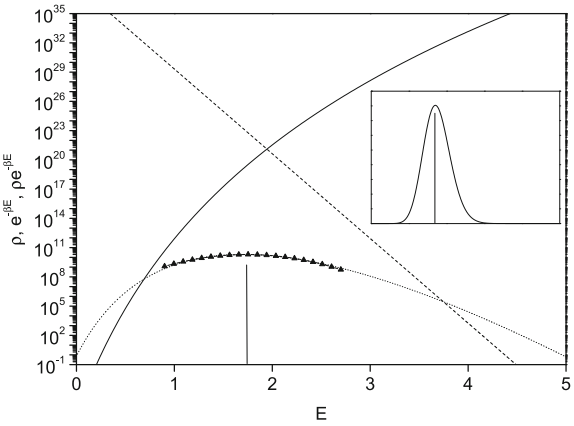


Fig. 3.6 Boltzmann factor (dashed line) for a temperature of 0.05 (in energy units), and level density for a system with heat capacity 50 and a small positive energy offset in the caloric curve (full line). The dotted line is the product of the two, with the maximum of the integrand, E_m, indicated with a vertical line, and the saddle point expansion with triangles. The factors and the product have been scaled with different constants for display purposes. The inset shows a linear plot of $\rho(E)\exp(-\beta E)$, for the same energy range

This estimate is already close to the final result.

In the more accurate procedure one uses a saddle point expansion of the integrand in the canonical partition function, $\rho(E)\exp(-\beta E)$. In that procedure the logarithm of the integrand is expanded around its maximum to second order which makes the integrand Gaussian and the integral easy to perform. The word 'saddle' refers to the fact that the curvatures have different signs in the real and imaginary directions, similar to the topography of a saddle. Figure 3.6 illustrates the procedure.

Assuming that the logarithm of the level density is sufficiently smoothly varying to be represented by an expansion to second order in energy, we can write the partition function as

$$Z(\beta) \approx \rho(E_m) \times \int_0^\infty \exp\Big(-\beta E + \frac{\partial \ln(\rho(E))}{\partial E}|_{E_m}(E-E_m) + \frac{1}{2}\frac{\partial^2 \ln(\rho(E))}{\partial E^2}|_{E_m}(E-E_m)^2\Big)\mathrm{d}E. \tag{3.39}$$

If we chose the expansion point E_m such that $\mathrm{d}\ln(\rho(E))/\mathrm{d}E|_{E_m} = \beta$, the first two terms in the exponential cancel and we have

$$Z(\beta) = \rho(E_m)\int_0^\infty \exp\left(-\beta E_m + \frac{1}{2}\frac{\mathrm{d}^2 \ln(\rho(E))}{\mathrm{d}E^2}|_{E_m}(E-E_m)^2\right)\mathrm{d}E \tag{3.40}$$

If the mean of the integrand, E_m, is bigger than a few times the width of the function, $1/\sqrt{-\mathrm{d}^2\ln(\rho(E))/\mathrm{d}E^2}$, the integration limit zero can be replaced with $-\infty$ and the integral performed to give

$$Z(\beta) = \rho(E_m)\mathrm{e}^{-\beta E_m}\left(\frac{2\pi}{-\frac{\mathrm{d}^2\ln(\rho(E))}{\mathrm{d}E^2}|_{E_m}}\right)^{1/2}. \tag{3.41}$$

The second derivative is negative as required for convergence of the integral, hence the sign in the square root. We find the value as

$$\frac{\mathrm{d}^2\ln(\rho(E))}{\mathrm{d}E^2}|_{E_m} = \frac{\mathrm{d}}{\mathrm{d}E}\frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E}|_{E_m} = \frac{\mathrm{d}}{\mathrm{d}E}\beta = -\beta^2 C^{-1}. \tag{3.42}$$

or

$$\left(\frac{\mathrm{d}^2\ln(\rho(E))}{\mathrm{d}E^2}|_{E_m}\right)^{-1} = -T^2 C. \tag{3.43}$$

This last equality assumes an approximate identity of the microcanonical and the canonical temperatures and heat capacities. This can be verified a posteriori. Equation (3.43) is consistent with the result from Chap. 1 (1.70) that heat capacities in a canonical ensemble are related to the energy fluctuation. The reciprocal of the second logarithmic derivative of the level density sets the scale for the energy fluctuations because it is simply the variance of a certain Gaussian distribution.

With the relation (3.43) we can proceed to calculate the canonical mean thermal energy, $\overline{E}$, as $\overline{E} = -\frac{\partial\ln(Z)}{\partial\beta}$. With (3.40) for Z we have

$$\overline{E} = -\frac{\partial\ln(\rho(E_m))}{\partial\beta} + E_m + \beta\frac{\partial E_m}{\partial\beta} + \frac{1}{2}\frac{\mathrm{d}\ln(\beta^2)}{\mathrm{d}\beta}. \tag{3.44}$$

For the last term (3.43) was used. The first term is

$$-\frac{\partial\ln(\rho(E_m))}{\partial\beta} = -\frac{\mathrm{d}\ln(\rho(E_m))}{\mathrm{d}E_m}\frac{\mathrm{d}E_m}{\mathrm{d}\beta} = -\beta\frac{\mathrm{d}E_m}{\mathrm{d}\beta}. \tag{3.45}$$

This term appears with the opposite sign of the third term in (3.44), they cancel and we are left with

$$\overline{E}(T) = E_m + T. \tag{3.46}$$

This result says that when a system is so well behaved that a saddle point expansion of the population probability in the canonical ensemble is a good approximation, the product of level density and Boltzmann factor peaks at an energy which is T less than the average thermal energy. The peak was determined by equating the logarithmic derivative of the level density with the reciprocal temperature, which is the definition of the microcanonical temperature for the system at the (microcanonical) energy E_m. The T that appears in (3.46) is therefore the microcanonical temperature. This means that we can find this temperature by solving (3.46) for T, for a given value of E_m. This can be very convenient in practical applications, because it is often easier to calculate the canonical partition function than to calculate level densities. The equation will have to be solved numerically, but that often turns out to be very easy with an iterative procedure that will be sketched below.

We can now use the relation between the canonical and microcanonical energies and the microcanonical temperature to get the desired expression for the level density. With the microcanonical temperature found with (3.46) and inserted into (3.40) we have:

$$\rho(E_m) \approx Z(\beta)\mathrm{e}^{\beta E_m}\frac{1}{\sqrt{2\pi T^2 C_v}} \tag{3.47}$$

where also (3.43) was used. Equation (3.47) says that the level density at the energy E_m is equal to the combination of canonical values on the right hand side of the equation. We also see, using that the partition function is the exponential of Helmholtz' free energy, $Z = \exp(-\beta F) = \exp(-\beta(\overline{E} - TS))$, together with (3.47), that

$$\rho \simeq \frac{\mathrm{e}^S}{\sqrt{2\pi T^2 C_v}}. \tag{3.48}$$

This is the desired result. It tells us, among other things and once again, that the entropy is essentially the logarithm of the level density. The relation is made complete by the square root factor, which just accounts for the energy interval over which the states are summed to give the total number of states.

The accuracy of (3.47) is very good. If we want to increase it, we can do so by continuing the expansion of the level density in E to higher orders in (3.39), expanding the exponential and integrating term by term. The third order term gives zero because the integrand is an odd function around E_m. The leading order correction is therefore the fourth order term. We will leave the calculation of that correction to the reader and just mention that it will turn out to be of order $1/C$. This translates to 'small correction' for any but the smallest systems and/or lowest temperatures.

The practical use of (3.47) with (3.46) to calculate level densities is greatly facilitated by the facts that the canonical heat capacity is always non-negative and that this is also the case for its derivative with respect to temperature for harmonic oscillators. This means that an iterative solution to (3.46) has a simple convergence behavior and can be based on a first order expansion of the energy in the temperature for these systems. With the initial temperature T' we have that $\overline{E}(T') \approx \overline{E}(T) + C(T')(T' - T)$. With $\overline{E}(T) = E_m + T$, we have the solution to first order

$$T \approx T' + \frac{E_m + T' - \overline{E}(T')}{C(T')}. \tag{3.49}$$

Due to the positive curvature of the caloric curve, a single iteration will give an estimate for T that is above the solution of (3.46). The values of C are positive (except for the uninteresting point $T = 0$) and this will ensure convergence once the input temperature in (3.49) is above the solution. Hence convergence is ensured for all initial guesses. The situation is illustrated in Fig. 3.7. The appearance of the extra term T in the numerator (3.49) does not change this behavior as long as $C - k_\mathrm{B}$ is still positive.The choice of initial temperature in the iteration is therefore not crucial, and one may use the energy per oscillator, E/s.

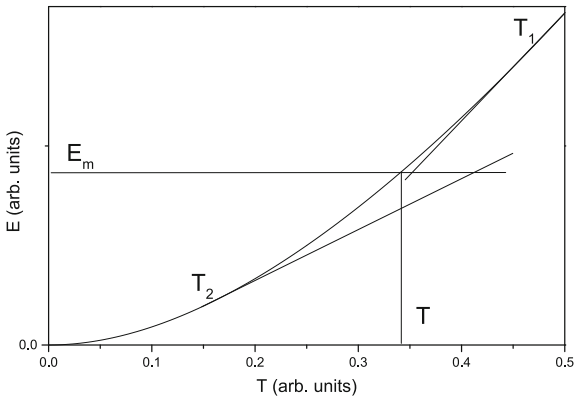


Fig. 3.7 A generic heat capacity with first order solutions to (3.46) for two initial temperatures, above (T_1) and below (T_2) the exact solution to (3.46). The first iteration gives the new temperatures found at the abscissa of the intercept of the two straight lines with the horizontal line at energy E_m

Exercises

3.1 Calculate the level density for 20 harmonic oscillators with all the vibrational frequencies equal to 10^{13} Hz, and a total energy of 3 aJ.

3.2 Estimate the magnitude of the third order term in the expansion formula (3.8) for an arbitrary number of oscillators by setting $\varepsilon = T$. Guess the right solution for all higher order terms. Does the expansion make sense? If not, why? If it does, why?

3.3 Invert (3.20) with the third term included. You can do this either by solving a cubic equation (see Handbook of Mathematical Functions referenced in Appendix A), or by use of successive approximations. With frequency factors of 10^{15} s^{-1} and observation times of 1 ms, determine the particle size (heat capacity) for which the next to leading order term must be included, and the size for which the next term is needed.

3.4 The canonical partition for a particle in a cubic box with side lengths L,

$$Z = \sum_{n,l,k} \mathrm{e}^{-\beta \frac{\hbar^2 \pi^2}{2mL^2}(n^2+l^2+k^2)} \tag{3.50}$$

can be calculated as an integral

$$Z_3 \approx \int_0^\infty \mathrm{d}n \int_0^\infty \mathrm{d}l \int_0^\infty \mathrm{d}k \mathrm{e}^{-\beta \frac{\hbar^2 \pi^2}{2mL^2}(n^2+l^2+k^2)}. \tag{3.51}$$

This gives the leading order in T. The next-to-leading order contribution is calculated by improving this integration. In one dimension the integral is

$$Z_1 \approx \frac{1}{2} \int_{-\infty}^\infty \mathrm{e}^{-\frac{\hbar^2 \pi^2}{2mL^2} n^2} \mathrm{d}n - 1/2 = \frac{\sqrt{\pi}}{2} (\beta \epsilon_0)^{-1/2} - 1/2, \tag{3.52}$$

where the last term is introduced to avoid counting the zero quantum numbers and the shorthand $\epsilon_0 \equiv \hbar^2\pi^2/2mL^2$ is used. Since the problem is separable in coordinates, we have

$$Z_3 = Z_1^3 = \left(\frac{\sqrt{\pi}}{2}(\beta\epsilon_0)^{-1/2} - 1/2\right)^3 . \tag{3.53}$$

Find the level density to the accuracy corresponding to the inclusion of the next-to-leading order term in the partition function.

3.5 An Au_8^+ cluster evaporates an atom, with an activation energy of $E_a = 2.25$ eV at a parent microcanonical temperature of 1700 K. Use harmonic oscillator level densities to estimate the ratio of rate constants calculated with and without the leading order finite heat bath correction, i.e. with and without subtracting the term $E_a/2C_v$ from the microcanonical temperature. You can set $E_r = E_p = 0$.

3.6 Quantify the condition that allows us to calculate (3.40) as a Gaussian integral.

Chapter 4
Thermal Properties of Vibrations

The thermal properties of a particle is to a large degree determined by the motion of the nuclei, and it is therefore of great interest to have a good idea about their behavior. Vibrational motion involving the nuclei have frequencies of typically $\omega = 10^{13}\,\mathrm{s}^{-1}$ to $10^{14}\,\mathrm{s}^{-1}$, which translates into quantum energies of 0.01–0.1 eV. In contrast, the energy scale of electronic excitations, which is the only other serious player in this game when it comes to counting the d.o.f., is 1 eV and above, for molecules and small particles at least.

4.1 Normal Modes

The motion of the nuclei in any piece of material is in principle included in the description of the wave function of the entire system, including the electrons. This description is simplified considerably by separating the motion of the electrons and the nuclei with the Born-Oppenheimer separation where the nuclear motion is taking place on a so-called Born-Oppenheimer (BO) surface which is an effective potential energy surface for the nuclei, provided by the wave function of the electrons and the electrostatic internuclear repulsion. Even though the Born-Oppenheimer separation is an approximation, a determination of reliable potential surfaces for even a small particle still requires a serious amount of quantum mechanical calculations. And quantum mechanics does not stop with the calculation of these potential energies. The nuclear motion on a BO surface is of course also quantized.

In order to understand this quantized motion, one begins by considering the corresponding classical motion. When the energy per degree of freedom is small compared with the binding energy of a monomer, say, the motion can be described as coupled harmonic oscillators. The potential energy can then be written as a bilinear sum in the

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_4

displacement, x_i, of the atoms from their equilibrium positions, collectively denoted by the vector $\overline{x}_0$:

$$V \approx V(\overline{x}_0) + \sum_{i,j} {}^1\!/_2 \, \frac{\partial^2 V}{\partial x_i \partial x_j} x_i x_j. \tag{4.1}$$

The zero of energy, $V(\overline{x}_0)$, can be set to zero without loss of generality, and the linear term vanishes because the expansion points are the equilibrium positions. Note that V can be considered a function of the coordinates or the displacements from the equilibrium positions without any change in the second derivatives. The matrix

$$\frac{\partial^2 V}{\partial x_i \partial x_j} \equiv V_{i,j}, \tag{4.2}$$

calculated at the equilibrium position, is called the Hessian matrix. Together with the masses of the atoms it determines the equations of motion of the vibrations to the extent they can be considered harmonic, both classically and quantum mechanically. If the interactions are local, only the displacements of neighboring atoms couple and contribute to V. In general, however, interactions are non-zero for other combinations of i and j than just those that correspond to coordinates on nearest-neighbor atoms. Irrespective of the precise range of the interaction, the matrix will couple the vibrational motion of individual atoms directly or indirectly.

To find the motion, the kinetic energy must also be included. It is

$$K = \sum_i {}^1\!/_2 \, m_i \dot{x}_i^2. \tag{4.3}$$

This is diagonal but in general not proportional to the unit matrix because the masses need not be identical. The m_i's are only guaranteed identical for the three values of i that correspond to a displacement of a specific atom in the three directions of space. Only for a particle of all-identical atoms is K proportional to the unit matrix.

To find the vibrational frequencies, one has to decouple the vibrations that are coupled by the Hessian matrix from each other with the purpose of writing the Hamiltonian as a sum of independent oscillators. This can be accomplished by a direct diagonalization of the Hessian, but only in the special cases where all the masses are identical, because then the kinetic energy is proportional to the unit matrix. In these situations the Hessian diagonalizes directly to give the frequencies

$$\omega^2 = \lambda / m. \tag{4.4}$$

If the particle contains two or more different atoms of different masses, the matrix that diagonalizes the Hessian will un-diagonalize the kinetic energy term. The different masses can be cause by the presence of different elements but the effect is there also for different isotopes of the same element.

The solution to this problem is to define mass-weighted displacement coordinates,

$$x_i \rightarrow \tilde{x}_i \equiv x_i \sqrt{m_i}. \tag{4.5}$$

In terms of these coordinates the Hamilton function becomes

$$\begin{aligned} H = K + V &= \sum_i {}^1\!/_2\, \dot{\tilde{x}}_i^2 + \sum_{i,j} {}^1\!/_2\, \frac{\partial^2 V}{\partial \tilde{x}_i \partial \tilde{x}_j} \tilde{x}_i \tilde{x}_j \\ &= \sum_i {}^1\!/_2\, \dot{\tilde{x}}_i^2 + \sum_{i,j} {}^1\!/_2\, \frac{1}{\sqrt{m_i m_j}} \frac{\partial^2 V}{\partial x_i \partial x_j} \tilde{x}_i \tilde{x}_j \\ &\equiv \sum_i {}^1\!/_2\, \dot{\tilde{x}}_i^2 + \sum_{i,j} {}^1\!/_2\, \tilde{V}_{i,j} \tilde{x}_i \tilde{x}_j. \end{aligned} \tag{4.6}$$

This makes the kinetic energy proportional to the unit matrix and the mass weighted Hessian $\tilde{V}_{i,j}$ can therefore be diagonalized without un-diagonalizing the kinetic energy matrix. The Hessian is a real symmetric matrix and those may be diagonalized with the method of Jacobi transformations, which is described in texts on numerical analysis. After diagonalization, which gives the eigenvalues λ_k, one has a Hamiltonian which is the sum of independent degrees of freedom:

$$H = \sum_k {}^1\!/_2\, \dot{\overline{u}}_k^2 + \sum_k {}^1\!/_2\, \lambda_k \overline{u}_k^2, \tag{4.7}$$

where u_k is the k'th linear combination of $\tilde{x}_i$, the square is the inner (vector) product and λ_k is the k'th eigenvalue of the transformed Hessian $\tilde{V}_{i,j}$. Comparing the expression (4.7) with the Hamilton function for a harmonic oscillator,

$$H_{ho} = {}^1\!/_2\, m\dot{x}^2 + {}^1\!/_2\, m\omega^2 x^2, \tag{4.8}$$

we see that the frequency for the kth eigenvector is $\omega_k = \sqrt{\lambda_k}$. The eigenvectors can be converted to the amplitudes of the physical vibrations by multiplication of each component by the square root of the relevant mass.

As a simple but very instructive demonstration of the procedure we will consider the vibrations of a diatomic molecule. The potential energy is given by the square of the deviation of the bond length relative to the equilibrium value x_0:

$$V = {}^1\!/_2\, k \left(|\overline{x}_1 - \overline{x}_2| - x_0 \right)^2, \tag{4.9}$$

where $\overline{x}_k$ are the Cartesian coordinates of atom k. The second derivative of this potential energy with respect to coordinates η_i and η_j is

$$\frac{\partial^2 V}{\partial \eta_i \partial \eta_j} = k \frac{\partial |\overline{x}_1 - \overline{x}_2|}{\partial \eta_i} \frac{\partial |\overline{x}_1 - \overline{x}_2|}{\partial \eta_j}, \tag{4.10}$$

plus a term which is zero in the equilibrium position. The coordinates η_i, η_j can each run over all of the six Cartesian coordinates of the two atoms. Performing the derivatives one gets

$$\frac{\partial^2 V}{\partial \eta_i \partial \eta_j} = \frac{k}{x_0^2}(\eta_i - \eta_i')(\eta_j - \eta_j'), \tag{4.11}$$

where the primed coordinate is the Cartesian coordinate of the other atom, along the same spatial axis, e.g. η_i' is the y-coordinate of atom 1 if η_i is the y-coordinate of atom 2 etc. We place the molecule along the y-axis and find that only the following matrix elements are non-zero:

$$\begin{aligned} V_{y_1,y_2} = V_{y_2,y_1} &= -k \\ V_{y_1,y_1} = V_{y_2,y_2} &= k. \end{aligned} \tag{4.12}$$

After weighting with the masses, these elements become

$$\begin{aligned} \frac{1}{m_1}\frac{\partial^2 V}{\partial y_1^2} &= \frac{k}{m_1} \\ \frac{1}{m_2}\frac{\partial^2 V}{\partial y_2^2} &= \frac{k}{m_2} \\ \frac{1}{\sqrt{m_1 m_2}}\frac{\partial^2 V}{\partial y_1 \partial y_2} = \frac{1}{\sqrt{m_1 m_2}}\frac{\partial^2 V}{\partial y_2 \partial y_1} &= -\frac{k}{\sqrt{m_1 m_2}}. \end{aligned} \tag{4.13}$$

The eigenvalues are found, as usual, as the zeros of the characteristic polynomial, which in this case gives the equation

$$\lambda^4 \left(\left(\frac{k}{m_1} - \lambda \right) \left(\frac{k}{m_2} - \lambda \right) - \frac{k^2}{m_1 m_2} \right) = 0, \tag{4.14}$$

with the solutions

$$\lambda = 0, \quad \sqrt{k\left(\frac{1}{m_1} + \frac{1}{m_2} \right)}. \tag{4.15}$$

The zero frequency solution has a five-fold multiplicity. Zero frequencies generally correspond to the rotational or translational motion of the particle, because there is no potential energy associated with those degrees of freedom, hence no restoring force and the vibrational frequencies are then zero. The zero eigenvalues in (4.15) correspond to the translational motion of the molecule along the three axes and the two rotational modes possible for a linear molecule. Degenerate modes can be mixed as you desire. Any sets of linear combination of these modes go for these, as long as the orthonormality is conserved. We can chose the zero mode ($\lambda = 0$) corresponding

to the motion along the y direction. This is represented by a vector with six entries of which four are zero. In the mass weighted coordinates it is:

$$\overline{u}_1 = \begin{pmatrix} 0 \\ \sqrt{m_1} \\ 0 \\ 0 \\ \sqrt{m_2} \\ 0 \end{pmatrix}, \tag{4.16}$$

where the first three components give the x, y, z amplitudes of atom 1 and the three last the same amplitudes of atoms 2. Division of the two components of this mode by their respective weight $\sqrt{m_i}$ removes the mass weighting. Leaving out all the inessential zeros associated with the motion in the x, z directions from the vector, it becomes

$$\overline{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{4.17}$$

which is indeed the expected translation in the y direction.

The non-zero eigenvalue gives the frequency $\sqrt{k\left(\frac{1}{m_1} + \frac{1}{m_2}\right)}$. This is equal to $\omega = \sqrt{k/\mu}$ where μ is the reduced mass of the system,

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}, \tag{4.18}$$

as it must be for a two-body problem. The mode associated with this frequency is

$$\overline{u}_2 = \begin{pmatrix} \sqrt{m_2} \\ -\sqrt{m_1} \end{pmatrix}, \tag{4.19}$$

or, after undoing the mass weighting,

$$\overline{x}_2 = \begin{pmatrix} \sqrt{\frac{m_2}{m_1}} \\ -\sqrt{\frac{m_1}{m_2}} \end{pmatrix}. \tag{4.20}$$

If we go a distance δ along this mode, the displacement of the center of mass will be

$$y_{CM} = \delta\left(m_1\sqrt{\frac{m_2}{m_1}} - m_2\sqrt{\frac{m_1}{m_2}}\right) = 0, \tag{4.21}$$

which is precisely as expected for a vibration.

The procedure, with or without mass weighting, provides the desired linear combinations of atomic displacements that decouple from each other. They are called normal coordinates and their motion normal modes. When the basically harmonic

motion of the nuclei of sufficiently large particles is quantized, these modes are known as phonons. They are basically quantized sound waves, hence the name.

Diagonalization of the potential energy means that the motion of a single normal coordinate will involve a large number of atomic coordinates, even when atoms only interact with nearest neighbor interactions. Conversely, if just a single atom is excited it will also involve a number of normal modes.

There are $s = 3N$ coordinates for N atom particles. If the particle is not fixed to some point in space, however, some of these will have vanishing frequencies, as we already saw for the diatomic particle. A free particle will have an unrestricted translational motion of the center of mass (that's the definition of free in physics). This reduces the number of vibrational normal modes by three. A free particle can also rotate freely. A linear particle has two rotational axis and a nonlinear particle three of these. This adds up to a number of vibrational normal modes of $s = 3N - 6$ for a non-linear particle and $s = 3N - 5$ for a linear.

The case of a large particle composed of identical-mass atoms will be illustrated with the calculation of the Hessian matrix for Lennard-Jones (LJ) clusters. Another example of a spectrum is given in Chap. 9, where the normal modes of a one-dimensional linear chain is calculated analytically with the purpose of comparing it to numerical simulations.

LJ clusters are model systems that describe neutral rare gas clusters fairly well and have a potential energy which is

$$V = 2\varepsilon \sum_{i \neq j} \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right) = 4\varepsilon \sum_{i} \sum_{j>i} \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right), \quad (4.22)$$

where r_{ij} is the distance between atoms i and j. The parameter ε is an energy and σ is a length, both of which have values characteristic of the rare gas in question. For more details about this potential, see Chap. 12. It is convenient to scale the energy and the length to get the expression

$$V = \sum_{i \neq j} 2 \left(r_{ij}^{-12} - r_{ij}^{-6} \right) = 4 \sum_{i} \sum_{j>i} \left(r_{ij}^{-12} - r_{ij}^{-6} \right). \quad (4.23)$$

The factor 4 makes the minimum energy of two atoms equal to -1. Because the distances involve all three Cartesian coordinates of the two atoms, it is most convenient for the present calculation to number the atoms and indicate the Cartesian coordinate explicitly next to the atom number. The Hessian is then written

$$V_2(k, \alpha, l, \beta) \equiv \frac{\partial^2 V}{\partial x_{k,\alpha} \partial x_{l,\beta}}, \quad (4.24)$$

where roman subscripts number the atoms and run from 1 to N, and Greek subscripts label the Cartesian coordinates x, y, z which for bookkeeping reasons is identified with 1, 2, 3. We calculate the first derivative, the negative of the force, to be

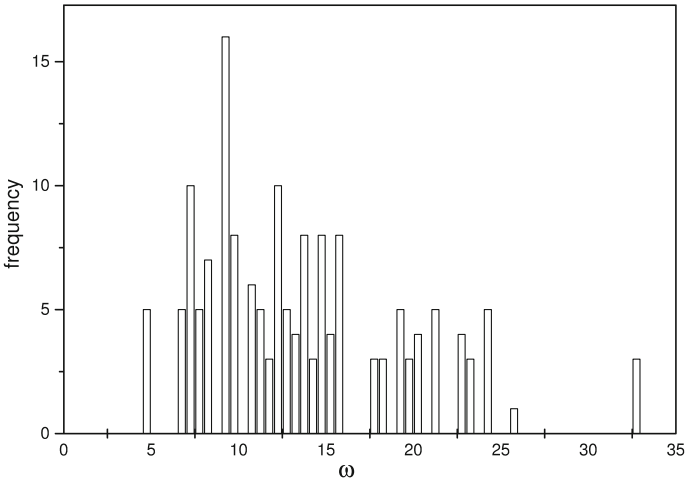


Fig. 4.1 The vibrational spectrum of the $N = 55$ Lennard-Jones cluster, calculated with the methods described in the text. For this specific particle the frequently used Debye model described in the text below is of marginal validity

$$\begin{aligned}\frac{\partial V}{\partial x_{k,\alpha}} &= 2\sum_{j\neq i}\left(-12r_{ij}^{-14}+6r_{ij}^{-8}\right)(x_{i,\alpha}-x_{j,\alpha})\left(\delta_{i,k}-\delta_{j,k}\right) \\ &= 4\sum_{i}\left(-12r_{ik}^{-14}+6r_{ik}^{-8}\right)\left(x_{k,\alpha}-x_{i,\alpha}\right)\left(1-\delta_{k,i}\right).\end{aligned} \tag{4.25}$$

In this expression it is to be understood that terms of the kind $(1-\delta_{i,k})/r_{ik}^{p}$ are identically zero for all powers p when $i = k$. The second derivative gives the Hessian:

$$\begin{aligned}&\frac{\partial^2 V}{\partial x_{k,\alpha}\partial_{l,\beta}} \\ &= 4\sum_{i}\left(-12r_{ik}^{-14}+6r_{ik}^{-8}\right)\left(1-\delta_{k,i}\right)\left(\delta_{k,l}-\delta_{i,l}\right)\delta_{\alpha,\beta} \\ &+ 4\sum_{i}(12\cdot 14r_{ik}^{-16}-6\cdot 8r_{ik}^{-10})(x_{k,\alpha}-x_{i,\alpha})(x_{k,\beta}-x_{i,\beta})(\delta_{i,l}-\delta_{k,l})(1-\delta_{k,i}) \\ &= 24\sum_{i}\left(1-\delta_{i,k}\right)\left(\delta_{k,l}-\delta_{i,l}\right)r_{ik}^{-8} \\ &\times\left[\left(1-2r_{ik}^{-6}\right)\delta_{\alpha,\beta}+\left(28r_{ik}^{-8}-8r_{ik}^{-2}\right)(x_{k,\alpha}-x_{i,\alpha})(x_{k,\beta}-x_{i,\beta})\right],\end{aligned} \tag{4.26}$$

When one places this in an array in a computer, the Hessian matrix element corresponding to this derivative can conveniently be numbered as $(3(k-1)+\alpha, 3(l-1)+\beta)$. Figure 4.1 shows the result of a calculation of the $3N-6 = 159$ non-zero vibrational frequencies of the $N = 55$ Lennard-Jones cluster.

4.2 Thermal Properties of Harmonic Oscillators

After having given a procedure to determine the vibrational frequencies of a particle, we can now consider the thermal properties of such a particle. We will start with the high energy limit of the level density.

For vibrational motion, high energy is defined as an energy per degree of freedom, E/s, which is large compared with any vibrational energy quantum of the system. We have seen earlier that the harmonic oscillator level density for identical frequencies can be written, in the simplest approximation, as

$$\rho \approx \frac{E^{s-1}}{(s-1)!\,(\hbar\omega)^s}, \tag{4.27}$$

where ω is the vibrational frequency and s is the number of oscillators. This result was found by convoluting the level densities of all s oscillators, treating the energy as a continuous variable. We will now improve on this result and generalize to different frequencies. The method will be to calculate the canonical partition function both exactly and in terms of a postulated level density, and adjust the parameters in the postulated level density to get agreement.

Let's first calculate the canonical partition function. The oscillators are decoupled and the partition function can be written down immediately as the product of partition functions of the individual oscillators, but it may be instructive to see the calculation made explicitly, so let's do this. The total energy of a given state is the sum of all the oscillator energies,

$$E = \sum_{i=1}^{s} n_i \hbar\omega_i, \tag{4.28}$$

where n_i can take all non-negative integer values, the sum runs over all vibrational modes, and the zero point of energy is chosen as the quantum mechanical ground state. The partition function is

$$Z = \sum_{n_i} \mathrm{e}^{-\beta \sum_i n_i \hbar\omega_i}. \tag{4.29}$$

The exponential can be factored into exponentials and the sum over states can be performed for each oscillator separately. This gives

$$Z = \prod_{i=1}^{\infty} \left(1 - \mathrm{e}^{-\beta\hbar\omega_i}\right)^{-1}. \tag{4.30}$$

The partition function is the product of the partition functions of the individual oscillators, as it should be for independent degrees of freedom. The excitation energy in the high temperature limit for a single harmonic oscillator is

$$\overline{E} = -\frac{\partial \ln\left(1 - \mathrm{e}^{-\beta\hbar\omega}\right)^{-1}}{\partial\beta} = \frac{\hbar\omega \mathrm{e}^{-\beta\hbar\omega}}{1 - \mathrm{e}^{-\beta\hbar\omega}} = T - 1/2\hbar\omega + O(\hbar\omega/T). \tag{4.31}$$

Because the oscillators are independent, the total energy is the sum of energies of the individual oscillators.

$$\overline{E} = \sum_i \frac{\hbar\omega_i \mathrm{e}^{-\beta\hbar\omega_i}}{1 - \mathrm{e}^{-\beta\hbar\omega_i}} \approx sT - 1/2 \sum_i \hbar\omega_i. \tag{4.32}$$

The first term on the right hand side, sT, corresponds to a canonical high temperature heat capacity of s and a level density which is proportional to energy to the power $s - 1$, $\rho(E) \propto E^{s-1}$, as shown in Sect. 2.3. The second term in (4.32) is also easy to include into the level density. It simply corresponds to a shift of the energy by the same amount:

$$\begin{aligned} Z &\propto \int_0^\infty \left(E + 1/2 \sum_i \hbar\omega_i\right)^{s-1} \mathrm{e}^{-\beta E} \mathrm{dE} \\ &= \int_{1/2\sum_i \hbar\omega_i}^\infty E^{s-1} \mathrm{e}^{-\beta E} \exp\left(\beta/2 \sum_i \hbar\omega_i\right) \mathrm{dE} \\ &\approx \int_0^\infty E^{s-1} \mathrm{e}^{-\beta E} \exp\left(\beta/2 \sum_i \hbar\omega_i\right) \mathrm{dE} \\ &= \exp\left(\beta/2 \sum_i \hbar\omega_i\right) T^s (s-1)! \end{aligned} \tag{4.33}$$

Calculating $\overline{E}$ for this partition function confirms that it indeed gives the caloric curve in (4.32). The extension of the integral to zero in the third step is justified because we are considering the high temperature limit which means the integrand peaks at the energy $E + 1/2 \sum_i \hbar\omega_i = sT$ which is high compared to the lower limit on the integral. The contribution to the integral from the low energy part can therefore be ignored.

This establishes the functional form of the high energy section of the level density to next-to-leading order in the energy, apart from a multiplicative factor. Calculations of energies or heat capacities cannot give us information about this, because one takes the logarithmic derivative of Z to get the energy, and multiplicative constants cancel in that process. Instead we can use the free energy, F, because it contains the entropy and the entropy is essentially the logarithm of the number of states and therefore depends on the unknown numerical constant multiplying E^{s-1}. The free energy calculated with the partition function in (4.30) is

$$F = -T \ln(Z) = T \sum_i \ln\left(1 - \mathrm{e}^{-\beta\hbar\omega_i}\right) \approx T \sum_i \ln\left(\beta\hbar\omega_i - 1/2(\beta\hbar\omega_i)^2\right) \tag{4.34}$$

$$\approx T \sum_i \ln(\beta\hbar\omega_i) - 1/2 \sum_i \hbar\omega_i .$$

The partition function for the level density $\rho(E) = c(E + 1/2 \sum_i \hbar\omega_i)^{s-1}$ was calculated in (4.33), apart from the multiplicative factor c. Introducing it into the level density gives the corresponding free energy as:

$$\begin{aligned} F &= -T \ln\left(c\, \mathrm{e}^{\beta/2 \sum_i \hbar\omega_i}\, T^s\, (s-1)!\right) \\ &= -T \ln(c) - 1/2 \sum_i \hbar\omega_i - sT \ln(T) - T \ln\left((s-1)!\right) . \end{aligned} \tag{4.35}$$

If we compare this with (4.34) term-by-term, we can check our previous calculation and find the constant c. We have identical expressions provided

$$c = \frac{1}{(n-1)! \prod_i \hbar\omega_i}. \tag{4.36}$$

Hence the harmonic oscillator level density in the high temperature limit is

$$\rho(E) = \frac{\left(E + 1/2 \sum_i \hbar\omega_i\right)^{s-1}}{(s-1)! \prod_i \hbar\omega_i}. \tag{4.37}$$

Figure 4.2 illustrates how this approximation fares. The figure shows ratios of the exact to the approximate high energy level densities from (4.37) for a number of systems of different sizes, all to the power $1/s$. The exact values is calculated with the Beyer-Swinehart procedure that will be derived below. The curves thus correspond loosely speaking to the ratio of exponentials of the exact and approximate values of the entropies per degree of freedom. As expected from the nature of the high temperature approximation, the calculated quality of the approximation is good when the energy per oscillator is large compared to the average frequency and becomes worse at low energies.

We used the type of level density given by (4.37) in Chap. 3 for the definition of microcanonical temperatures. We can now also write down an improved version of the high temperature limit for the microcanonical temperatures for harmonic oscillators:

$$T_m = \left(\frac{\partial \ln\left(\rho(E)\right)}{\partial E}\right)^{-1} = \frac{E + 1/2 \sum_i \hbar\omega_i}{s-1} \quad \text{(microcanonical)} \tag{4.38}$$

The canonical relation between (mean thermal) excitation energy and the (fixed value) temperature is;

$$\overline{E} = sT - 1/2 \sum_i \hbar\omega_i \quad \text{(canonical)}. \tag{4.39}$$

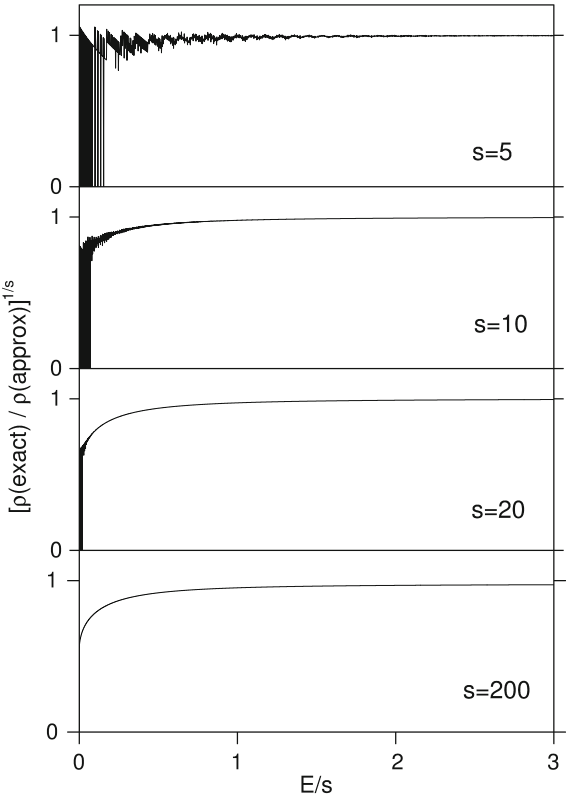


Fig. 4.2 A comparison of the harmonic oscillator level densities calculated with the Beyer-Swinehart algorithm and the high temperature limit in (4.37). The vibrational frequencies were selected randomly in the interval from 0 to 2 and the average normalized to unity. The frames show the ratios of the two calculations for $s = 5$, 10, 20 and 200 oscillators in the system, to the power $1/s$. The abscissa is the energy per degree of freedom. Several curves were calculated but only one of each is shown. The others were similar. The representation of the data is chosen to emphasize the scaling of the error. Both the absolute and the relative error increases rapidly with size, but the error in the entropy per degree of freedom stays constant

As before, where only the first term was calculated, the only difference between the canonical and microcanonical expression is the replacement of the canonical value s with the microcanonical $s - 1$, equivalent to the relation

$$E = \overline{E}(T_m) - T_m, \tag{4.40}$$

in agreement with (3.46).

4.3 Debye Particles

The method to calculate the vibrational spectrum of a particle, given the second derivatives of the potential energy function, was presented above, but no general property of the spectra of vibrational frequencies has actually been given so far, apart from the number of non-zero frequencies of a free particle. It is clear that the vibrational spectrum depends on the nature of a system and in order to make specific statements about it in detail, we need to know the interaction potential to apply the diagonalization procedure. This information is rarely available for nanoparticles.

This lack of information is mainly due to the irregular structures of small particles. The simplifications one can apply to a periodic lattice are usually not available to real nanoparticles. Worse, it is rarely possible to even know the precise structure of a particle, even composed of as little as a few tens of atoms, so even if the effective interaction is known, the Hessian cannot be calculated.

An alternative to the diagonalization of a specific structure Hessian is to extrapolate vibrational spectra from bulk values. For sufficiently large particles this is expected to be meaningful. The model we will use for this purpose is derived using the nature of the motion of the atoms in the solid as sound waves. This is a long wavelength approximation and it is most interesting for low temperatures where long wavelength vibrations are the only ones that are thermally excited. It relates the frequency to the speed of sound, c, by the dispersion relation $\omega = ck$, where k is the wave vector. The dispersion relation is identical to that of electromagnetic waves (light) with the speed of sound substituting the speed of light.

In a solid there are two transverse and one longitudinal types of modes. For a single one of these polarizations the density of states is similar to that of a single polarization mode of light. Light will be treated in Chap. 6 where it will be shown that the level density has a quadratic dependence on the frequency. Per polarization direction it is:

$$g(\omega)d\omega = \frac{\omega^2 V}{2\pi^2 c^3}\mathrm{d}\omega. \tag{4.41}$$

For a more realistic parametrization of the phonon spectrum, one must take into account that the two transverse and the single longitudinal types of mode present in monoatomic crystals have different speeds. The speed of the longitudinal modes, $c_\parallel$, and the speed of the two perpendicular types of modes, $c_\perp$, are different and one could in principle operate with two different distributions of the type in (4.41). We will simplify and define the effective speed of sound as the sum of inverse powers:

$$3c^{-3} = 2c_\perp^{-3} + c_\parallel^{-3}. \tag{4.42}$$

The number of modes in the particle determines the frequency, ω_D, at which the spectrum terminates. In bulk where the difference between $3N$ and $3N - 6$ is irrelevant, we have

$$3N = \int_0^{\omega_D} \frac{3\omega^2 V}{2\pi^2 c^3}\mathrm{d}\omega = \frac{V\omega_D^3}{2\pi^2 c^3}, \tag{4.43}$$

or

$$\omega_D = c\left(6\pi^2 \frac{N}{V}\right)^{1/3}. \tag{4.44}$$

The final result for the spectrum is therefore

$$g(\omega)\mathrm{d}\omega = 9N\frac{\omega^2}{\omega_D^3}\mathrm{d}\omega \quad \omega \leqslant \omega_D. \tag{4.45}$$

The cutoff frequency ω_D in the theory (the Debye frequency) is often given in temperature units, $T_D = \hbar\omega_D/k_B$, and is then naturally called the Debye temperature.

The spectrum was suggested by Debye to explain the low temperature heat capacity of solids, which were found experimentally to vary as T^3, and the spectral density in (4.45) is best used to describe bulk crystals composed of monoatomic species at low temperature excitation energy per d.o.f. In an older theory by Einstein, which we will meet below, all frequencies were assumed to be identical. This is often called an Einstein spectrum and a piece of matter described with such a spectrum an Einstein crystal.

All thermal properties of a Debye nanoparticle are determined by the spectrum in (4.45). We will first use this to calculate the low temperature heat capacity, with the expected T^3 dependence. The heat capacity is the sum of the heat capacities of the individual oscillators and is (see Exercise 4.2):

$$C = \sum_i \left(\frac{\hbar\omega_i}{T}\right)^2 \frac{\mathrm{e}^{-\hbar\omega_i/T}}{\left(1-\mathrm{e}^{-\hbar\omega_i/T}\right)^2} \approx \int_0^{\omega_D} 9N\frac{\omega^2}{\omega_D^3}\left(\frac{\hbar\omega}{T}\right)^2 \frac{\mathrm{e}^{-\hbar\omega/T}}{\left(1-\mathrm{e}^{-\hbar\omega/T}\right)^2}\mathrm{d}\omega. \tag{4.46}$$

With the substitution $x = \hbar\omega/T$ we have

$$C = 9N\left(\frac{T}{\hbar\omega_D}\right)^3 \int_0^{\hbar\omega_D/T} x^4 \frac{\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2}\mathrm{d}x. \tag{4.47}$$

First a check: For high temperatures the upper integration limit is small, $\hbar\omega_D/T \ll 1$, and the exponential in the numerator can be set to unity, whereas the denominator is approximated by x^2. This gives

$$C \approx 9N\left(\frac{T}{\hbar\omega_D}\right)^3 \int_0^{\hbar\omega_D/T} x^2\mathrm{d}x = 9N\left(\frac{T}{\hbar\omega_D}\right)^3 1/3\left(\frac{\hbar\omega_D}{T}\right)^3 = 3N, \tag{4.48}$$

as expected (right?). This confirms that we got the numerical factors right.

In the more interesting limit of low temperatures, $\hbar\omega_D/T$ is big and the integrand in (4.47) is essentially zero when the upper limit is reached. We can therefore approximate the integral by setting the upper limit to infinity. The integral is then simply a temperature independent number and this gives us the T^3 law without any further ado. If we also want the numerical coefficient we need to perform the integral. It is calculated to

$$\int_0^\infty x^4 \frac{\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2}\mathrm{d}x = \zeta(4)\,4! = \frac{4\pi^4}{15}, \tag{4.49}$$

where ζ is Riemann's zeta function. The low temperature heat capacity then becomes

$$C \approx \frac{12\pi^4 N}{5}\left(\frac{T}{\hbar\omega_D}\right)^3 \equiv \frac{12\pi^4 N}{5}\left(\frac{T}{T_D}\right)^3, \tag{4.50}$$

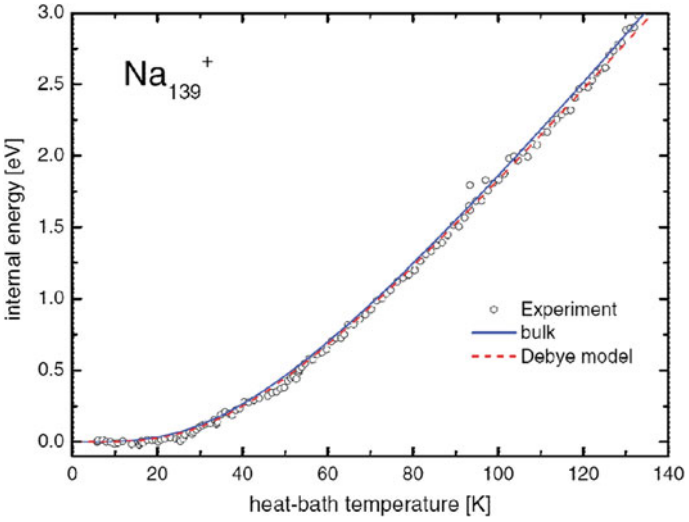


Fig. 4.3 The thermal excitation energies at low temperatures of the positively charged 139 atom sodium cluster. The dark line is the bulk heat capacity, scaled to the $3N - 6$ vibrational degrees of freedom. The dashed line is the fit to the experimental data with the Debye model. The fitted temperature is 164 ± 10 K, consistent with the bulk value of 154 K in Table 4.1. Reprinted from C. Hock et al., *Phys. Rev. B*, **84** (2011) 113401, http://prb.aps.org/abstract/PRB/v84/i11/e113401. Copyright (2001) by the American Physical Society

or, expressing the Debye temperature in terms of the speed of sound (4.44),

$$C \approx \frac{2\pi^2 V}{5}\left(\frac{T}{\hbar c}\right)^3. \tag{4.51}$$

Figure 4.3 shows measured thermal energies in the Na_{139}^+ cluster at low temperatures. The good agreement between the experimental data and the Debye model suggests a strong similarity between the vibrational spectra of bulk matter and the $N = 139$ sodium particle. In particular there is no gap in the low energy part of the spectrum. This would have been manifested as a strong decrease of the heat capacity at low energies.

Measured Debye temperatures vary a lot across the periodic table. Table 4.1 lists a compilation of literature data. For some elements several rather different values can be found in the literature and the numbers in the table should be taken with a grain of salt.

To find the high energy/temperature limit for the level density specific to a Debye particle, we can use the general formula given in (4.37). This requires the calculation of a couple of numbers related to the spectrum. The average is easily calculated:

$$\langle\omega\rangle = \frac{\int_0^{\omega_D} \omega g(\omega)\mathrm{d}\omega}{3N} = \frac{3}{4}\omega_D. \tag{4.52}$$

Table 4.1 Measured Debye temperatures (in kelvin) for some elements

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Li 376	Be 1000						
Na 154	Mg 318						
K 94	Ca 232	Sc 359	Ti 325	V 358	Cr 525	Mn 400	Fe 445
Rb 55	Sr 147	Y 256	Zr 299	Nb 241	Mo 450	Tc 454	Ru 555
Cs 40	Ba 110	La 207	Hf 252	Ta 258	W 330	Re 416	Os 467
Group 9	Group 10	Group 11	Group 12	Group 13	Group 14	Group 15	Group 18
				B 1250	C 1860		Ne 75
				Al 409	Si 675		Ar 85
Co 385	Ni 383	Cu 330	Zn 200	Ga 240	Ge 400	As 285	Kr 73
Rh 415	Pd 279	Ag 221	Cd 120	In 119	Sn 185	Sb 200	Xe 64
Ir 425	Pt 230	Au 167	Hg 100	Tl 96	Pb 86	Bi 120	

The product of the frequencies is calculated as

$$\prod_{j=1}^{3N} \omega_j = \exp\left(\ln\left(\prod_{j=1}^{3N} \omega_j\right)\right) = \exp\left(\sum_{j=1}^{3N} \ln\left(\omega_j\right)\right). \tag{4.53}$$

To an accuracy which is not worse than the physical assumptions involved in the use of the spectrum, we can replace the sum of logarithms with an integral,

$$\sum_{j=1}^{3N} \ln\left(\omega_j\right) \approx \int_0^{\omega_D} g(\omega)\ln(\omega)\mathrm{d}\omega = \int_0^{\omega_D} 9N\frac{\omega^2}{\omega_D^3}\ln(\omega)\mathrm{d}\omega. \tag{4.54}$$

Partial integration gives

$$\prod_{j=1}^{3N} \omega_j \approx \omega_D^{3N}\mathrm{e}^{-N}. \tag{4.55}$$

With the help of the general expression in (4.37) we can then approximate the high energy limit of the level density of a Debye particle as

$$\rho_D(E) \approx \frac{\mathrm{e}^N\left(E + \frac{3}{8}N\hbar\omega_D\right)^{3N-1}}{(3N-1)!\,(\hbar\omega_D)^{3N}} \quad \text{(high energy limit)}. \tag{4.56}$$

For a free particle, replace $3N$ with $3N - 6$.

An important fact of life is that interatomic potentials are in general not absolutely harmonic. The finite, and usually positive thermal expansion coefficients are due to the anharmonic components in the potential. This will render the Debye model less precise at high temperatures, when T exceeds some fraction of $\hbar\omega_D$. Another factor

that blurs our nice and simple picture is the non-isotropy of the material that must be expected to influence the vibrational spectrum, for example at the surface which is an important factor in shaping the vibrational spectrum of nanoparticles. All this said, we can nevertheless consider the Debye spectrum as a useful first approximation for not too small particles.

4.4 Degenerate Oscillators

In an Einstein crystal all vibrational frequencies are identical, and for this problem one can calculate the level density exactly. It is a sum of δ functions, located at multiples of the quantum energy. The canonical thermal properties of this system are easy to calculate and we will leave it as an exercise for the reader (if it is still needed at this point). The determination of the microcanonical partition function reduces to the combinatorics problem of how many ways one can distribute $N_E \equiv E/\hbar\omega$ units into s boxes. The problem is equivalent to counting how many ways one can put N_E cows into a stable with s enclosures when there is room for more than one cow in each enclosure.[1] The trick to solve this problem is that, instead of placing animals in enclosures, one lines up the animals and places the separators around them. There are $s-1$ separators. The first separator can be located in either of N_E+1 different positions, the next in N_E+2 different positions etc. This gives a total number of partitionings of $(N_E+s-1)!/N_E!$ if we could tell the separators apart. But we can't and therefore we need to divide by the number of permutations of them, $(s-1)!$ This gives us a total of $(N_E+s-1)!/N_E!(s-1)!$ ways to distribute N_E identical units into s boxes.

Leaving agriculture and translating the result into level densities with the proper unit conversions we get:

$$\rho(E) = \frac{1}{\hbar\omega}\frac{\left(\frac{E}{\hbar\omega}+s-1\right)!}{\left(\frac{E}{\hbar\omega}\right)!(s-1)!}. \tag{4.57}$$

The formula is still discrete and the level density should be understood as the amplitude of the δ functions at the energies indicated in the argument.

The situation described by the equation is rather unrealistic, one almost never[2] has a collection of completely degenerate oscillators. It is nevertheless very useful because it provides us with a quick estimate of thermal quantities in closed form.

It is useful to develop an intuition of how the factorials work in the ratios. As a start we can find the high energy limit. The factors $1/\hbar\omega$ and $1/(s-1)!$ already appear in the general expression in (4.37) and it is only necessary to check the

[1] To avoid any misunderstanding of any kind whatsoever, it should be pointed out that the calculation also works with pigs.

[2] This is the expression mathematicians use when they mean never.

behavior of the rest of the expression, i.e. the factor $\frac{(N_E+s-1)!}{N_E!}$ in the dimensionless variables. The check is best done on the logarithm of the fraction. Instead of using the approximations of the Γ function, we calculate directly the logarithm of the ratio:

$$\ln\left(\frac{(N_E+s-1)!}{N_E!}\right) = \sum_{i=1}^{s-1} \ln(N_E+i). \tag{4.58}$$

If each term is expanded around $N_E + s/2$ we get the $s-1$ leading order constant terms $\ln(N_E + s/2)$. The first order terms in i/N_E sum to zero because of our choice of expansion point. To second order in $s/(2N_E+s)$ the result is

$$\ln\left(\frac{(N_E+s-1)!}{N_E!}\right) \approx (s-1)\ln(N_E+s/2) - 1/2 \sum_{i=1}^{s-1} \frac{(i-s/2)^2}{(N_E+s/2)^2}. \tag{4.59}$$

We see that the leading term in this approximation combined with the two factors that are already factored out give precisely the same high temperature limit as calculated for the general case, (4.37).

It is worthwhile spending a moment on the last term in (4.59). The sum can be calculated but we approximate it with an integral and get

$$\sum_{i=1}^{s-1} \frac{(i-s/2)^2}{(N_E+s/2)^2} \approx 1/12 \frac{s^3}{(N_E+s/2)^2}. \tag{4.60}$$

For a given energy per degree of freedom, N_E/s, this correction is proportional to size, s. A numerical comparison of the left hand side of (4.59) with the first term on the right hand side shows that this correction is rather small for $N_E/s \gtrsim 1$ and that the leading order approximation can thus be used for the logarithm, at least when we are interested in values per oscillator. What the approximately constant value of the corrections per degree of freedom also says is that the quality of the approximation calculated for 10 oscillators in Fig. 4.4 can be considered essentially universal for Einstein crystals.

The exact solution in (4.57) allows us to find an exact expression for the microcanonical temperature defined in Chap. 3 for this system. Instead of a derivative we use a finite difference over the interval $2\hbar\omega$, centered at E:

$$\beta = \frac{\ln(\rho(E+\hbar\omega)) - \ln(\rho(E-\hbar\omega))}{2\hbar\omega}. \tag{4.61}$$

Inserting (4.57) gives, with $N_E \equiv E/\hbar\omega > 1$,

$$\beta = \frac{\ln\left(\frac{(N_E+s)(N_E+s-1)}{N_E(N_E+1)}\right)}{2\hbar\omega}. \tag{4.62}$$

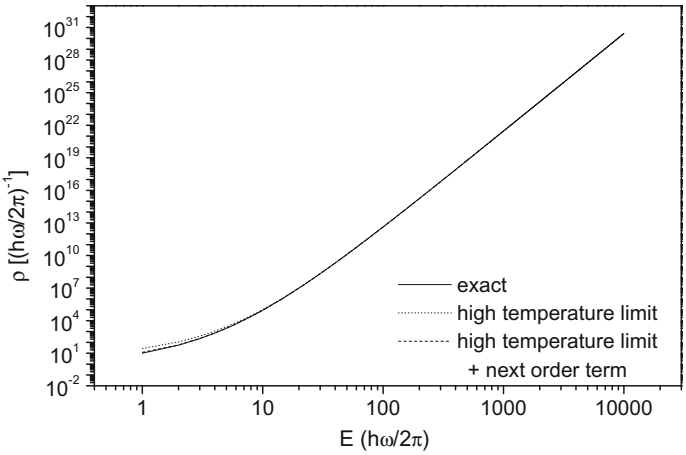


Fig. 4.4 The level density for 10 degenerate harmonic oscillators. The full line is the exact calculation and the dotted line the high temperature limit for the ten degenerate oscillators. The dashed line is the high energy result including the next-to-leading order correction, as calculated in the text. Even for energies as low as $E = 10$, which corresponds to an average of one quantum per oscillator, is the high temperature approximation rather good. Including the calculated correction makes the approximation at $E = 10$ good to a $3 \cdot 10^{-3}$ relative precision

Then we can use that $(N_E + 1)N_E \approx (N_E + 1/2)^2$ and $(N_E + s - 1)(N_E + s) \approx (N_E + s - 1/2)^2$ to get

$$\beta = \frac{\ln\left(\frac{N_E+s-1/2}{N_E+1/2}\right)}{\hbar\omega} = \frac{\ln\left(1+\frac{(s-1)\hbar\omega}{E+\hbar\omega/2}\right)}{\hbar\omega}, \tag{4.63}$$

or, equivalently,

$$T = \frac{\hbar\omega}{\ln\left(1+\frac{(s-1)\hbar\omega}{E+\hbar\omega/2}\right)}. \tag{4.64}$$

It is interesting also to consider the inverted expression, $E(\beta)$:

$$E + \frac{\hbar\omega}{2} = \frac{(s-1)\hbar\omega}{\mathrm{e}^{\beta\hbar\omega} - 1} \quad \text{(microcanonical)}. \tag{4.65}$$

This looks a lot like the analogous relation for the canonical case of s degenerate harmonic oscillators:

$$\langle E \rangle = \frac{s\hbar\omega}{\mathrm{e}^{\beta\hbar\omega} - 1} \quad \text{(canonical)}. \tag{4.66}$$

The differences are the small offset of $\hbar\omega/2$ in energy on the left hand side of (4.64), and the apparent reduction in the number of degrees of freedom, from s to $s - 1$

when going from canonical to microcanonical. This reduction was already seen in the calculation of the high energy limit. In the case of degenerate oscillators we have now seen that it works for arbitrary temperatures.

4.5 The Beyer-Swinehart Algorithm

The sections above have provided us with an approximate solution at high excitation energies for any distribution of vibrational frequencies, and an exact solution for all energies for the special case where all frequencies are identical. Those solutions cover a large selection of the problems you will encounter in practice, and since the solutions can be written down in closed form, they are particularly useful. But obviously not all situations can be reduced to one of those cases.

Fortunately there exist a couple of methods to use in other situations. One is the very powerful method known as the Beyer-Swinehart algorithm after the inventors. It is numerical in nature, must be implemented on a computer and is therefore limited to a reasonable number of oscillators. Precisely how many you will have to ask your computer about, but compared to quantum mechanical calculations of realistic frequencies, the computer requirements are usually minuscule.

The idea is very simple. The harmonic oscillators are independent and the energy can be written as a sum of energies over individual oscillators. For such a system the level density is a convolution of the level densities of the individual degrees of freedom, as was already used when we derived the Boltzmann factor and defined the temperature in Chap. 1. This means that if the set of n harmonic oscillators ω_i have a total level density of $\rho(E, n)$, one can find the total level density for the $n + 1$ oscillator system, where an oscillator with frequency ω_{n+1} has been added, as the convolution

$$\rho(E, n+1) = \int_0^E \rho(E - \varepsilon, n) \left(\sum_{k=0} \delta(\varepsilon - k\hbar\omega_{n+1}) \right) \mathrm{d}\varepsilon. \tag{4.67}$$

The sum in the bracket is the level density of the $n + 1$'th oscillator. The sum over k runs to infinity, but $\rho(E - \varepsilon, n)$ can only be non-zero when the integer k is less than or equal to $[E/\hbar\omega_{n+1}]$ ($[x]$ is the integer part of x, i.e. the highest integer not larger than x). If we integrate over ε, the result is a discrete sum over all possible energy partitionings;

$$\rho(E, n+1) = \sum_{k=0}^{[E/\hbar\omega_{n+1}]} \rho(E - k\hbar\omega_{n+1}, n). \tag{4.68}$$

This is basically it. It is a recurrence relation and is well suited for implementation on a computer. The recurrence is initialized by calculating the level density of the first oscillator as the sum of δ functions at the multiples of the energy quantum. The final result does not depend on the numbering of the frequencies, although the

intermediate level densities do. The result is exact up to the numerical precision one choses to work with. The standard check for sufficient precision is to improve the resolution until you reach a level density that does not change any more.

Because the Beyer-Swinehart algorithm is a convolution by nature, it can be easily modified to account for anharmonicities in the vibrational spectrum. The presence of anharmonicities mean that the harmonic spectrum $\varepsilon_n(k) = k\hbar\omega_n$ where k is a non-negative integer, is replaced by a general spectrum $E_n(k)$. The recurrence then reads

$$\rho(E, n+1) = \sum_k \rho(E - E_{n+1}(k), n). \tag{4.69}$$

For molecules investigated by spectroscopic methods, anharmonicities in the potential are often represented by a Taylor series in the vibrational quantum number. The leading order correction is negative in most, if not all cases, indicating a tendency for the energy levels to become more closely spaced as one moves up in vibrational energy. Such behavior is easily taken into account in the implementation of (4.69). For completeness it should be mentioned that the sum in (4.69) must be limited to terms where $E_n(k+1) > E_n(k)$. Otherwise levels move down in energy with increasing quantum number. This is forbidden by mathematical theorems on eigenvalues as well as by common sense, and you can be sure that you have stretched the fitted formula beyond its breaking point.

When anharmonicities are not important, (4.68) can be rewritten in a form which allows a considerable saving in computational effort for systems with many degrees of freedom. If one writes (4.68) with the energy $E - \hbar\omega_n$ in the argument, it reads

$$\rho(E - \hbar\omega_n, n) = \sum_{k=0}^{[(E-\hbar\omega_n)/\hbar\omega_n]} \rho(E - (k+1)\hbar\omega_n, n-1) = \sum_{k=1}^{[E/\hbar\omega_n]} \rho(E - k\hbar\omega_n, n-1). \tag{4.70}$$

Forming the difference between this equation and (4.68) at energy E gives

$$\rho(E, n) = \rho(E - \hbar\omega_n, n) + \rho(E, n-1). \tag{4.71}$$

The reduction in computing time is a factor which is on the order of the total energy in units of a typical frequency, $E/\hbar\omega$. This saving will often translate into a reduction of computer time with a factor on the order of the number of atoms in the particle, which can be important if the algorithm is used in a trial-and-error fitting of experimental data, for example.

Incidentally, the problem of counting how many ways one can distribute a certain amount of energy on a specific number of oscillators with specified frequencies is equivalent to the problem of how many ways one can change a certain amount of money. Figure 4.5 shows the calculated values for the Swedish kronor which used to have the denominations 0.5, 1, 5, 10 kr (all coins), 20, 50, 100, 500 and 1000 kr (notes). The high numbers would be very time consuming to find with a direct count.

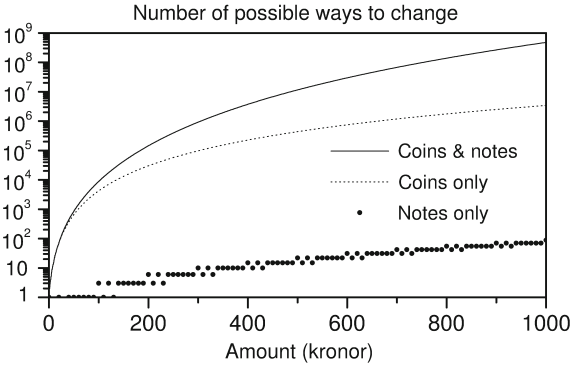


Fig. 4.5 The number of ways to change the amount of money given on the abscissa with the denominations of the money units given in the text. Also shown are the number of ways if only coins or only notes are used. The trivial change (100 kr → 100 kr) is included in the number. Clearly the number is not additive with respect to the number of different denominations used, although the logarithms seem to be approximately additive (see Exercise 4.7)

4.6 Vibrational Level Densities from Canonical Quantities

Recurrence relations are by nature very compact and easy to program but are also completely opaque; there is no way one can understand from one of these algorithms how the final result looks like and it will be very difficult to see whether one ends up with exponentially growing solutions, solutions that grow like powers of the energy or even super-exponential solutions. Fits of experimental data must then very often be performed as trial and error in the vibrational frequencies, running a code a large number of times. For molecules with a large number of vibrational degrees of freedom this can be very taxing for a computer (and for the patience of a scientist). So however much we appreciate a procedure that, like the Beyer-Swinehart algorithm, gives us exact results, it is always good to have a general and analytical approximation for a problem. A couple of more or less schematic methods were given earlier in this chapter.

It is possible to combine the transparency of analytical solutions with the accuracy of exact numerical procedures. The procedure derived in Chap. 3, which is based on inversion of the canonical partition function, (3.46 3.47) is of almost similar accuracy as the Beyer-Swinehart algorithm when the discrete nature of the excitations is ignored, and for large systems it is vastly faster because the level density for a specific energy can be calculated with a few iterations of (3.46) and a couple of algebraic operations. Because we now have the tools to judge the quality of the equation by comparison with exact results calculated with the Beyer-Swinehart algorithm, we will do this.

Figure 4.6 shows the errors in the level densities calculated with (3.47, 3.46) for the vibrational frequencies of a Lennard-Jones cluster of 55 atoms. To appreciate the precision of the result obtained with the approximate inversion formula, two

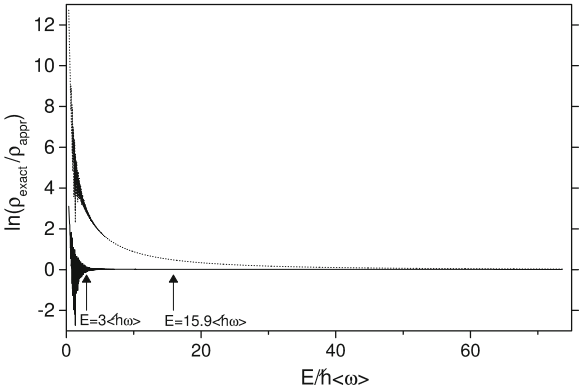


Fig. 4.6 Comparison of approximate equations for level densities with the numerically exact Beyer-Swinehart calculation for the 55 atom Lennard-Jones cluster. The full line is the ratio of the Beyer-Swinehart calculation to the values given by (3.47), and the dotted line is the Beyer-Swinehart result divided by Haarhoff's approximation (see Sect. 4.7)

arrows in the figure have been inserted at energies of $E = 0.1s\hbar\langle\omega\rangle = 15.9\hbar\langle\omega\rangle$ and $E = 3\hbar\langle\omega\rangle$. The good agreement already at the total energy content of three average quanta is quite remarkable. The oscillations below that value are due to the discreteness of the excitations and cannot be reproduced with (3.47).

Another advantage of the algorithm is that its applicability extends beyond the reign of the Beyer-Swinehart algorithm, which is designed only to handle vibrational degrees of freedom. All types of excitations can be handled with a method that converts canonical to microcanonical partition functions.

4.7 Other Computational Schemes

In addition to the method developed above there are two other functions in frequent use in the literature, due to Haarhoff and to Witten & Rabinovitch, and this chapter would not be complete without mentioning these methods. Both attempt to solve the problem of the low energy end of the spectrum, which is *the* obstacle to obtaining reliable vibrational level densities.

The idea behind the Witten & Rabinovitch method is to approximate the low energy level density of a system of oscillators with a varying, energy dependent number of oscillators. When the excitation energy per degree of freedom is reduced to values that compare with the vibrational frequencies of the particle, degrees of freedom will begin to freeze out. Vibrations with quantum energies several times the average energy per degree of freedom will be excited so rarely that they contribute only little to the total level density.

This freeze-out is mimicked by an energy-dependent reduction of the parameters that appear in the high energy limit. One of these is the energy offset, which we have seen is $1/2 \sum_i \hbar\omega_i$ in the high temperature limit. The other is the number of degrees of freedom, which appears as the power on the energy. This is sound physical reasoning, but in practice the determination of the effective values of these parameters, valid

for any molecule, is less than straightforward, and they were found by Witten & Rabinovitch by a fit to values calculated numerically for fairly small molecules. Two different expressions were needed, covering different energy regions, and containing a lower energy cut-off below which the theory is undefined. For large particles with 1000 vibrational degrees of freedom, say, the algorithm is often off by a very serious amount, and in summary it cannot be recommended for nanoparticles.

The method developed by Haarhoff is essentially a high temperature expansion. The quantity expanded is the Laplace transform of the level density, i.e. the canonical partition function. The expansion is then transformed back. Some tricks are applied to include higher order terms, and the procedure gives a single formula for the level density, in terms of scaled parameters. The theory has the appealing feature that the result is obtained in closed form, but it fails at low energies and for large particles. The low energy failure is almost unavoidable because the method is based on an expansion in powers of the reciprocal excitation energy. Given the other tools presented in this chapter (the Beyer-Swinehart algorithm for small particles and the canonical-microcanonical inversion for larger particles) the Haarhoff equation is practically obsolete.

4.8 Level Densities from Bulk Properties

We have spent some time calculating level densities based on the assumption that excitations can be build by vibrational excitations, and sometimes even harmonic ones, to boot. This is not completely realistic. There are usually a number of effects that will destroy this simple picture. One is thermal excitation of valence electrons. This effect is treated in detail in Chap. 10. Another is anharmonicities in the vibrational motion of the nuclei. This effect can be treated with a modification of the Beyer-Swinehart algorithm, as mentioned above, but only up to some temperature. Yet another effect is the very strong anharmonicities associated with melting.

For a number of elements these effects can be quite strong. Heat capacity measurements of bulk materials at elevated temperatures will often show values higher than the Dulong-Petit value of $3k_B$ per atom, even below the melting temperature. Figure 4.7 illustrates the problem with plots of experimentally measured bulk heat capacities of several elements.

Often one does not want to attempt to account for these effects one by one but simply chose to extrapolate the bulk thermal properties to finite sizes. There is no guarantee that this will give a good representation of the finite size particle, of course, and the judgement must ultimately be made by experiments. If one does chose to do the extrapolation, level densities can be extracted with simple means now we have established a connection between canonical values and level densities.

Thermal properties will often be represented by a caloric curve, $E(T)$ or by a table of heat capacities vs. temperature, $C_p(T)$, i.e. the heat capacity at constant and ambient pressure and not at constant volume, $C_v(T)$. For condensed matter the difference is small and we will ignore it and use C_v to denote the measured

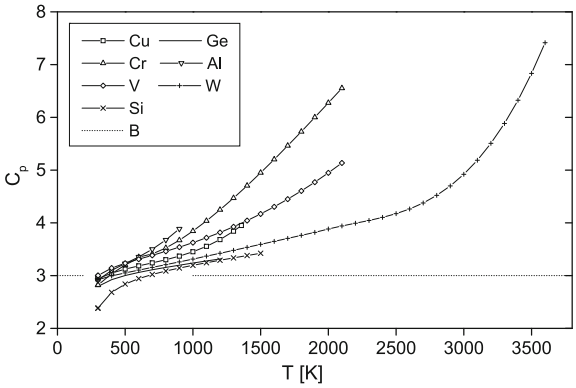


Fig. 4.7 Experimentally measured value of the heat capacity per atom for some metallic elements. The curves start at room temperature and end at the melting points. The dotted line is the Dulong-Petit value. Measured values are heat capacities at constant pressure, not volume, hence the subscript p. The difference between the two for a condensed phase is usually marginal

values. By division with the number of atoms one gets the values per atom, $e(T)$, and $c_v(T)$. To get the corresponding curves for a particle of N atoms one would to a first approximation multiply with N. For free particles this overestimates the number of degrees of freedom by 6, viz. the number of d.o.f.'s that represent translational and rotational motion and that are not present as vibrational motion. Effectively there are two thermal atoms fewer than inertial atoms in a free particle. The reduction of the effective number of atoms holds only if the heat capacity is carried predominantly by the motion of the nuclei, but this should be a good approximation in most cases. The particle thermal properties are therefore given as

$$\overline{E} = (N-2)e(T), \tag{4.72}$$

$$C_v = (N-2)c_v(T), \tag{4.73}$$

$$S(T) = (N-2)\int_0^T \frac{c_v}{T'}\mathrm{dT}', \tag{4.74}$$

if only internal degrees of freedom are included. With these input functions all ingredients are in place to use the connection between canonical properties and microcanonical level densities given in (3.46, 3.47) that can be used without any further ado.

Exercises

4.1 Show that the mode in (4.16) is indeed a zero mode for its Hamiltonian.

4.2 Show that the heat capacity for harmonic oscillators is given by

$$C = \sum_i \left(\frac{\hbar\omega_i}{T}\right)^2 \frac{\mathrm{e}^{-\hbar\omega_i/T}}{\left(1-\mathrm{e}^{-\hbar\omega_i/T}\right)^2}, \tag{4.75}$$

and show that the leading order high temperature limit is independent of the frequencies. This limit is the Dulong-Petit law when given for a mole. Calculate this value in units J/K mol.

4.3 Find the leading order correction to the low temperature T^3 heat capacity of a Debye particle.

4.4 Find the high energy entropy of a Debye crystal.

4.5 Use the Debye temperatures to find the speed of sound in five different elements of your choice.

4.6 We will try to verify the calculation of the level density of degenerate harmonic oscillators. For this purpose we will use induction and the solution in (4.57) in the dimensionless form with $N = E/\hbar\omega$.

Let's first verify that the equation holds for an arbitrary positive integer N and one oscillator, $s = 1$. Since $0! = 1$, one calculates one possible state, which is indeed true. The question is whether it is true for $s + 1$ if it is true for s. To show this, we calculate the convolution of the number of states at energy N for s oscillators with the corresponding number for a single oscillator. This should give number of states at energy N for $s + 1$ oscillators. (Note that if you want to check this formula with (4.57) that the values of the s that appear in the two equations differ by one.) Use that the level density of a single harmonic oscillator is a sum of equidistant δ functions to show that the number is

$$\sum_{i=0}^{N} \frac{(N+s-1-i)!}{(N-i)!(s-1)!} \tag{4.76}$$

To calculate this sum, first prove the relationship between ratios of factorials:

$$\frac{(p+1)!}{m!(p+1-m)!} = \frac{p!}{m!(p-m)!} + \frac{p!}{(m-1)!(p+1-m)!} \tag{4.77}$$

where p and m are arbitrary non-negative integers with $p \geqslant m$. With standard notation this can be written as

$$\binom{p+1}{m} = \binom{p}{m} + \binom{p}{m-1}. \tag{4.78}$$

Use this relation recursively to show that

$$\binom{p+1}{m} = \sum_{j=0}^{m} \binom{p-j}{m-j}. \tag{4.79}$$

Make the proper identifications and show that this is the desired result.

4.7 The number of ways you can change a certain amount of money is illustrated in Fig. 4.5 with values calculated with the Beyer-Swinehart algorithm. For Swedish currency the number depends on the different denominations of coins and notes put into circulation by Sveriges Riksbank. From the figure it seems that the number obtained for coin-changes-only and the number for notes-changes-only multiply approximately to the correct number, i.e. they are multiplicative, not additive.

You can show that this is expected, under some conditions. Assume that the coins-only curve, $C(A)$, is proportional to A^α, and the notes-only curve has a similar behavior, $N(A) \propto A^\beta$, where A is the amount and α and β are dimensionless numbers. Calculate the total number of possible changes. What is the condition α and β must fulfill for the logarithms of the two curves to be strictly additive, and what happens if this condition is not fulfilled?

4.8 Integrate (4.71) with a Boltzmann factor to find a recurrence relation for the calculation of the corresponding canonical partition functions.

4.9 The subject of this problem is to derive an alternative expression for the harmonic oscillator level density in (4.37). Like that equation, it will be valid for high energies. The strategy is to take the known canonical partition function and approximate it with a form for which the level density is known. Step one is to multiply and divide the canonical partition function with the partition function for the same number, N, of harmonic oscillators but with a common frequency, ω'. Next, multiply and divide by the ratio $\prod_{i=0}^{N} \exp(-\beta(\omega' - \omega_i)/2)$, where the ω_i's are the frequencies of the problem. Then expand the factors $(1 - \exp(-\beta\omega))\exp(\beta\omega)$ in $\beta\omega$. This gives factors of the form $\beta\omega + (\beta\omega)^3/24 + \cdots$ (with subscripts or primes, as it may be). Extract the factors $\beta\omega$. This gives factors of the type

$$\frac{\omega'}{\omega_i}\frac{1 + \frac{1}{24}(\beta\omega_i)^2}{1 + \frac{1}{24}(\beta\omega_i)^2}. \tag{4.80}$$

Expand the product of these to first order in $(\beta\omega)^2$ and set the coefficient of β^2 to zero. This defines the choice of ω'. Show that the canonical partition function then becomes

$$Z \approx (1 - \mathrm{e}^{-\beta\omega'})^{-N} \prod_{i=1}^{N} \mathrm{e}^{\beta(\omega_i - \omega')/2} \frac{\omega'^N}{\prod_{i=1}^{N} \omega_i}, \tag{4.81}$$

with the choice $\omega' = \sqrt{\sum \omega_i^2/N}$. Use this to extract the level density.

Chapter 5
Rate Constants for Emission of Atoms and Electrons

One of the most important types of processes in the physics of free nanoparticles is the emission of energy in some form. Particles flying around in vacuum with excitation energies beyond those that correspond to ambient temperature will sooner or later emit their excess internal energy. This is bound to happen, because even in vacuum will the particles interact with the rest of the world by radiation and sooner or later equilibrate to the temperature they see from the walls of the vacuum chamber. But emission (and absorption) of photons is not the only possible relaxation mechanism. Particles can also emit massive fragments like electrons, atoms or small molecules. Emission of massive particles or electrically charged particles is much easier to detect than photon emission. Lose one photon and you may or may not (usually not) change the external attributes of the particle a lot. Lose one electron and the motion in electric and magnetic fields is completely changed.

Although equilibrium will set in sooner or later, 'later' may be after a very long time. This open the possibility to use the quasi-equilibrium processes as diagnostic tools, and it is therefore of major interest to be able to describe the behavior of the particles before any external equilibrium can be established.

The general framework for the description of dissipation of excess energy is provided by the theory of unimolecular decay. As the name suggests, unimolecular decay involves only one precursor molecule, that is, the reaction is not provoked by, for example, the collision of two molecules. This chapter will teach you the basics of the theory for these types of decay. It is implicit in all the following considerations that the internal equilibration times in these reactions is much shorter than any equilibration time due to interactions with the surroundings.

The possible thermal reactions are, representing particles as composed of a single species of monomers, X, for simplicity:

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_5

$$X_N^q \rightarrow X_{N-M}^q + X_M \qquad \text{evaporation} \tag{5.1}$$

$$X_N^q \rightarrow X_N^{q+1} + \mathrm{e}^- \qquad \text{thermionic emission} \tag{5.2}$$

$$X_N^q \rightarrow X_N^q + h\nu \qquad \text{radiation} \tag{5.3}$$

$$X_N^q \rightarrow X_{N-M}^{q'} + X_M^{q-q} \qquad \text{fission} \tag{5.4}$$

Only the first two will be treated in this chapter. In the chemical literature, the theory describing this type of reactions is often referred to as transition state theory. A separate chapter is devoted to thermal radiation. Fission processes which, like the other processes in the list, can be partly direct and partly thermal, will not be treated at all. There are applications where fission plays an important role, as for example in electrospray ion sources, which is a very gentle method for producing gas phase ions, but their description usually involves only thermal processes as a minor ingredient.

The starting point for a theory of something changing, $A \rightarrow B$, is to be able to tell A and B apart. Is it a meaningful question to ask if this particle has lost an atom during the preceding 10 μs? The answer is yes. We may not know the answer if we didn't do the measurement, but at least we can tell the difference between a particle containing 92 potassium atoms and one containing 91. There is only a small part of the total number of states explored by a free particle that cannot be assigned either one of these two labels. As another example, consider a neutral W_{10} cluster which emits an electron 3 μs after being hit by a light pulse from a laser. During perhaps 10 fs the electron is on the way out, not part of the cluster but not yet free to do as it pleases. We cannot really say if the cluster is neutral or charged. After the electron has been emitted, the cluster is charged and flies through a Time-of-Flight mass spectrometer for 50 μs and is then detected. The bottom line is that the state of uncertainty for us only lasts a very short time compared to all the time the cluster spends as an object with well-defined mass, charge and energy.

In terms of phase space, there are therefore different parts that can each be labelled practically unambiguously by the size of the particle, by the charge state and by the energy. Considering a single, specific reaction, this implies that we can divide the phase space into a reactant (or parent) part and a product (or daughter) part. Unimolecular reaction theory describes quantitatively the rate constant of crossing from the reactant part of this phase space to the product part. The idea is illustrated in Fig. 5.1.

Figure 5.1 is also intended to suggest that a decay is a rare process, properly understood. The system rattles around for a long time before it manages to hit the small connection between the two parts of phase space. Time constants of microseconds are routinely encountered in experiments, and although this seems extremely fast at first glance, it is not. The classification of fast or slow should be made by comparing to an equilibration time, i.e. the time it takes for an initial excitation to be dissipated into all possible degrees of freedom. If the time constant is long compared

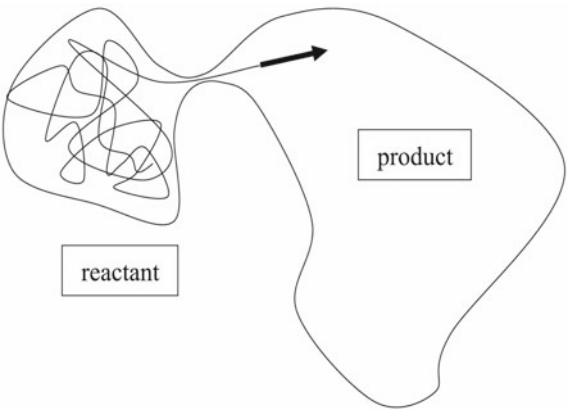


Fig. 5.1 A schematic representation of the phase space describing a unimolecular reaction. The process for which we want to determine the rate constant is indicated by the arrow

with this, the decay is slow, by definition.[1] A slow decay ensures that the decaying system samples a representative part of the phase space before it decays. This is known as complete statistical mixing. In nuclear physics, which has contributed a lot to the development of the contents of this chapter, this is known as the formation of a compound nucleus.

Statistical mixing or compound particle formation means that decays occur independently of the way the excitation energy was brought into the particle. When the particle decays, is has effectively forgotten how it was excited initially, apart from the conserved quantities of energy, momentum and angular momentum.[2] In this situation the decay is exponential

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -kP, \tag{5.5}$$

where P is the population and k is the rate constant we want to calculate in this chapter.

Before we start deriving equations for k, it is worth emphasizing the background for the exponential decay. In spite of the frequent occurrence and even more frequent use (and occasional abuse), it actually only pertains to the special situations where the system decays out of a single state. In radioactive decay, such single initial states are the norm because decaying nuclei tend to exist in a specific quasi-stable quantum state. It is also observed in thermal decay processes, in situations where the energy distributions are continuously being reestablished by collisions and whatever form of mechanisms one has for exchange of energy and other quantities. Then the population has a practically constant distribution of all the quantities that determine a reaction

[1] This is an example of the general principle that only dimensionless quantities can be considered big or small. Big or small, fast or slow or any other quantitative adjective always refers to some, often implicit, reference value.

[2] The linear momentum will turn out to be conserved automatically by implementation of energy conservation because it is a two-body problem.

rate, in particular the energy. Then the decay occurs, by definition, from a single state, although this is not a pure eigenstate of the Hamiltonian.

In settings where the rate constants compare with or are higher than the equilibration times, the situation changes. Precisely how decays then proceed will be the subject of a later chapter. On the other hand, it is important to keep in mind from the start that this does not in itself invalidates the characterization of the decay of a single, isolated particle as exponential in time.

The decay of an isolated particle with complete statistical mixing therefore belongs to the special class for which exponential decays occur, and it make good sense to continue with the development of the theory for this type of decay.

5.1 Atomic Evaporation

The first rate constant that will be calculated describes atomic evaporation. The calculation will serve as a template for the treatment of other types of decay, and we will go into some detail with the derivation. To define the involved phase space, imagine you enclose the system in a box with walls that reflect all particles when they hit and that you then leave the system alone for long enough for the particle to evaporate an atom and recombine with it a large number of times. We have a stationary distribution which, on the average, does not change with time. If probed at some random time, the system will then be found in the two states, 'reactant' and 'product', according to the statistical weights of these two parts of the system. The statistical weight of one of these systems is the number of quantum states available at a certain energy or, in the classical description, of the volume of phase space assigned to the state. Correspondingly, they are the level densities times some small energy interval, $\rho(E)\delta E$.

Stationarity requires that the rate of going from one part of phase space into the other is the same as the rate going the other way. These rates are the rate constants times the respective populations. We therefore have

$$\frac{\rho_P(E)}{\rho_P(E)+\rho_r(E)}k_{formation} = \frac{\rho_r(E)}{\rho_P(E)+\rho_r(E)}k_{decay}, \tag{5.6}$$

where subscripts P, r refer to product and reactant states, respectively. The product state includes two particles (of size $N-1$ and 1), hence the capitalized P. The argument leading to (5.6) is the principle of detailed balance. The principle is founded on microscopic reversibility which in turn expresses that the modulus of a matrix element for a process is the same as for the inverse process. It is the basis of the second law of thermodynamics. The second law of thermodynamics states that the entropy of an isolated system increases with time or remains constant when the system is in a steady state or during a reversible process. In equilibrium, regions of phase space are populated proportional to their volume, and this is also what gives the maximum entropy for the given constraints. Violation of detailed balance would

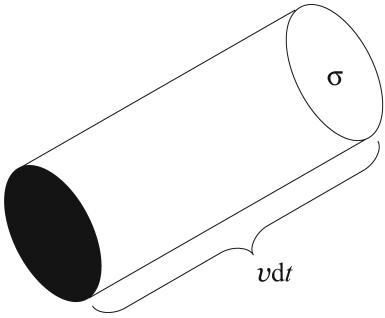


Fig. 5.2 The capture rate. The concentration of the particle is $1/V$ which, together with the volume covered during the time interval dt gives the probability of capture during that time interval

make it possible for a system to develop away from that distribution, toward lower entropy.

It is therefore hard to get around (5.6). The only requirement that must be fulfilled for the use of detailed balance is that decay and formation happens statistically, because only then can level densities be expected to represent the states of the system.

The ratio of level densities in (5.6) is the normalized population of the two species. If we know one of the rate constants $k_{formation}$ or k_{decay} and the level densities, clearly we know the other rate constant. It so happens that we can calculate the rate constant for formation. For the evaporation of a single atom from a particle the inverse process has a formation rate constant of

$$k_{formation} = \frac{1}{V}\sigma(v)v, \tag{5.7}$$

where v is the relative speed and σ is the cross section for fusion of the atom and the product particle, and $1/V$ is the concentration of the particle. The calculation is illustrated in Fig. 5.2.

As indicated in (5.7), the capture cross section may depend on the relative speed, or kinetic energy, of the two particles. One may also specify the relative angular momentum, which has some advantages when the problems are formulated quantum mechanically, but we will stick with the energy dependence. It should be stressed that the use of the cross section is not just a fancy way of introducing a fit parameter with no physical importance. The cross section is a real physical quantity that can be measured in independent experiments. It is sometimes parametrized in terms of a sticking coefficient which is the ratio of the attachment cross section to a geometric cross section. A convenient definition of a geometric cross section for X_{N-1} and X_1 colliding is $\sigma_{geo} \equiv \pi r_1^2(1 + (N-1)^{1/3})^2$, where the monomer radius r_1 is related to the bulk particle density, ρ_b, by the relation $\rho_b = 3/(4\pi r_1^3)$, provided the densities are identical to bulk densities. A little more general it can be written as $\sigma_{geo} = \pi(r_{N-1} + r_1)^2$.

However, we also know that cross sections need not refer to some geometric size values calculated from tabulated bond lengths or bulk densities. It can easily be different, both larger and smaller. A cross section may be exactly zero below a certain collision energy. This will be the case when you have a threshold for

attachment of the atom to the product. This situation is often found in chemical reactions and is dealt with by adding the barrier height to the activation energy for the process but can equally well be understood in terms of the cross section. Another type of cross section arises when the two fragments interact with attractive long range potentials. Section 5.6 will present calculations for several simple intermolecular potentials of both attractive and repulsive types and relate them to kinetic energy release distributions.

The classical approximation for the motion of the small fragment in the reverse process of attachment is usually a good description for atoms or heavier fragments colliding with a large particle, because the de Broglie wavelength of the atom is much smaller than the dimension of the particle (see Exercise 5.9), and we rarely need to worry about quantum corrections to the classical collision cross section.

With these caveats we can combine (5.6, 5.7) to get the decay rate constant:

$$k(E) = \frac{1}{V}\sigma v \frac{\rho_P(E)}{\rho_r(E)}. \tag{5.8}$$

We have eliminated the subscript *decay* on k because from now on we will only deal with this constant, unless otherwise mentioned. Equation (5.8) is almost the desired result, but we can be a little more specific. The level density of the product, ρ_P, is that of the system comprising the free atom and the product particle. We know the density of states for a free atom with mass m and degeneracy g which was given in (2.63):

$$\mathrm{d}n = g\frac{\mathrm{d}\overline{x}\,\mathrm{d}\overline{p}}{h^3}, \tag{5.9}$$

corresponding to a free particle translational level density of

$$\rho\,\mathrm{d}\varepsilon = g4\pi\sqrt{2}\frac{Vm^{3/2}}{h^3}\varepsilon^{1/2}\mathrm{d}\varepsilon, \tag{5.10}$$

where ε is the kinetic energy of the atom. Because we are dealing with a two-body problem, the mass m is the reduced mass of the atom and the product particle, and ε is the sum of the two translational energies in the center of mass system. Often these can be approximated by the values for the atom alone. The factor g is any degeneracy the atom may have, for example due to an intrinsic angular momentum. If we specify the rate constant with respect to the kinetic energy of the atom, the product level density is

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\rho_P(E,\varepsilon) = g4\pi\sqrt{2}\frac{Vm^{3/2}}{h^3}\varepsilon^{1/2}\rho_p(E - E_a - \varepsilon), \tag{5.11}$$

where ρ_p is the level density of the size $N - 1$ product particle. This level density denotes the internal degrees of freedom of the product particle alone. The parameter E_a is the activation energy required to remove an atom from the reactant particle.

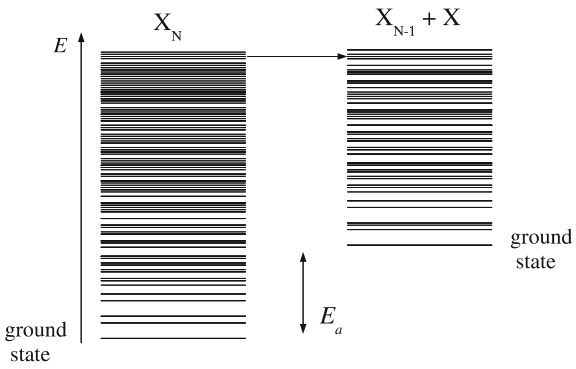


Fig. 5.3 The energetics of unimolecular decay. Vibrational levels are indicated schematically by the horizontal lines. The arrow indicates the decay at constant total energy. The kinetic energy release and the internal degrees of freedom of X are set to zero for simplicity. When added, they will tilt the arrow slightly downward. The activation energy is the difference of ground state energies, as indicated. By energy conservation, the excitation energy of the product is reduced by this amount

This energy, and the kinetic energy ε, are both extracted from the internal degrees of freedom of the particle and hence must be subtracted from E in the argument of the level density. Figure 5.3 illustrates the situation.

Together with (5.8) the level density in (5.11) gives the rate constant for atomic evaporation

$$k(E,\varepsilon)\mathrm{d}\varepsilon = g\frac{m}{\pi^2\hbar^3}\sigma(\varepsilon)\varepsilon\frac{\rho_p(E-E_a-\varepsilon)}{\rho_r(E)}\mathrm{d}\varepsilon. \tag{5.12}$$

The volume has disappeared from this expression, which is certainly comforting. On physical grounds we do not expect some arbitrary volume to influence the evaporation rate constant (if it is not too small, see Exercise 5.2). The volume would have been relevant if we had instead considered the relative populations of N and $N-1$, because the translational entropy depends on the volume.

Equation (5.12) is the Weisskopf formula for evaporation of particles from ... whatever. It was originally derived to describe neutron evaporation from excited nuclei but is rather general and equally good for nanoparticles. There is no size or energy restrictions on the applicability of detailed balance. The application of the formula is particularly simple for particles that interact with a spherical or almost spherical potential, because the capture cross section, which may not be available as an experimentally measured parameter, can then be modelled. If the capture cross section is known there is little reason not to use this formula.

We need to account for momentum conservation. One may wonder why energy conservation is implemented but conservation of linear momentum is not in (5.12). Surely linear momentum is as conserved as energy! But this conservation law is actually taken into account implicitly. Formally one could write the level density of the final state as a phase space integral over the two particles in the final state and implement momentum conservation with a δ function in the difference between the

initial and final momenta. Integrating over one of the particles' momentum, that of the heavy particle say, will cancel that momentum and the δ function and leave us with the phase space integral over one momentum. But this is exactly what we are doing when we interpret the mass and the kinetic energy as pertaining to the outgoing channel and not to the atom, so we do not need to go through that particular piece of x-erei.[3]

We can integrate over the kinetic energy of the decay channel with the use of the microcanonical temperature of the product (daughter) particle, T_d. We expand the level density of the product

$$\rho_p(E - E_a - \varepsilon) \simeq \mathrm{e}^{-\varepsilon/T_d}\rho_p(E - E_a). \tag{5.13}$$

With the definition

$$\langle\sigma\rangle \equiv \frac{\int_0^E \sigma(\varepsilon)\varepsilon \exp(-\varepsilon/T_d)\mathrm{d}\varepsilon}{\int_0^E \varepsilon \exp(-\varepsilon/T_d)\mathrm{d}\varepsilon}, \tag{5.14}$$

we get the total decay rate

$$k(E) = g\frac{m}{\pi^2\hbar^3}\langle\sigma\rangle T_d^2 \frac{\rho_p(E - E_a)}{\rho_r(E)}. \tag{5.15}$$

For cross sections that do not vary too strongly with energy, the integrand in (5.14) peaks around $\varepsilon \approx T_d$, and in these cases we have $\langle\sigma\rangle \approx \sigma(T_d)$. This is useful for a first estimate and often also sufficient for the second, but should be used only for the calculation of the total rate constant. Describing kinetic energy release distributions requires that the energy-specified cross section is retained.

5.2 Rate Constants with Microcanonical Temperatures

The ratio of the level densities in (5.15) is the most strongly varying factor of any unimolecular rate constant, not just for atomic evaporation, but equally well for thermal emission of both larger and smaller particles. As discussed in Chap. 3, this ratio can be approximated with a Boltzmann factor to give an Arrhenius rate constant if the temperature used is the finite heat bath temperature, roughly the average of the precursor and the product microcanonical temperatures. It also requires a redefinition of the activation energy. The expression is, with the prefactors calculated here, given by

$$k(E) = \frac{mg}{\pi^2\hbar^3}\langle\sigma\rangle T_d^2 \frac{a_p}{a_r}(E + E_r - E_a')^{s_p - s_r}\mathrm{e}^{-E_a'/T}, \tag{5.16}$$

[3] Einstein's expression for unnecessary use of formulae. Wordplay on 'Hexerei' (witchcraft).

where the parameters a_i, E_i $(i = r, t)$ are defined through the level density $\rho_i = a_i(E + E_i)^{s_i}$, $E'_a \equiv E_a + E_r - E_p$, and T is defined as

$$T = \frac{E + E_r - E'_a/2}{\overline{s}} - \frac{1}{12\overline{s}}\frac{E'^2_a}{E + E_r} + \frac{1}{24\overline{s}}\frac{E'^3_a}{(E + E_r)^2} + \cdots, \tag{5.17}$$

with $\overline{s}$ the average of s_r and s_p. m is the reduced mass and g the degeneracy of the atomic fragment.

As already mentioned, a few points should be kept in mind when using these equations. One is the cross sections which, as already mentioned, do not necessarily have values equal to any simply defined geometric size. The second is the reliability of the level density calculations. Very precise calculations of harmonic oscillator level densities may be completely invalidated by anharmonicities and other effects and, as we have seen from Fig. 4.7, anharmonicities do occur.[4]

It is possible to test the validity of the formula for ω by comparison with macroscopic data. The factor in front of the ratio of the level densities needs to be integrated over ε to be understood as a frequency factor, ω. For a comparison with the bulk data the capture cross section must be set energy independent and geometric, assuming that a gas molecule sticks to a surface of the same material with unit probability, i.e. has a sticking coefficient of one. This is usually the case. Zinc behaves differently, but this is one of the few cases where soft like-on-like collisions do not lead to automatic fusion. The frequency factor then becomes

$$\omega_N = \frac{mg}{\pi^2\hbar^3}\sigma T_d^2 \frac{a_p}{a_r}(E + E_r - E'_a)^{s_p - s_r} \tag{5.18}$$

where the capture cross section is the geometric, $\sigma = \pi r_N^2 = \pi r_1^2 N^{2/3}$. We will absorb g into the a-coefficients and use the harmonic oscillator values, $a_p/a_r \approx (\hbar\omega_v)^3(3N - 8.5)^3$. Together with the energy in the parenthesis and the fact that the reduction in heat capacity, $s_p - s_r$ is 3 for harmonic oscillators and evaporation of a single atom, we get

$$\omega_N = \frac{m}{\pi^2\hbar^3}\pi r_1^2 N^{2/3} T_d^{-1}(\hbar\omega_v)^3, \tag{5.19}$$

where ω_v is the geometric average of the vibrational frequencies, assumed identical for reactant and product.

This value can now be compared with observed macroscopic values, as derived from measured vapor pressure data. With the parametrization

$$P = A\mathrm{e}^{-B/T}, \tag{5.20}$$

the empirical parameter A is related to ω. Using the ideal gas law to relate the condensation and evaporation in equilibrium gives

[4]This is a case where the distinction between precision, accuracy, and reliability is essential (see preface to first edition).

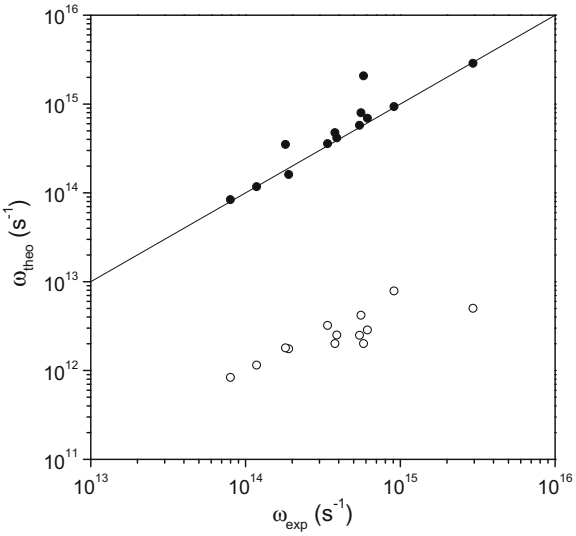


Fig. 5.4 The calculated frequency factors, exclusive of the $N^{2/3}$ factor, for some metallic elements calculated with (5.19) vs. the experimental values, converted from vapor pressure data with (5.22). The line is the ideal 1:1 relationship. Also shown, as open circles, are the Debye frequencies, $\nu = T_D k_B / h$

$$\omega \mathrm{e}^{-E_a'/T} = \frac{N}{V}\overline{v}S = \frac{P}{T}\overline{v}S = \frac{A}{T}\overline{v}S\mathrm{e}^{-B/T}, \tag{5.21}$$

where $\overline{v}$ is the average thermal speed of an atom in the gas, and S a surface area. Cancelling the exponentials we get

$$\omega = \frac{A}{T}\overline{v}S = A\sqrt{\frac{8}{\pi m T}}S. \tag{5.22}$$

Values for ω calculated with (5.22) and (5.19) are shown in Fig. 5.4 for a number of elements, plotted against each other. The elements shown are the ones for which the experimental data have been fitted with the equation $P = A\exp(-B/T)$ in the range from room temperature to the melting point. The two equations for the ω's have different temperature dependences, but both are weak, and the values for the figure are calculated with the midpoint of the range used to fit A. The areas S and σ were both set to πr_1^2, i.e. the area of one atom, with r_1 calculated from the bulk density.

The values for the ω's given by the Debye frequency are also shown in the figure from the same elements. For the vibrational quantum energy $\hbar\omega_v$ in (5.19), the values $\hbar\omega_D \mathrm{e}^{-1/3}$, calculated in (4.55), were used. All Debye temperatures were below the lowest fitted range, except for one (Li, with $T_D = 376$ K). The vapor composition is in all cases shown measured to be overwhelmingly or exclusively monomers. Note that the different notation for the Debye frequency and the observed/calculated frequency factors is purely notational and no factor of 2π has been left out in the comparison.

The agreement between calculated and measured values is fairly good, with only a single outlier (tin). The good agreement is partly due to the relatively low temperatures; all the data points refer to solids. At higher temperatures, if not before than at least above the melting point, one expects deviations from any calculation that uses harmonic vibrations as an essential input.

In heuristically justified theories the value of ω is often identified with the typical vibrational frequency of the material, often represented by the Debye frequency, although this is manifestly not correct. The misconception has been kept alive by the fact that the two time scales differ by only two orders of magnitude. A difference of this magnitude can easily be hidden in experimental uncertainties if the data are measured on millisecond time scales, say, but Fig. 5.4 leaves no doubt that the Debye frequency gives a poor prediction for the frequency factor. This should not surprise; it is not designed for that purpose. Below we will see calculations of even much larger frequency factors which you need to be very crude, even on a logarithmic scale, to approximate with any intrinsic vibrational frequency in the particles.

The fact that ω is often larger than the vibrational frequency also means that it will be larger than the inverse thermal equilibration time, or the time needed for an excitation to dissipate into accessible degrees of freedom, because equilibration times cannot be shorter than the typical vibrational period. Rate constants that are higher than the inverse of this time cannot therefore be expected to be correctly described by (5.12). If one takes the hint from nuclear physics, the breakdown of the complete mixing assumption occurs when the energy per degree of freedom is a good fraction of the activation energy, E_a, 0.3–0.4. This suggests that (5.12) is good after equilibration times of 0.1–10 ps and longer. If the system decays before statistical mixing is achieved, the description depends on more system specific parameters and must include some dynamics. Also the precise time for the crossover from a dynamics-specific decay to a statistical decay depends on the system. We will continue to work with the assumption of complete statistical mixing unless explicitly mentioned.

5.3 Large Fragments

When the small fragment evaporated is a molecule instead of an atom, it will contain internal degrees of freedom beyond just degeneracies, viz. rotational, vibrational and electronic. Some of these will be activated, in particular rotations, but often also vibrational motion and occasionally even electronic motion, for low-lying excited electronic states. This is the meaning of the word ‘larger’ here; more than just translational degrees of freedom. With the low quantum energies of rotations, large is almost automatically equivalent to two or more atoms. The degeneracy factor g in (5.12) is a special case of a small fragment with internal degrees of freedom. It is energy independent and just gives a multiplicative factor. This special case need not change our definition of large.

Any degree of freedom which is thermally activated will contribute to the total level density for large fragment and add to the entropy of the final state. This will usually increase the decay rate constant, although this increase will be reduced by the loss of degrees of freedom of the larger fragment. The rotational degrees of freedom play an important role here, beyond their low quantum energies. If we calculate the degrees of freedom before and after the decay and for simplicity assume all rotational and vibrational degrees of freedom are active and electronic are not, we see that we lose some vibrational degrees of freedom and gain some rotational. For two products emerging from the single reactant, all considered non-linear, the changes are -6 for the vibrational and +3 for the rotational degrees of freedom. For a single linear product, as produced in dimer evaporation, the changes are -5 and +2.

Although the bulk part of the level density is vibrational, the loss of a few of those usually does not change it by a significant amount compared with the gain in entropy from the appearance of just a few more rotational degrees of freedom, because of the large ratio of the quantum energies of the vibrational and the rotational degrees of freedom. The larger the fragments, the larger the moment of inertia and therefore the smaller the rotational constants and the higher the rotational entropy. For decays of e.g. biological molecules which can break up into fragments composed of tens of atoms or more, the gain in rotational entropies can be very large.

Before we embark on a quantitative calculation of these effects, it is necessary to consider to which degree detailed balance is capable of describing such a situation. The critical point is the question whether or not the inverse reaction capture cross section is easily parametrized and does not require a geometric factor. If present such a factor can potentially suppress the cross section by a very large factor if, for example, precise orientations of both molecules are needed in the collision.

To formulate the question quantitatively, we note that the Ansatz of compound particle formation does not imply that the fragments need to fuse on the time scales that characterize a collision. The requirement is less severe, namely that the product is incorporated into the final state of the particle/cluster/molecule and equilibrates to internal degrees of freedom before it decays again. Otherwise a calculated rate constant would not describe the decay of the fused state, i.e. the state we are interested in describing the decay of. This is illustrated schematically in Fig. 5.5.

By the same token, decay processes need not happen in one step but can occur through a partial detachment of the fragment followed by a complete dissociation at a later time. One advantage of the detailed balance description is that such details of the process need not be known. Complications only arise if the intermediate configurations carry appreciable statistical weight relative to the total density of states, in which case the problem reduces to the technical problem of including them into the level density.

For a more quantitative consideration let's reduce the question to the bare bones, and make the simplifying assumptions that there is a single intermediate state, denoted by i, and it is similar to the ground state, denoted by 0, with respect to vibrational properties, specifically the level density. The indicator i can also refer to isomer, which is the standard cluster notation for a configuration different from the ground state and for which all the vibrational frequencies are real, corresponding

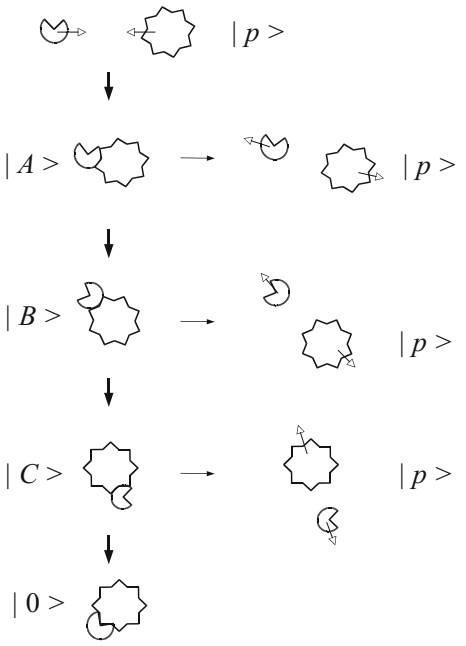


Fig. 5.5 An illustration of the possible capture of a fragment state, $|p\rangle$, and the establishing of the final bond between the two pieces into state $|0\rangle$, with an range of possible intermediate dizzy state of loose bonding, indicated as $|A\rangle$, $|B\rangle$, $|C\rangle$. As long as the re-dissolution rate of these, indicated by the small arrows pointing right, is small, the capture cross section is close to the one for formation of the dizzy state

to the absence of negative eigenvalues in the Hessian. Chemists prefer to call them conformers. We can group the states $|A\rangle$, $|B\rangle$, $|C\rangle$ into one intermediate, $|i\rangle$. Then the whole reaction scheme for attachment and dissociation is

$$|p\rangle \rightleftarrows |i\rangle \rightleftarrows |0\rangle. \tag{5.23}$$

The attachment reaction moves from left to right, and the dissociation back again. The dissociation may happen after several attempts where state $|i\rangle$ is populated and decays to $|0\rangle$ without causing any modifications of the theory. The crucial question is if such processes also happen for $|p\rangle$ and $|i\rangle$.

An example of a situation approximately described by this reaction scheme is the attachment and evaporation of a water molecule to a water cluster. Water molecules are highly polarized and the bond between the hydrogen atoms and an oxygen atom of another molecule, aptly named the hydrogen bond, is a fairly large part of the total binding energy of a molecule to the cluster. Molecules in the surface of a water cluster can break one of the hydrogen bonds and have the hydrogen atom dangling out in vacuum, while still being bound to the cluster by the other. This corresponds roughly to the isomeric state, although the two vibrational frequencies of the dangling hydrogen may not be positive and the state not truly isomeric in the above sense of the simplified reaction in (5.23), albeit hopefully close enough for the analogy to work. Figure 5.6 depicts the situation and defines the energies.

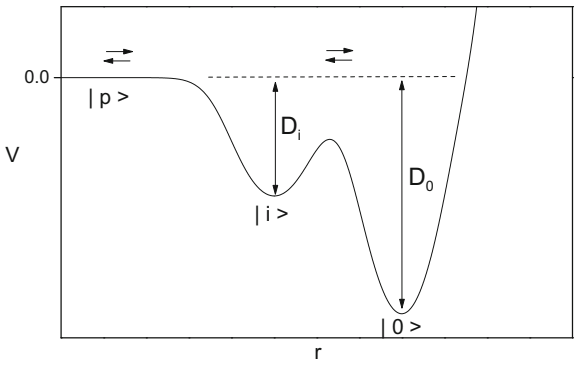


Fig. 5.6 A schematic illustration of the potential energy surface involved in an attachment and decay process occurring through an intermediate isomer/conformer state

The decay constant in the absence of the isomer is proportional to the ratio of level densities

$$k \propto \frac{\rho_p(E - D_0)}{\rho_0(E)}, \quad \text{(no isomer)}, \tag{5.24}$$

where E is the total excitation energy. For the decay from the isomeric state, the analogous expression is, with E_i the excitation energy in the isomeric state, equal to

$$k \propto \frac{\rho_p(E_i - D_i)}{\rho_i(E_i)} P_i, \quad \text{(decay through isomeric state)}, \tag{5.25}$$

where P_i is the probability of the particle being in the isomeric state, as opposed to the global minimum state;

$$P_i = \frac{\rho_i(E_i)}{\rho_0(E) + \rho_i(E_i)}. \tag{5.26}$$

The final state is the same in the presence and absence of an isomeric gateway state and the difference in the numerators of (5.24, 5.25) is in the argument of the level density. But these two energies are clearly the identical;

$$E_i - D_i = E - (D_0 - D_i) - D_i = E - D_0. \tag{5.27}$$

Equation (5.25) can therefore be written as

$$\begin{aligned} k \propto \frac{\rho_p(E - D_0)}{\rho_i(E_i)} P_i &= \frac{\rho_p(E - D_0)}{\rho_i(E_i)} \frac{\rho_i(E_i)}{\rho_0(E) + \rho_i(E_i)} \\ &= \frac{\rho_p(E - D_0)}{\rho_0(E) + \rho_i(E_i)} = \frac{\rho_p(E - D_0)}{\rho_0(E) + \rho_i(E - (D_0 - D_i))}. \end{aligned} \tag{5.28}$$

The only difference between the rate constants of the two situations is therefore that the term $\rho_i(E - (D_0 - D_i))$ is added to $\rho_0(E)$ in the isomeric case. This is usually a small correction (see Exercise 5.3).

The conclusion is that if a low energy collision of the fragments does not lead to a too rapid decay, the presence of a complicated decay trajectory involving an intermediate state will not cause any problems in the application of detailed balance. This will be assumed in the following.

To proceed with a quantitative estimate of the rate constants, we will make the approximation that the energy, ε, of the fragments are written as a sum of independent contributions from translational (ε_t), vibrational (ε_v), rotational (ε_r) and electronic degrees of freedom (ε_{el}),

$$\varepsilon = \varepsilon_t + \varepsilon_v + \varepsilon_r + \varepsilon_{el}. \tag{5.29}$$

This is definitely not true in any rigorous sense. There are couplings between rotations and vibrations, as is well known, and vibrational frequencies depend on the molecules' Born-Oppenheimer surfaces, i.e. the electronic quantum numbers, and so on. Nevertheless, the inclusion of these cross terms will usually only give rise to relatively minor corrections. For the vibrational-rotational (vib-rot) coupling, this is so because the coupling is weak. The vibrational frequencies, rotational properties and electronic degeneracies of electronically excited states may be significantly different from those of the ground state. For relatively small particles or insulator or semiconductor particles, the probability of electronic excitations are usually also small because electronic excitations in those species are high in energy, and for that reason will potentially large differences in state properties usually not appear with any appreciable weight. For large metal particles, however, electronic excitations need to be considered.

Two situations involving large fragments will be now be analysed. One where the 'large' fragment, although larger than an atom, is still composed of only a few atoms. Then its internal degrees of freedom will carry away only a small fraction of the total energy, typically $T/2$ or T per degree of freedom. This facilitates the calculation of the total level density of the product state which, as always, is a convolution of the level densities of all the degrees of freedom. Denote with subscript x the quantities pertaining to a generic d.o.f. Then the combined level density is

$$\rho_{t,x}(E') = \int_0^{E'} \rho_x(\varepsilon_x)\, \rho_t(E' - \varepsilon_x) \mathrm{d}\varepsilon_x. \tag{5.30}$$

The upper limit on the integral can be replaced with infinity without measurable loss of precision and we will do that in the following without necessarily mentioning it. Because in this case the vibrational degrees of freedom of the product particle are sufficiently numerous, we can use the same expansion as when we calculated the integral over ε_t for the atomic evaporation,

$$\rho_{t,x}(E') \approx \rho_t(E') \int_0^{\infty} \rho_x(\varepsilon_x) \mathrm{e}^{-\varepsilon_x / T(E')} \mathrm{d}\varepsilon_x = \rho_t(E')\, Z_x(T(E')), \tag{5.31}$$

identifying E' with the product energy $E - E_a$. Hence the convolution corresponds to a multiplication with the canonical partition function of the integrated degree of

freedom. When the rotations, vibrations and electronic degrees of freedom of the small fragment decouple, we integrate independently over these in the same way and get the rate constant

$$k(E,\varepsilon_t)\mathrm{d}\varepsilon_t = \frac{m}{\pi^2\hbar^3}\sigma(\varepsilon_t)\varepsilon_t Z_v Z_r Z_{el}\frac{\rho_p(E-E_a-\varepsilon_t)}{\rho_r(E)}\mathrm{d}\varepsilon_t. \tag{5.32}$$

All the Z's will be functions of the product temperature, $T(E-E_a)$, as will the kinetic energy release distributions. This expression only differs from the one for atomic evaporation by the additional factors given by the canonical partition functions, and as before the total rate constant is found by integration over the translational kinetic energy ε_t:

$$k(E) = \frac{m}{\pi^2\hbar^3}\langle\sigma\rangle T_d^2 Z_v Z_r Z_{el}\frac{\rho_p(E-E_a)}{\rho_r(E)}. \tag{5.33}$$

The degeneracy factor g which appeared earlier has disappeared because it is most reasonable to consider it part of Z_{el}. Note that in the convolution we have implicitly assumed that the capture cross section for the inverse process only depends on the kinetic energy of the decay channel. This may of course be wrong. If you know a better expression for the cross section, nothing stops you from deriving a better version of (5.33) for your system.

When the internal degrees of freedom of the small fragment couple too strongly to allow the factorization of the partition functions, the rate constant must be calculated as

$$k(E) = \frac{m}{\pi^2\hbar^3}\langle\sigma\rangle T_d^2 Z_{v,r,el}\frac{\rho_p(E-E_a)}{\rho_r(E)}, \tag{5.34}$$

where $Z_{v,r,el}$ is the combined canonical partition function for all degrees of freedom of the small, molecular fragment. Smallish large fragments will often be characterized in the literature and there will therefore often be data available that allow a calculation of the most important contributions to this sum.

The product temperature was used in (5.33) to calculate the canonical partition functions. For better precision one may prefer to use the temperature estimated at the excitation energy $E-E_a-\overline{E}_p-\overline{E}_v-\overline{E}_{el}$, where the last three terms are the mean canonical values. This can be iterated if one so desires.

The most important new features in (5.33) are the canonical partition functions. In order of increasing importance they are Z_{el}, Z_v, Z_r, which is not surprisingly also the order of decreasing quantum energies. Table 5.1 gives the values for the relevant quantum energies for some dimers. As mentioned above, the rotational quantum energies in particular will be low compared to the product temperatures in most situations where unimolecular reactions are involved, and Z_r will therefore often give a large enhancement of the rate constant frequency factor for molecular evaporation, on the order of T/B (the precise meaning of the rotational constant is discussed in Chap. 8).

Table 5.1 Constants of some diatomic molecules. Isotope numbers are given because values depend on the masses

Element	B (cm^{-1})	ω (cm^{-1})	E_{el} (cm^{-1})
$^{7}Li_2$	0.67	351	14,070
$^{23}Na_2$	0.155	159	14,680
$^{39}K_2$	0.057	92	11,680
$^{12}C_2$	1.82	1855	716
$^{63}Cu_2$	0.109	265	20,400
$^{107}Ag_2$	0.051	192	23,000
$^{197}Au_2$	0.028	191	19,670
$^{51}V_2$	0.21	537	1,874
$^{52}Cr_2$	0.23	452	14,100

One must consider if the inclusion of the new degrees of freedom into the expressions for the rate constant obeys the known conservation laws, and if not, how important the errors committed are. The conserved quantities are energy, momentum, and angular momentum. Energy conservation is included to a good approximation but still only approximately because of the expansion used to get the microcanonical temperature used for the integration over ε_r, ε_v, and ε_{el}. We walked into this approximation with open eyes and will accept it without any further ado. The linear momentum is also conserved, for the same reason it is in the atomic evaporation case. The last quantity that must be conserved is angular momentum. This conservation law has not been implemented rigorously. We have implicitly assumed that it is possible to integrate with an identical weight for all rotational states of the small fragment. This is a decent approximation if the large fragment carries an angular momentum which is large compared to that of the small fragment.

The second case of large fragments deals with fragments of comparable size. For those, the expansion in (5.31) must be examined. It seems obvious that the expansion must fail when the break-up becomes more symmetric. Then there is no way of deciding which one of the product particles should be treated as the smaller fragment and which one as the heat bath. The question is if one can use the simple expansion in (5.34) to assign a temperature calculated from the energy $E - E_a$ to the fragments.

To give a specific answer to this question we use the level densities, ρ, for the two fragments labelled 1 and 2 given by;

$$\rho_i(E) = a_i E^{s_i - 1}, \quad i = 1, 2. \tag{5.35}$$

The average energy deposited into the fragment 2, say, is then

$$\begin{aligned} \langle E_2 \rangle &= \frac{\int_0^E \rho_1(E-\varepsilon)\varepsilon\rho_2(\varepsilon)\mathrm{d}\varepsilon}{\int_0^E \rho_1(E-\varepsilon)\rho_2(\varepsilon)\mathrm{d}\varepsilon} = E\frac{\Gamma(s_2+1)\Gamma(s_1+s_2)}{\Gamma(s_2)\Gamma(s_1+s_2+1)} \\ &= E\frac{s_2}{s_1+s_2}, \end{aligned} \tag{5.36}$$

which is hardly a surprising result. The same result holds for $\langle E_1 \rangle$ with 1 and 2 exchanged.

The question is then a question of the magnitude of the relative change in T_1 with this value of E_2. With the above result one gets:

$$\frac{1}{T_1}\frac{\mathrm{d}T_1}{\mathrm{d}E_1}E_2 = \frac{s_2}{s_1}. \tag{5.37}$$

For a variation of the temperature below, say, 10%, one therefore requires a ratio of fragment sizes of 0.1 or less, independent of the absolute size.

For the decays that do not allow a treatment of the microcanonical-to-canonical type, another procedure must be devised. First determine the level density of both fragments, including the numerical constants. If one or both of these cannot be simply parametrized, use a numerical integration to find the level density of the combined system. If they can be parametrized as simple functions that can be convoluted analytically, just do that. This may actually be easier done than said. The vibrational level densities of two particles, for example, is simply the level density of the combined collection of oscillators. In the high temperature limit, where the each take the form $\rho_i(E) = (E + E_{0,i})^{(s_i-1)/(s_i-1)!}\prod_j \hbar\omega_{j,i}$, the convoluted high energy level density is simply

$$\rho(E) = \frac{(E + E_{0,1} + E_{0,2})^{s_1+s_2-1}}{(s_1 + s_2 + 1)!\prod_j \hbar\omega_{j,1}\prod_j \hbar\omega_{j,2}}. \tag{5.38}$$

The technique holds for all energies and all pairs of level densities that can be parametrized similarly, though, and is not restricted to these standard forms.

5.4 RRKM Theory

The theory presented so far is based on detailed balance and is used mainly by physicists. Chemists tend to favor an alternative formulation, from which the term transition state theory derives, and which is identified with the main architects by the abbreviation RRKM (Rice, Ramsperger, Kassel, Marcus). As the term 'transition state' suggests, the basic idea behind the theory is to consider a decaying particle at the point where the atom is just leaving on a trajectory, called the reaction coordinate. The motion along this coordinate is free at the top of a barrier. The transition state may be equal to the asymptotically separated fragments in the absence of an activation barrier for the reverse process, i.e. when the top of the barrier is not a hill but a plateau. All other degrees of freedom are described as harmonic oscillators. The transition state is assumed to be populated according to the statistical weight of the state, and the total statistical weight of the transition state is the convolution of the vibrational degrees of freedom for the transition state, ρ_t, with the translational level density of the reaction coordinate, relative to the total level density of the reactant, $\rho_r(E)$:

$$W_t \mathrm{d}x\mathrm{d}p = \frac{\rho_t(E - E_a - \varepsilon)}{\rho_r(E)}\frac{\mathrm{d}x\mathrm{d}p}{h}. \tag{5.39}$$

The phase space element $\mathrm{d}x\mathrm{d}p/h$ is the one-dimensional phase space integral in the semiclassical approximation, E_a is the height of the barrier and ε is the kinetic energy of the reaction coordinate, $\varepsilon = p^2/2m$, where m is the reduced mass of the two products. The length of the reaction coordinate is L and we can then integrate over $\mathrm{d}x$. The integration is trivial and we get

$$W_t \mathrm{d}p = \frac{\rho_t(E - E_a - \varepsilon)}{\rho_r(E)} \frac{L}{h} \mathrm{d}p. \tag{5.40}$$

This weight is specified with respect to ε. The total rate constant is this population of the transition state times the flow out, which is the speed divided by L,

$$k_{RRKM} = W_t \frac{v}{L}. \tag{5.41}$$

Writing $v\mathrm{d}p = \mathrm{d}\varepsilon$ gives

$$k_{RRKM}\mathrm{d}\varepsilon = \frac{\rho_t(E - E_a - \varepsilon)}{h\rho_r(E)}\mathrm{d}\varepsilon, \tag{5.42}$$

which is the desired expression.

The reader may wonder if we have not double-counted the number of states leading to dissociation in this expression. After all, the velocity in the reaction coordinate can point in either direction and only one leads to dissociation. This is not the case. The level density used was derived from $\mathrm{d}x\mathrm{d}p/h$, and the transformation of the derivative $\mathrm{d}p = \sqrt{m/2\varepsilon}\mathrm{d}\varepsilon$ implicitly assumed that only positive momenta were considered. Had both directions been included in the calculation, the number of states would have been twice that. The factor of one half needed to pick out the right direction would then cancel that extra factor of two.

It is worth recalling to what the level densities entering (5.42) refer. The denominator is clear. It is the total vibrational level density of the reactant. The numerator is the level density of all vibrational degrees of freedom in the transition state. The number of these is one less than the number for the reactant particle, i.e. $3N - 7$, although the final product will only have $3N - 9$ vibrational degrees of freedom (for evaporation of an atom). The reason is that the reaction is assumed to take place along a one-dimensional trajectory, and hence only one vibrational degree of freedom is converted into a translational d.o.f. The two remaining d.o.f.'s which will ultimately be converted into translational d.o.f.'s are considered vibrational at this point. And since there is no reason these vibrations should not be excited, they contribute to the level density ρ_t. The RRKM rate constant derived here thus corresponds to the emission of an atom through a saddle point in the potential. It is a saddle point in $1 + 2$ dimensions, it should be remembered, and not the usual $1 + 1$ dimensions of a real saddle.

The physical picture described by the saddle point motion has consequences for the total rate constant. It will also have consequences for the predicted kinetic energy

release distributions which will be treated in the Sect. 5.7. Integrating over all possible kinetic energy releases in (5.42) corresponds to summing over all partitionings of the energy between the translational kinetic energy in the reaction coordinate and the energy stored in the vibrations in the transition state particle, or

$$k_{RRKM} = \frac{\int \rho_t(E - E_a - \varepsilon)\mathrm{d}\varepsilon}{h\rho_r(E)}, \tag{5.43}$$

where the ε's are the possible reaction coordinate kinetic energies. Thus one can rewrite the rate constant equivalently with a summation over all vibrational states of the product:

$$k_{RRKM} = \frac{\sum_j \rho_t(E_j)}{h\rho_r(E)}, \tag{5.44}$$

where E_j is any energy the product can take, below the limit $E - E_a$ imposed by energy conservation.

There is a good deal of similarity between the detailed balance and the RRKM theory. The similarities reflect that these two theories both describe activated processes. The differences are reflected in the frequency factor multiplying the ratio of level densities, which is due to the different treatment of the transition states in the two theories. Detailed balance is best suited to describe the situation where the transition state is equal to the asymptotically separated state, whereas RRKM is designed to describe situations where particle emission occurs along a coordinate that crosses a saddle point.

The kinetic energy release distributions is a very efficient test of the nature of the transition state, and can provide a clear experimental signature for the different transition states. Decays describable by the RRKM theory produce gaps in the energy distributions, which is absent in decays through transition states that are not saddle points. The experimental literature seems to favour the latter situation, even for molecular dissociation for which RRKM was developed, although bona fide RRKM spectra also appear occasionally. For nanoparticles a natural guess would disfavour saddle point transition states.

5.5 Electron Emission

Thermal emission of electrons is so similar to atomic evaporation that one almost doesn't need a separate section to describe it. There are a few differences, however. One is the word. For electrons it is called thermionic emission. Another is that the degeneracy g for electron emission is always 2, because this is the electron spin degeneracy. More important numerically is the difference in mass between an electron and an atom. The phase space factor of the emitted particle is directly proportional to the mass, which causes a considerable numerical difference between the rate constants for these two types of emission processes.

The mass of the electron also causes us to consider the quantization of angular momentum. The angular momentum quantum number is on the order of

$$l_{max} \sim \frac{r\sqrt{m_e T}}{\hbar}, \tag{5.45}$$

where r is the radius of the particle, m_e the mass of the electron and T the product ion temperature. The effect of the small electron mass may therefore have consequences for absorption cross sections. As an example we calculate that the l_{max} is 1.5 for the values $r = 1$ nm, $T = 2000$ K. A classical description is therefore borderline and we must expect quantum corrections to some of these. We will not calculate these here, although they may be needed for the interpretation of precise experiments.

Yet another difference is the capture cross section. Electrons interact with a particle differently from what an atom does. When an atom hits a particle with thermal speeds, it will gain some energy from the short range attraction experienced when it is close to the particle. Some fraction of this interaction energy and the initial thermal kinetic energy of the channel will dissipate into the vibrations of the particle during the collision. This reduces the energy available for a bouncing re-exit and if the total energy of the atom ends up negative before the rebounce, which does not take a lot of dissipation, it will be bound. The sticking coefficient will then be effectively unity for such collisions.

Two things make the energy transfer relatively inefficient in electron-particle collisions compared to atom-particle collisions. One is that the masses of the colliding particles are so dissimilar compared with the two particles in atom-atom collisions. This, by plain Newtonian dynamics, reduces the energy transfer. The other reason is that the excitation spectrum of moving atoms (vibrations, phonons) have low quantum energies compared to electronic excitations. The energy transfer is then not prevented because of a mismatch between the energy and the energy of a reception state. For electron-particle collisions things are less favorable. Either an impinging electron excites a vibration, which is a suppressed process because of the mass mismatch, or the electron excites an electronic state, which is also disfavored because electronic excitation energies are high and far between.

What these hand waving arguments say is that generally the capture of electrons can not be expected to be as efficient as that of atoms. This is important in particular for attachment of electrons to neutral particles and hence for thermal emission of electrons from anions. For emission from neutral species, the problem is most likely much less severe, although few experimental data are available on absolute attachment cross sections. This is due to an effect which partly compensates the mass ratio effect mentioned. Electrons impinging onto an ion are caught in a strong potential and gain a kinetic energy roughly corresponding to the ionization energy before they collide with the particle. This means that the relative energy loss (the degree of inelasticity) in the collision does not have to be very large to reduce the energy to less than the energy needed to escape, and hence capture the electron.

If one assumes that capture is efficient and the electron is captured on contact, the cross section can be calculated classically to a good approximation for a spherically

symmetric, attractive potential as (see Sect. 5.6),

$$\sigma(\varepsilon) = \pi R^2 \left(\frac{|V(R)|}{\varepsilon} + 1 \right), \tag{5.46}$$

where R is the capture distance and $V(R)$ is the interaction energy at that distance. For metallic particles, defined as particles containing freely mobile valence electrons, one may want to add a contribution from the image charge. Effectively that adds more short range attraction to the potential. We will stick with (5.46). With that cross section detailed balance gives the thermionic emission rate constant from neutral species

$$k(E, \varepsilon)\mathrm{d}\varepsilon = 2\frac{m_e}{\pi^2\hbar^3}\pi R^2(|V(R)| + \varepsilon)\frac{\rho_p(E - \Phi - \varepsilon)}{\rho_r(E)}\mathrm{d}\varepsilon, \tag{5.47}$$

The activation energy for the process is the ionization energy (or ionization potential), Φ. The rate constant for thermal electron detachment from an anion can be calculated (classically) with the cross sections found in the last part of this chapter.

5.6 Kinetic Energy Release in Unimolecular Reactions

A very important measurable quantity in molecular beam experiments is the kinetic energy of the fragments produced in a unimolecular reaction. The kinetic energy is determined by a combination of the reverse process cross section and the microcanonical temperature of the particle. The basic formalism required to solve this problem has been established previously in the chapter. We saw little difference between the expressions for the electron- and atom emission rate constants and this is true also for kinetic energy release. The differences appear when the equations are applied due to differences in the capture cross section where they differ significantly.

In the following sections we will calculate the kinetic energy release distributions for a couple of cross sections of different types. The material covered here does not exhaust the subject, by far. One interesting class of processes not included here are emission involving particles with permanent dipole moments. The problem of emission of an electron in the image charge potential of a metal cluster is also not given explicitly but can be solved with the methods given here.

The kinetic energy distributions can be calculated with (5.12). The relevant part reads:

$$k(E, \varepsilon) \propto \sigma(\varepsilon)\varepsilon\rho_p(E - E_a - \varepsilon) \propto \sigma(\varepsilon)\varepsilon \exp(-\varepsilon/T), \tag{5.48}$$

where E is the excitation energy of the parent particle, ε is the kinetic energy of the decay channel, i.e. the sum of the translational kinetic energy of the two fragments, $\sigma(\varepsilon)$ is the cross section for capture of the fragment in the inverse process, ρ_p is the level density of the product particle, E_a is the activation energy of the process, and T is the microcanonical temperature of the product particle. Because the expansion

of ρ_p is made in ε and not in $E_a - \varepsilon$, the convergence is rapid and we do not need to involve the finite heat bath correction in this case.

In the simplest application of this equation, the cross section is independent of the collision energy ε, as is the case for a geometric capture cross section. Then the distributions are given by

$$k(E, \varepsilon)\mathrm{d}\varepsilon \propto \varepsilon \mathrm{e}^{-\varepsilon/T}\mathrm{d}\varepsilon, \tag{5.49}$$

It is instructive to compare this distribution with a gas in equilibrium at temperature T (a real, canonical temperature). The distribution of momenta or energies in such a gas is given by the Maxwell-Boltzmann distribution

$$P(T, \varepsilon) \propto \mathrm{e}^{-\varepsilon/T}\mathrm{d}^3 p \propto \sqrt{\varepsilon}\mathrm{e}^{-\varepsilon/T}\mathrm{d}\varepsilon. \tag{5.50}$$

Apart from the different origins of the temperature in the Boltzmann factors, the difference between this distribution and (5.49) is the pre-exponential, a power of 1/2 and 1, respectively, on the energy. These powers both refer to free particles, and if one considers a sufficiently large surface in the application of (5.49) one would naively expect that the difference between these two distributions should disappear. For the microcanonical distribution we have even gone to great pains to define a temperature so that the two distributions have all the chances in the world to look similar.

The reason they are different, also in the bulk limit, is that they are distributions of different things. The Maxwell-Boltzmann distribution in (5.50) is the distribution of kinetic energies you would measure if you picked an atom at random from the gas sufficiently many times to measure the distribution. The average energy of a Maxwell-Boltzmann distribution is the well-known $\frac{3}{2}T$. The kinetic energy release in (5.49), on the other hand, is the distribution of kinetic energies one measures if one puts up an imaginary surface somewhere around the emitting particle and measures the kinetic energies of the atoms passing through this surface. In fact, you don't even need an emitting particle to do this Gedanken experiment, just put up a screen in a gas and measure the kinetic energies of the atoms passing through it. This is not the same as picking an atom at random because the flux through the surface is biased in favour of the fast molecules. The bias is the speed (see (5.8)) and the speed is essentially the square root of the energy. This is what causes the difference between the two distributions. Indeed, the average energy of a particle from a Maxwell-Boltzmann distribution that hits a surface located at $z = 0$ is given by

$$\langle E_k \rangle = \frac{\int_0^\infty \mathrm{d}p_z \int_{-\infty}^\infty v_z \frac{1}{2}mv^2 \mathrm{e}^{-\beta(\frac{1}{2}mv^2)}\mathrm{d}p_x\mathrm{d}p_y}{\int_0^\infty \mathrm{d}p_z \int_{-\infty}^\infty v_z \mathrm{e}^{-\beta(\frac{1}{2}mv^2)}\mathrm{d}p_x\mathrm{d}p_y} = 2T. \tag{5.51}$$

(The number of particles hitting the surface during a time interval δt is proportional to $v_z \delta t$). This is also the mean value of the distribution in (5.49), as we have now understood.

After this exercise, let's turn to a more general situation where the interaction is not simply an absorbing sphere but also has a long range tail which will be assumed

spherically symmetric. Due to the surface tension, particles tend to be close to spherical because this shape minimizes the surface area, and therefore also tend to produce a spherical interaction potential, in particular at long distances. Moving on an inward trajectory in such an approximately spherical potential, the small fragment will be captured on arrival at a certain distance. The sticking at this critical distance does not have to be assumed unity, but for the purpose of calculating kinetic energy release distributions it will be taken as independent of the relative speed of the particles at that point. If not independent of the speed, the sticking coefficient $s = 1$ must be replaced by a function of the energy, $s(\varepsilon)$. We will continue with the $s = 1$ description, with the assumption implicit in that choice that it is also independent of the angular momentum of the large particle. That assumption is definitely not universally valid for molecular fragmentation, for which care should therefore be exercised when applying the results here.

We will begin with a potential relevant for the interaction of an ionic and a polarizable fragment or for two neutral polarizable particles. It has the form

$$V(r) = -\frac{1}{2}\alpha\frac{e^2}{(4\pi\varepsilon_0)^2 r^4} - \sum_{n=0}^{\infty}\frac{C_{6+2n}}{r^{6+2n}}, \tag{5.52}$$

where the first term is the potential due to the polarizability α of the neutral fragment in an ion-neutral system. For a pair of neutral particles, the first term is therefore absent.

Most of the literature data on polarizability are in Gaussian units. This is not going to change for a long time so we need a conversion factor. The Gaussian unit of polarizability is volume, specifically cm^3, but one also finds units of a_0^3 or Å^3. In any case, the conversion between the SI unit and the Gaussian unit is

$$\alpha = 4\pi\varepsilon_0\alpha'(\text{ Gaussian}). \tag{5.53}$$

This makes the polarizability part of the potential equal to

$$V(r) = -\frac{1}{2}\alpha'\frac{e^2}{4\pi\varepsilon_0 r^4}. \tag{5.54}$$

For convenience we can use the definition

$$\alpha'' \equiv \alpha'\frac{e^2}{4\pi\varepsilon_0} = \alpha' 14.4\,\text{eVÅ} \tag{5.55}$$

Some values for elements are given in Table 5.2.

Initially we will consider point-like particles captured in a potential of the form given in (5.52). The cross section can be calculated by application of energy and angular momentum conservation. The angular momentum in the relative motion of the small fragment and the product is conserved due to the symmetry of the potential.

Table 5.2 Measured atomic polarizabilities in atomic units (see Appendix B for conversion factors)

Group 1	Group 2	Group 8
		He 1.384
Li 164	Be 37.8	Ne 2.668
Na 163	Mg 71.3	Ar 11.09
K 291	Ca 159	Kr 16.74
Rb 319	Sr 202	Xe 27.34
Cs 402	Ba 272	

With the impact parameter b, this means that $L = v_0 bm = b(2m\varepsilon)^{1/2} = m\dot{\theta}r^2$ is constant, where v_0 the relative speed of the asymptotically separated particles, m is the reduced mass of the two fragments, and θ the angular coordinate in the plane where the motion takes place. The impact parameter is defined as the smallest distance between the two centres of mass one would have in a trajectory where the interactions were hypothetically switched off; mathematically

$$b \equiv \left| \overline{r} - \frac{\overline{v} \cdot \overline{r}}{|\overline{v}|} \right|, \tag{5.56}$$

for the asymptotically separated fragments. The two conservation laws give the relation

$$\varepsilon = V(r) + \frac{m\dot{r}^2}{2} + \frac{(\dot{\theta}r)^2 m}{2} = V(r) + \frac{m\dot{r}^2}{2} + \varepsilon\frac{b^2}{r^2}, \tag{5.57}$$

where the angular momentum conservation was used in the last equality together with $\varepsilon = mv_0^2/2$.

In the turning point of the radial motion, if it exists, $\dot{r} = 0$. For this point the equation reduces to

$$V(r) + \varepsilon\frac{b^2}{r^2} - \varepsilon = 0. \tag{5.58}$$

The solution to this equation can be illustrated graphically with the rewrite

$$\frac{r^2 V(r)}{\varepsilon} = r^2 - b^2. \tag{5.59}$$

The right hand side of this equation is, as a function of r^2, a straight line with unit slope and a negative offset, and the left hand side is a negative function with positive slope and a negative curvature. The lowest curve in Fig. 5.7, labeled $r^2 V(r)/\epsilon_2$, illustrates the behavior of both sides of the equation for the point-like particles we are considering here.

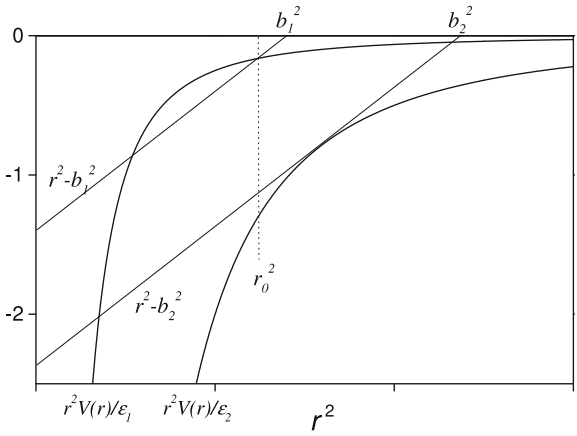


Fig. 5.7 Schematics of the left and right hand side of (5.59). The low energy curve with ε_2 corresponds to the cross section given by (5.61) with the value πb_2^2. The high energy curve, for ε_1, corresponds to the modified geometric cross section in (5.65) with the value πb_1^2. The capture radius r_0 is indicated by the dotted vertical line, and the values of b_1^2, b_2^2 given on the top axis

For a fixed energy (5.59) has solutions for values of b^2 larger than or equal to the values where the straight line is tangential to the $r^2V(r)$ curve. The two conditions that the curves are tangential are (5.59) and the equation

$$\frac{\mathrm{d}}{\mathrm{d}r^2}\left(\frac{r^2V(r)}{\varepsilon}\right) = \frac{\mathrm{d}}{\mathrm{d}r^2}\left(r^2 - b^2\right) = 1. \tag{5.60}$$

The negative curvature of $r^2V(r)$ ensures that there is precisely one real solution for b^2 to these two equations. A little algebra gives the value of this critical impact parameter b_c and hence the capture cross section;

$$\sigma_{cap} = \pi b_c^2 = \pi\left(r_c^2 + \frac{1}{\varepsilon}\left(\frac{\alpha''}{2r_c^2} + \sum_{n=0}^{\infty}\frac{C_{6+2n}}{r_c^{4+2n}}\right)\right), \tag{5.61}$$

where r_c is determined as the root of:

$$\frac{\alpha''}{2r_c^4} + \sum_{n=0}^{\infty}(n+2)\frac{C_{6+2n}}{r_c^{6+2n}} - \varepsilon = 0. \tag{5.62}$$

Physically, r_c is the smallest radial turning point that for a given energy does not lead to capture. If the impact parameter is reduced, the particle is guaranteed to be captured, and if it is increased, the particle is guaranteed to turn around in the radial motion and escape. In the special case when $C_{6+2n} = 0$ ($n \geqslant 2$) and only α'' is non-zero, (5.62) is easily solved to give the Langevin cross section,

$$\sigma_{cap} = \pi\left(\frac{2\alpha''}{\varepsilon}\right)^{1/2}. \tag{5.63}$$

So far, all results have referred to point particles. That will be sufficient for a range of parameters, as we will see below, but the finite size of the colliding particles will in general modify the cross section. So let's add a finite capture radius, r_0, to the problem. The different lengths, r_c, r_0 and b, are illustrated in Fig. 5.8. From the above equation for r_c and the properties of the potential it is easy to see that r_c increases monotonically when ε decreases. Furthermore, the derivative $db_c^2/d\varepsilon$ is negative, which can be verified by differentiation of the cross section and using the equation which determines r_c. The cross section is therefore given by (5.61, 5.62) when

$$\varepsilon \leqslant \tilde{\varepsilon} \equiv \frac{\alpha''}{2r_0^4} + \sum_{n=0}^{\infty}(n+2)\frac{C_{6+2n}}{r_0^{6+2n}}. \tag{5.64}$$

This energy is in general *not* the negative of the potential energy at r_0.

For energies exceeding $\tilde{\varepsilon}$ we have two turning points in the radial motion, of which the smallest, unphysical one, is inside the particle and corresponds to the turning point in a hypothetical radially confined motion. The cross section is determined by the largest turning point and can be found with reference to Fig. 5.7 from (5.59), evaluated at the capture radius, r_0. We have

$$\sigma_{cap} = \pi\left(r_0^2 + \frac{1}{\varepsilon}\left(\frac{\alpha''}{2r_0^2} + \sum_{n=0}^{\infty}\frac{C_{6+2n}}{r_0^{4+2n}}\right)\right), \qquad \varepsilon \geqslant \tilde{\varepsilon}. \tag{5.65}$$

The cross section is continuous across the critical energy $\tilde{\varepsilon}$.

The trajectories for different impact parameters and two different values of α'' and ε are shown in Fig. 5.8 for a numerical calculation of the $-1/r^4$ potential with two different energies. The capture distance r_0 has arbitrarily been set to unity. In the solution of the equations of motion the time is scaled as $\tau \equiv t\sqrt{2\varepsilon/m}$. The equations of motion that are solved are then

$$\frac{\mathrm{d}\theta}{\mathrm{d}\tau} = \frac{b}{r^2}, \quad \frac{\mathrm{d}r}{\mathrm{d}\tau} = \pm\left(1 - \frac{b^2}{r^2} + \frac{\alpha''}{2\varepsilon r^4}\right)^{1/2}, \tag{5.66}$$

where the sign depends on whether the particle is on the incoming or outgoing branch of the trajectory. The two situations shown correspond to a high ($\varepsilon > \tilde{\varepsilon}$) and a low energy ($\varepsilon < \tilde{\varepsilon}$) cross sections.

As another illustration we have calculated the kinetic energy releases for the case where α is the only non-vanishing coefficient in the potential. Equations (5.61, 5.65) give, after scaling with the temperature, the kinetic energy release distributions

$$P(\varepsilon) \propto \left(\frac{\varepsilon}{T} + \frac{\alpha''}{2r_0^4 T}\right)\mathrm{e}^{-\varepsilon/T}, \qquad \varepsilon \geqslant \tilde{\varepsilon} \equiv \frac{\alpha''}{2r_0^4} \tag{5.67}$$

$$P(\varepsilon) \propto 2\left(\frac{\alpha''}{2r_0^4 T}\right)^{1/2}\left(\frac{\varepsilon}{T}\right)^{1/2}\mathrm{e}^{-\varepsilon/T}, \quad \varepsilon \leqslant \tilde{\varepsilon},$$

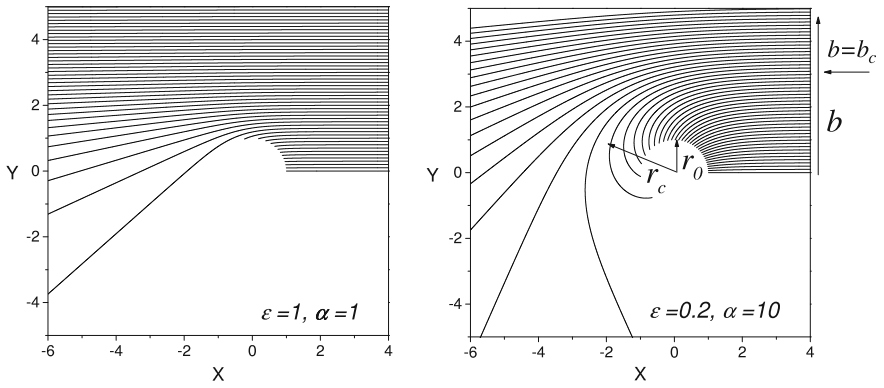


Fig. 5.8 Two sets of trajectories in a $1/r^4$ attractive potential. The left frame is a high energy and the right frame a low energy collision, as indicated by the value of ε. The right figure also shows some of the important lengths used in the text. $r_0 = 1$ by definition, r_c has a definite value because the energy is specified, equal to $\sqrt{5}$ in the units used (5.62). The impact parameter b takes all positive values. The critical value b_c is equal to $\sqrt{10}$ in the right frame (see (5.61))

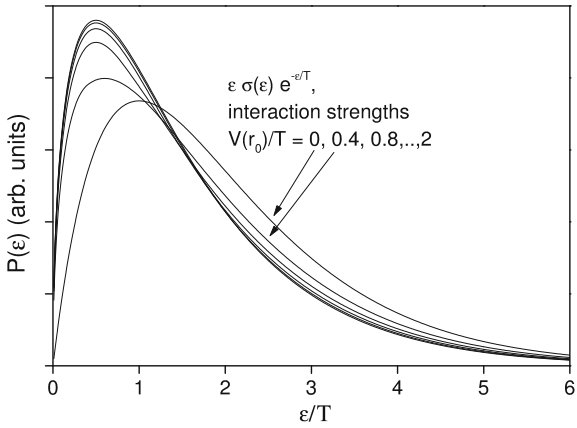


Fig. 5.9 Kinetic energy distributions calculated with (5.67) for six different polarizabilities. $V(r_0) = -\tilde{\varepsilon}$ is the interaction potential at the capture radius

where the constant of proportionality is the same. Curves for different values of $\alpha''/2r_0^2T$ are shown in Fig. 5.9.

It is relatively easy to derive the capture cross sections for the situation where the potential is determined by the C_6 coefficient. This is the leading approximation to most interactions between neutral particles. The equation for the critical distance r_c becomes

$$2\frac{C_6}{r_c^6} - \varepsilon = 0 \Rightarrow r_c = \left(\frac{2C_6}{\varepsilon}\right)^{1/6}, \tag{5.68}$$

and the cross section for low energy therefore

$$\sigma_{cap} = \frac{3\pi}{2}\left(\frac{2C_6}{\varepsilon}\right)^{1/3}, \qquad \varepsilon \leqslant \frac{2C_6}{r_0^6}. \tag{5.69}$$

At high energies the cross section becomes

$$\sigma_{cap} = \pi r_0^2 \left(1 + \frac{1}{\varepsilon}\frac{C_6}{r_0^6}\right), \qquad \varepsilon \geqslant \frac{2C_6}{r_0^6}. \tag{5.70}$$

Another important type of capture cross sections is the one generated by a Coulomb potential. It is relevant for electron emission from neutral and positively charged particles. Both particles, the product and the electron are charged. We will also here assume a finite capture radius, r_0, below which the electron is captured. Equation (5.58) for the turning point of the radial motion still holds for this potential because it was derived using only the spherical symmetry of the long range potential. Introducing the special form of the potential one has that at the radial turning point

$$\varepsilon = -\frac{Ze^2}{4\pi\epsilon_0 r} + \varepsilon\frac{b^2}{r^2}, \tag{5.71}$$

where Z is the charge after electron emission ($Z \geqslant 1$). The physical (positive) solution for $1/r$ is

$$\frac{1}{r_c} = \frac{Ze^2}{4\pi\epsilon_0 2b^2\varepsilon} + \sqrt{\frac{1}{b^2} + \left(\frac{Ze^2}{4\pi\epsilon_0 2b^2\varepsilon}\right)^2}, \tag{5.72}$$

showing that

$$\frac{\mathrm{d}r_c}{\mathrm{d}b} > 0. \tag{5.73}$$

Consequently, the largest impact parameter that will give capture for a given energy, b_c, is given by (5.71) with $r_c = r_0$. This gives

$$\sigma = \pi r_0^2 \left(1 + \frac{Ze^2}{4\pi\epsilon_0 r_0}\frac{1}{\varepsilon}\right). \tag{5.74}$$

The Coulomb capture cross section differs from the polarizability cross section in the analytical behavior; it is described by a single function for all energies. It is similar in the sense that this function is the same as the one for the high energy part of the polarizability potential, although the constants with dimension energy are obviously different in the two cases.

One may wonder how an extended attractive force can lead to a *larger* emission rate than, for example, no external force. One would intuitively expect the emission rate to be smaller, because the potential has some time to act on the small fragment when it flies away and to pull it back into the particle.

The simple explanation is to consider how the momentum and kinetic energy is divided into the radial and angular parts. For the sake of the argument, one can imagine that the potential is constant and less than the asymptotic value out to a radius which is larger than the capture radius, so the particle is described by two radii, one small within which the fragments equilibrate, and one large within which the fragments

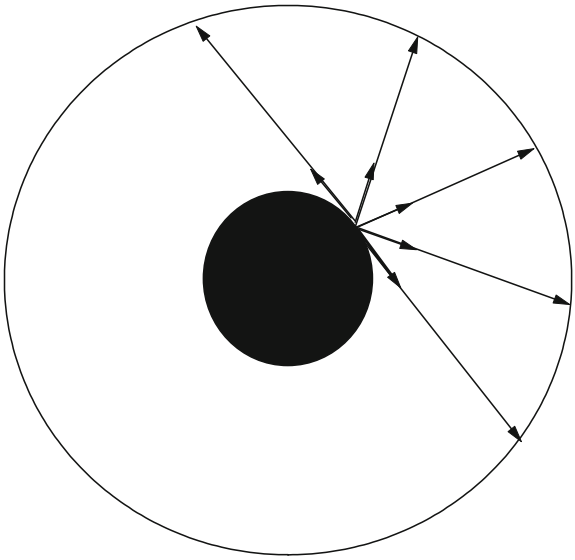


Fig. 5.10 Illustration of the effect of a long range potential on the escape probability for a fragment moving outward. The black part signifies the part of the particle where the velocity distribution is created. The larger open circle is the range of the potential. A few velocities are indicated with arrows. It is clear from inspection of the figure that the radial component increases when the outer radius increases

only move. Fragments emitted from the small core will have all possible ratios of radial to angular velocities, but only the radial momentum can be used to climb the activation energy barrier. As the fragments move radially outward, the distribution of radial to angular velocities will change, also in the absence of any thermalizing action. This is a consequence of simple geometry, as illustrated in Fig. 5.10. Some of the angular velocity will be converted into radial velocity, increasing the chance of having enough energy in the radial motion to overcome the activation energy barrier. Hence a long range potential will have a larger fraction of particles with sufficient radial energy to climb the potential and be emitted than a short range potential, everything else equal.

5.7 Kinetic Energy Release in RRKM Theory

The kinetic energy release distributions in RRKM theory are calculated on the basis of the idea of a translational degree of freedom equilibrated at and moving through a potential energy saddle point along a reaction coordinate, just like the total rate constant was derived, as consistency requires. This motion will give rise to a thermal distribution corresponding to a single translational degree of freedom, multiplied with the flux, which is proportional to the speed of the particle. The kinetic energy distribution along the reaction coordinate is therefore

$$P_{RRKM}(\varepsilon) \propto v\mathrm{e}^{-\varepsilon/T}\mathrm{d}v \propto \mathrm{e}^{-\varepsilon/T}\mathrm{d}\varepsilon, \tag{5.75}$$

where T is the microcanonical temperature of the transition state. This is consistent with the expression for the rate constant, found in Sect. 5.1:

$$k_{RRKM}\mathrm{d}\varepsilon = \frac{\rho_t(E - E_a - \varepsilon)}{2\pi\hbar\rho_r(E)}\mathrm{d}\varepsilon, \tag{5.76}$$

where ρ_t is the level density at the transition state, and $\rho_r(E)$ is the parent level density (E_a is the activation energy, and h is $2\pi\hbar$). With the usual expansion of the logarithm of the level density in ε, (5.76) is identical to (5.75).

Equation (5.75) predicts a Boltzmann distribution for the kinetic energy along the reaction coordinate. If there is no reverse activation barrier present, i.e. if the saddle point is flat in the outgoing direction, the resulting measured energy has the mean value T. At the transition state there will be two other directions that in the standard RRKM case are taken to give a parabolic potential. For a particle with a spherical long range potential and without a reverse activation barrier, this saddle point must be replaced by a plateau where the kinetic energy gets contributions from all three directions. Hence this gives the same result as detailed balance when the flux factor is included. When one has a transition state which is a real saddle point, where the asymptotic potential is below that of the transition state, equivalent to the presence of a reverse activation barrier, the difference in potential energy ΔE should simply be added to the energy in thermal distribution to get the observed distribution,

$$P(\varepsilon)\,\mathrm{d}\varepsilon \propto \mathrm{e}^{-(\varepsilon-\Delta E)/T}\mathrm{d}\varepsilon, \tag{5.77}$$

for $\varepsilon \geqslant \Delta E$, and zero otherwise. T is the microcanonical temperature at the transition state.

Exercises

5.1 Calculate the ratio of density of states of product and reactant for a process that occurs with a rate constant of 10^4 s^{-1}, an inverse process cross section of 10^{-18} m^2, a speed of the emitted product (electron, say) of 10^3 m, and a volume of 100 L. How long do you need to wait on the average before the electron is attached to the particle again?

5.2 The volume cancels in the derivation of (5.12), and similarly also in the expressions for thermal emission of electrons and photons. This is true if the volume is large enough for the quantized level structure of the free particle to be ignored. Estimate for atoms and electrons the size of the confining box where the level structure begins to influence the emission rate constants. Use your favorite atom and typical values, or an evaporated helium atom and the energy 0.4 K.

5.3 Consider the effect of an isomer on the decay constant, following (5.24, 5.25). Make the assumption initially that $\rho_i(E) = \rho_0(E)$, and that they can be calculated

as the vibrational values for harmonic oscillators in the high temperature limit in the leading order approximation. Show that if the vibrational frequencies of the two states are identical, the isomeric contribution to the level density will always be smaller than that of the minimum energy state.

Next assume that the vibrational frequencies of the isomer are lower than the minimum energy state. Determine how much softer they must be to contribute equally to the level density. The answer depends on the number of oscillators. Assume that the average energy is per oscillator is $D_i/10$. Hint: Expand the logarithm of the level densities to leading order.

5.4 Calculate the atomic emission rate constant for a particle composed of harmonic oscillators with a Debye vibrational spectrum. Use your favorite atom and calculate numbers for a wide range of energies. Guesstimate a reasonable activation energy or look up the bulk value.

5.5 Using a range of Arrhenius-expression frequency factors $\omega = 10^{14}N^{2/3}$ to $\omega = 10^{15}N^{2/3}$, calculate the range of values of $\ln(\omega t)$ for experimental times t ranging from 10 ns to 1 ms (this will be relevant in a following chapter).

5.6 In (5.33) the degrees of freedom of the small fragment beyond the translational channel were expressed in terms of their canonical partition functions. This can also be done for the translational degree of freedom. Do this.

5.7 Show that the units of the rate constants calculated in this chapter are correct.

5.8 Calculate the value of $\overline{v}$ in (5.22).

5.9 Find the typical values of the de Broglie wavelength of an evaporating atom for two of your favorite systems and compare them with the size of the system. Estimate the angular momentum, in units of $\hbar$, of the relative motion of the small fragment and the product particle. Conclude about the classical vs. quantum mechanical description of the process.

5.10 For evaporation of a large fragment, find the change in the number of vibrational and rotational degrees of freedom in the process for the possible reactant and product geometries not mentioned in the text.

5.11 Discuss whether a statistical decay can give rise to discrete kinetic energy release distributions.

5.12 Show that the turning point in the radial motion in a $-1/r^4$ potential for the critical impact parameter corresponds to the maximum of the effective radial potential

$$V_{eff}(r) = -\frac{\alpha''}{2r^2} + \frac{L^2}{2\mu r^2}, \tag{5.78}$$

with α the redefined polarizability, L the angular momentum and μ the reduced mass.

5.13 Consider a surface composed of atoms that interact with an atom above the surface with a long range potential proportional to r^{-n}. Show that the effective potential from the surface varies as r^{-n+2} if the surface layer is thin, and as r^{-n+3} if it is infinitely or just sufficiently thick. Discuss qualitatively how this is modified for a finite size and spherical particle at long distances.

5.14 Verify that the numbers given in the caption to Fig. 5.8 are correct.

Chapter 6
Radiation

The scientific history of thermal photon emission is undoubtedly glorious. As every science student knows, we owe Planck's constant and the birth of quantum mechanics to work on precisely this problem. And just as small particles have analogues of many other macroscopic phenomena, so does radiation from small particles. There is a rich literature on the emission of thermal infrared radiation from molecules, partly driven by the desire to understand astrophysical phenomena. There has been considerably less interest in the thermal part of the bluer side of the electromagnetic spectrum. Presumably this priority has been motivated by the expectation that high energy excitations will not happen in small systems by thermal means, and visible photons therefore not emitted thermally.

Well, we know better now. This chapter is dedicated to the description of the thermal radiation from particles, small and very small. It will build on the detailed balance presented in the previous chapter, with the proper modifications.

6.1 Photon Level Density

Photon emission incorporates several new features compared to the emission of massive particles such as atoms and electrons. The first is that the level density of the emitted particle is different. Photons are relativistic and full-blown quantum mechanical particles, which means that we cannot use the semiclassical and non-relativistic free particle level density in (5.10), and we must retrace the calculation of level densities.

The starting point of the calculation is identical to the one used for massive particles; We enclose a photon in a volume, which for simplicity will be considered cubic with side lengths L, and quantize the levels in this volume. The standing wave solutions in one dimension give, as for massive particles;

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_6

$$n\frac{\lambda}{2} = L, \quad n = 1, 2, \ldots, \infty, \tag{6.1}$$

where λ is the wavelength, but where the energy of the particle is now $h\nu = hc/\lambda$ instead of the semiclassical $(h/\lambda)^2/2m$.

To generalize to three dimensions, write the quantization condition in terms of the photon wave vector

$$k = \frac{2\pi}{\lambda} = n\frac{\pi}{L}, \tag{6.2}$$

which in three dimensions becomes

$$\overline{k} = \frac{\pi}{L}(n_x, n_y, n_z), \quad n_i \in Z_+. \tag{6.3}$$

The energy squared is

$$E^2 = (h\nu)^2 = (\hbar c)^2\overline{k}^2 = (\hbar c)^2\left(\frac{\pi}{L}\right)^2(n_x^2 + n_y^2 + n_z^2). \tag{6.4}$$

If we rewrite this as

$$\left(\frac{2\nu L}{c}\right)^2 = n_x^2 + n_y^2 + n_z^2, \tag{6.5}$$

we see that the n's that fulfil this condition are the (integer) coordinates of those points that reside on the positive octant part of the spherical shell with radius $2\nu L/c$. Figure 6.1 illustrates the situation with the 2d analogue. The states are numbered with consecutive positive integers, and each state therefore has a volume of one in these units. The total number of states up to the energy $h\nu$ is then $1/2^3$ of the volume enclosed by this shell. Inclusion of the factor of two due to the two possible polarization directions of the photon gives the number of states up to photon energy $h\nu$

$$N(h\nu) = \frac{2}{8}\frac{4\pi}{3}\left(\frac{2\nu L}{c}\right)^3. \tag{6.6}$$

The level density of a photon is then

$$\rho_{ph}(h\nu)\mathrm{d}(h\nu) = \frac{\mathrm{d}N(h\nu)}{\mathrm{d}(h\nu)}\mathrm{d}(h\nu) = \frac{8\pi\nu^2L^3}{hc^3}\mathrm{d}(h\nu), \tag{6.7}$$

which has units of reciprocal energy. The number of states per frequency interval is

$$\rho_{ph}(\nu)\mathrm{d}\nu = \frac{8\pi\nu^2L^3}{c^3}\mathrm{d}\nu. \tag{6.8}$$

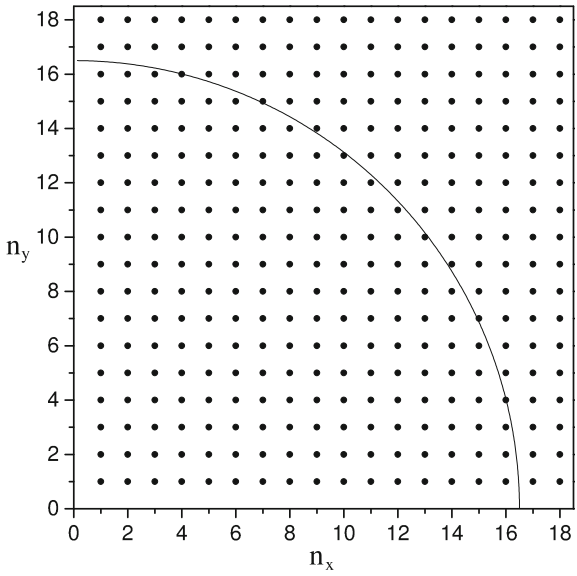


Fig. 6.1 The energy levels of a photon in a two-dimensional box of equal side lengths. A constant energy surface is a $\pi/2$ arc of a circle in the diagram, as shown. All states in the quarter disk have energies below that of the states residing on the surface

This level density is the number of ways a single photon can be distributed over a physical volume L^3. If one instead calculated the number of ways one could distribute a given energy over the same volume, irrespective over many photons this energy was distributed, it would be much larger. This can be seen if you split the energy into two photons of energy $h\nu_1$ and $h\nu_2$ with $h\nu_1 + h\nu_2 = h\nu$ (for details, see Exercise 6.1).

6.2 The Photon Emission Rate Constants

We now introduce the photon level density into the equation for rate constants. The formation rate constant is $\sigma(E - h\nu, \nu)c/V$, where E is the excitation energy of the particle. Volumes cancel, as before, and we get:

$$k(E, h\nu)\mathrm{d}\nu = \frac{8\pi\nu^2}{c^2}\sigma(E - h\nu, \nu)\frac{\rho(E - h\nu)}{\rho(E)}\mathrm{d}\nu. \tag{6.9}$$

This is the rate constant per frequency interval for emission of a photon with frequency ν, and it is therefore dimensionless.

The photon emission rate constant in (6.9) looks a lot like the emission rate constant for atoms and electrons, with a ratio of level densities and a prefactor which is reasonably energy independent.

One major difference between the emission of an atom/electron and a photon is the role of the activation energy. On emission of a photon, all excitation energy lost by the emitting particle is converted into the energy of the photon. For low energy photons this will effectively make photon emission a process without the very strong suppression of the emission caused by the presence of activation energies. Low energy photons means here that they have energies not significantly exceeding the excitation energy per degree of freedom;

$$h\nu \lesssim \frac{E}{s}. \tag{6.10}$$

In this situation, a substantial fraction of the particles can have emitting states populated and the emission rate is then mainly limited by the matrix element between the pre- and post-photon-emitting states or, equivalently, the inverse process cross section. This means that thermal photon emission may compete with other emission types under some conditions, even when the photon absorption cross section is small compared with the absorption cross section for massive particles.

Another difference to massive particle emission is the strong reduction of the factor $8\pi\nu^2/c^2\sigma(E-h\nu,\nu)$ compared with the corresponding factor for massive particle emission, even if the cross sections would be identical. Assume that this is the case, for the sake of argument. Inserting Planck's constant to make the photon emission factor one of rate per energy interval, the ratio of the two factors is then

$$\frac{\frac{8\pi\nu^2}{hc^2}}{\frac{m\varepsilon}{\pi^2\hbar^3}} = \frac{h^2\nu^2}{\varepsilon mc^2} \sim \frac{h\nu}{mc^2}, \tag{6.11}$$

where the last estimate follows from the fact that both the photon energy and the kinetic energy of an atomic evaporative or thermionic emission channel are on the order of the temperature. This ratio is an exceedingly small number which would suppress photon emission very strongly without the compensating reduction of the rate of the competing channel by the presence of its activation energy.

In spite of this suppressing factor and a potentially small photon absorption cross section, photon emission will dominate over other channels at sufficiently low excitation energies, precisely because it decreases much slower than these when the energy (temperature) is reduced. For the same reason, photon emission is also favoured by high dissociation energies, and it has been observed directly for clusters of different refractive materials (C, W, Nb), and indirectly from clusters of silicon, gold and boron.

The quantum statistics of photons introduces a third effect which makes thermal photon emission different from the other channels. The absorption of a photon competes with stimulated emission. Instead of absorbing a photon, a particle may emit an additional photon. The possibility this process can happen is a consequence of the bosonic nature of the photon (it has spin 1 $\hbar$).

As taught us by Einstein, stimulated emission happens with a rate which is proportional to the radiation density, which in our case is $h\nu/V$. This is the same volume, or

concentration, dependence as the absorption rate constant. If we want to relate a measured photo absorption cross section to the photon absorption cross section that appears in (6.9), we therefore need to consider both the stimulated emission and the absorption processes simultaneously.

When one measures the photon absorption cross section, the quantity recorded is the amount of transmitted light through a sample of the material. This is converted into a cross section, σ_{meas}, via Lambert-Beer's law,

$$I(L) = I_0 \mathrm{e}^{-\sigma_{meas} n L}, \tag{6.12}$$

where n is the density of particles in the sample of length L. This cross section includes both the absorption cross section $\sigma(E - h\nu, \nu)$, and the reduction of the absorption due to stimulated emission of photons from the particles. We can calculate this contribution if we take stimulated emission into account in the calculation of the formation cross section. So we need to redo the calculation of the cross section. For a single photon in the volume V, the relative rate of decrease of the photon number with time is

$$\frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\sigma(E - h\nu, \nu)c}{V} + B_{21}(E - h\nu, \nu)\frac{h\nu}{V}. \tag{6.13}$$

The first term has already appeared above. The second term is the contribution from stimulated emission which increases the number of free photons, as indicated by the positive value of the term. The fraction $h\nu/V$ is the energy density and we have redefined the meaning of the Einstein coefficient B_{21} to be the coefficient *per frequency interval*. With a rewrite of (6.13) we can get the effective photon absorption cross section;

$$\frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{c}{V}\left(\sigma(E - h\nu, \nu) - B_{21}(E - h\nu, \nu)\frac{h\nu}{c}\right). \tag{6.14}$$

The effective cross section is just the factor in the bracket;

$$\sigma_{meas}(E - h\nu, \nu) = \sigma(E - h\nu, \nu) - B_{21}(E - h\nu, \nu)\frac{h\nu}{c}. \tag{6.15}$$

This may not be awfully transparent (no pun intended). There is, however, a relation between this B-coefficient and the one for spontaneous emission, called the A-coefficient,

$$\frac{A(E, \nu)}{B_{21}(E, \nu)} = \frac{8\pi h\nu^3}{c^3}, \tag{6.16}$$

where the A has also been redefined to denote the rate constant per frequency interval. But A is nothing but our rate constant from (6.9). In this expression it should be evaluated at the energy $E - h\nu$, because that is the relevant energy for B_{21} in (6.15). If we then substitute (6.9) into (6.16) and that equation into (6.15) and cancel factors, we get

$$\sigma_{meas}(E - h\nu, \nu) = \sigma(E - h\nu, \nu) - \sigma(E - 2h\nu, \nu)\frac{\rho(E - 2h\nu)}{\rho(E - h\nu)}. \tag{6.17}$$

The second term in (6.17) is the contribution from stimulated emission. It is absent for atom and electron evaporation. It may also be absent for thermal radiation, which happens if the energy $E - h\nu$ is below $h\nu$. Then energy conservation prevents stimulated emission; Formally this is obtained if we set $\rho(E - 2h\nu) = 0$ for $E < 2h\nu$. But this is not the most interesting situation if we care about thermal photon emission, because in general energies will be so high that both terms will be present in (6.17).

If we don't know the dependence of $\sigma(E, \nu)$ on E, we are stuck. Absorption cross sections do in fact depend on the energy content of the particle, for example because thermal populations are temperature dependent, and cross sections depend on the initial state. We nevertheless want to have some idea of what the equations tell us, and therefore consider the emission from a particle with a heat capacity which is so high that the emission of a single photon does not significantly change the temperature and modify the spectrum. Then both terms on the right hand side of (6.17) include the same cross section, $\sigma(E, \nu) = \sigma(\nu)$. This gives the rate constant for emission of a photon for an absorption cross section which is independent of particle energy:

$$k(E, \nu) = \frac{8\pi\nu^2}{c^2}\sigma_{meas}(\nu)\frac{\frac{\rho(E-h\nu)}{\rho(E)}}{1 - \frac{\rho(E-2h\nu)}{\rho(E-h\nu)}}. \tag{6.18}$$

If the heat capacity is not too small, we can approximate the ratio of level densities as

$$\frac{\rho(E - 2h\nu)}{\rho(E - h\nu)} \approx \frac{\rho(E - h\nu)}{\rho(E)} \approx \mathrm{e}^{-h\nu/T}, \tag{6.19}$$

where T is the microcanonical temperature (see Chap. 3), and we get

$$k(T, \nu) = \frac{8\pi\nu^2}{c^2}\sigma(\nu)\frac{\mathrm{e}^{-h\nu/T}}{1 - \mathrm{e}^{-h\nu/T}}. \tag{6.20}$$

If the cross section is geometric and independent of frequency, (6.20) is the Planck radiation formula. To see this, first set the cross section to the constant value $\sigma = S/4$ (why 1/4?), where S is the surface area, and then integrate $h\nu k(T, \nu)$ over all frequencies to get the total emitted power:

$$P = \int_0^\infty h\nu\frac{8\pi\nu^2}{c^2}\frac{S}{4}\frac{\mathrm{e}^{-h\nu/T}}{1 - \mathrm{e}^{-h\nu/T}}\mathrm{d}\nu = \frac{2\pi Sh}{c^2}\int_0^\infty \nu^3\frac{\mathrm{e}^{-h\nu/T}}{1 - \mathrm{e}^{-h\nu/T}}\mathrm{d}\nu. \tag{6.21}$$

(black body, geometric cross section)

The integral is calculated to $(T/h)^4\, 3\,\zeta(4) = (T/h)^4\pi^4/15$. This gives an emitted power of

$$P = \frac{\pi^2 k_B^4}{60\hbar^3 c^2} S T^4 \tag{6.22}$$

where Boltzmann's constant has been reintroduced. The constant in front of the factor ST^4 is indeed Stefan-Boltzmann's constant. It has the value $\sigma_{SB} = (130\ \mathrm{K})^{-4}$ eV s^{-1} $\mathrm{Å}^{-2}$. Hence the bulk limit is reproduced correctly, as required.

To get some idea of the magnitude of total photon emission rate constants, we consider the maximum integrated photo-absorption cross section. This is given by a so-called sum rule. When the interaction between the particle and light is described in the dipole approximation, where the wavelength of the light is required to be much longer than the dimension of the particle, the photo absorption cross section obeys the Thomas-Reiche-Kuhn (TRK) sum rule:

$$\int_0^\infty \sigma(\omega)\mathrm{d}\hbar\omega = N_e\hbar\frac{e^2}{4\pi\varepsilon_0}\frac{2\pi^2}{m_e c} \approx N_e 1.09\, e\mathrm{V}\mathrm{Å}^2, \tag{6.23}$$

where N_e is the number of valence electrons. For the sake of argument we represent the cross section as a δ function,

$$\sigma = N_e 1.09\ \mathrm{eV}\mathrm{Å}^2\delta(h\nu - h\nu_0). \tag{6.24}$$

With a photon energy $h\nu_0 = 0.5$ eV, for example, $8\pi\nu_0^2/c^2 = 4.1 \cdot 10^{-8}$ $\mathrm{Å}^{-2}$, which gives the rate constant

$$\begin{aligned} k(E) = & \qquad (6.25)\\ & \int_0^\infty \frac{8\pi\nu^2}{c^2}\delta(h\nu - h\nu_0)N_e 1.09\ \mathrm{eV}\mathrm{Å}^2\frac{\rho(E-h\nu)}{\rho(E)}\frac{\rho(E-h\nu)}{\rho(E-h\nu)-\rho(E-2h\nu)}\mathrm{d}\nu \\ & = 1.1\cdot 10^7 \mathrm{s}^{-1}\, N_e \frac{\rho(E-h\nu_0)}{\rho(E)}\frac{\rho(E-h\nu_0)}{\rho(E-h\nu_0)-\rho(E-2h\nu_0)}. \end{aligned}$$

For this numerical example, the prefactor to the ratios of level densities is then 8 orders of magnitude or so smaller than for atomic evaporation, depending on the number of valence electrons.

6.3 IR Emission

The emission of radiation from optically active vibrations is so common and has features that allow us to specify the nature of the radiation in more detail than usual that it is worth an explicit treatment. The universally used term for this radiation is infrared (IR) radiation, as this is the wavelength region in which the radiation is emitted. Infrared stretches from 700 nm and upward to 1 mm or thereabouts (definitions differ slightly), corresponding to photon energies of 1.8 to $1.2 \cdot 10^{-3}$ eV. This covers all possible vibrations one encounters in practice, and includes the energy

scales of some molecular rotations at the low end and some electronic transitions at the high end. Rotational transitions occur simultaneously with vibrational, but we will ignore this complication in the following.

Vibrational transitions require a non-vanishing matrix element of the operator representing the photons between the initial and final state of the absorber. For vibrational transitions, the largest contribution is proportional to the matrix element of the first order term of an expansion of the dipole moment, μ, with respect to the vibrational coordinate,

$$m \propto \frac{\partial \mu}{\partial x} k \langle M|x|N\rangle, \tag{6.26}$$

where the states $|M\rangle$ and $|N\rangle$ represent the state of the oscillator with M respective N vibrational quanta. The zeroth order term in the expansion is the permanent dipole moment which may be non-zero, but it does not connect states with different values of M and N because these are orthogonal and the zeroth order term is simply a constant. That term will therefore not cause any IR transitions. It will cause rotational transitions, though.

Not all vibrations will have a non-vanishing derivative of the dipole moment and therefore associated radiative transitions. This will be the case for a dimer (a molecule composed of two identical atoms). The requirement on the dipole moment is intuitively clear; The only interaction between the particle and light is through the dipole moment. If the dipole moment does not change as a mode vibrates, the light is blind to this vibration and cannot excite it. A charge on the particle does not change this conclusion. The charge only causes the center of mass to move but will not induce vibrations.

What is less transparent is which modes have a non-vanishing derivative and are therefore optically active. This question can be answered by application of group theory. It would unfortunately take us too far astray if we tried to give this an adequate treatment, and we refer to textbooks in physical chemistry for this subject.

Assume now that we have identified a vibrational mode that can radiate. The quantum mechanical selection rules for the transition can be found easily when the position operator in (6.26) is written in terms of the raising and lowering operators a and $a^\dagger$ of the emitting mode as

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger), \tag{6.27}$$

where ω is the vibrational frequency and m is a mass, with a precise definition we won't specify except to note that it is related to atomic masses. The relation in (6.27) holds for harmonic oscillators. For anharmonic oscillators there will be more terms with higher powers of $a, a^\dagger$ that relax the selection rules and give rise to more possible transitions than found below. They are usually less intense and we will ignore them here.

The two lowering and raising operators $a, a^\dagger$ act by changing the number of vibrational quanta in a state as

$$a|N\rangle = \sqrt{N}|N-1\rangle, \tag{6.28}$$
$$a^{\dagger}|N\rangle = \sqrt{N+1}|N+1\rangle, \tag{6.29}$$

corresponding to stimulated emission (a) and absorption of radiation ($a^{\dagger}$). The stimulated emission will have the matrix element with modulus squared equal to

$$\left|m_{M,N,-}\right|^2 = Nh\nu\,|\langle M|N-1\rangle|^2 = Nh\nu\delta_{M,N-1}, \tag{6.30}$$

and the absorption the value

$$\left|m_{M,N,+}\right|^2 = (N+1)h\nu\,|\langle M|N+1\rangle|^2 = (N+1)h\nu\delta_{M,N+1}. \tag{6.31}$$

For both cases, the photon energy must conform to Bohr's frequency condition on the emitted and absorbed light, viz. that the photon energy is equal to the difference in energy of the involved states:

$$|E_N - E_M| = h\nu_0, \tag{6.32}$$

as required by energy conservation. The cross section for absorption in state N is then

$$\begin{aligned}\sigma_N &\propto \left|m_{M,N,+}\right|^2 \\ &= \nu\delta(\nu-\nu_0)(N+1).\end{aligned} \tag{6.33}$$

Introducing the constant of proportionality, σ_0 and integrating over photon energies cancels the δ function and gives

$$\sigma_N = \nu_0\sigma_0(N+1). \tag{6.34}$$

This cross section is not energy-independent, $\sigma(E-h\nu) \neq \sigma(E-2h\nu)$, as we had to assume to proceed from (6.17), and it does not make sense to use that equation as we don't need it here. Instead we can calculate the energy dependence directly as the value averaged over the population of excited states. After integration over the photon energy, cancelling the δ function, we have the cross section averaged over the populations $P_N(E)$ of the vibrational states of the IR active mode:

$$\begin{aligned}\sigma(E) \equiv \sum_{N=0}^{[E/h\nu]} \sigma_N P_N(E) &= \sum_{N=0}^{[E/h\nu]} \sigma_0(N+1)\frac{\rho'(E-Nh\nu)}{\rho(E)} \\ &= \sigma_0 \sum_{N=0}^{[E/h\nu]} (N+1)\frac{\rho'(E-Nh\nu)}{\rho(E)},\end{aligned} \tag{6.35}$$

where $\rho(E)$ is the total level density and $\rho'(E-Nh\nu)$ is the level density of the whole system less the contribution from the emitting mode.

Inserting this cross section, with the argument $E - h\nu_0$, into (6.9) we get

$$k(E) = \frac{8\pi\nu_0^3}{c^2}\sigma_0 \sum_{N=0}^{[E/h\nu]} (N+1)\frac{\rho'(E-(N+1)h\nu_0)}{\rho(E)} \tag{6.36}$$
$$= \frac{8\pi\nu_0^3}{c^2}\sigma_0 \sum_{N=1}^{[E/h\nu]} N\frac{\rho'(E-Nh\nu_0)}{\rho(E)}$$

Note that the rate constant here is the total rate constant, and not the differential value per frequency interval, as in (6.18). We can express the factors multiplying the ratios of level densities in terms of Einstein's A-coefficient, $A^{(0)}$, for the transition which gives

$$k(E) = A^{(0)} \sum_{N=1}^{[E/h\nu]} N\frac{\rho'(E-Nh\nu_0)}{\rho(E)}, \tag{6.37}$$

which is the IR emission rate constant for IR radiation of an isolated microcanonical particle.

It seems that stimulated emission has disappeared from (6.36). This is not the case. This is easiest seen in the canonical ensemble. Exercise 6.7 helps you to see this.

Figure 6.2 shows three examples of the temperature dependence of the IR rate constant from a single oscillator calculated with (6.36).

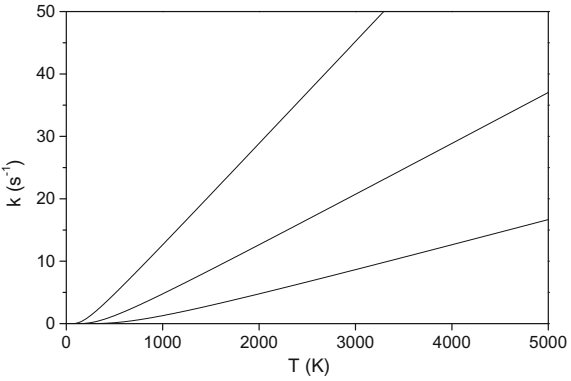


Fig. 6.2 The IR emission rate constants vs. temperature for a particle with a single IR active mode of quantum energies $h\nu = 500, 1000$, and 2000 K, and a cross section corresponding to 100 km/mol. The curves do not depend on the size of the particle because the mode is decoupled from all others in the particle, i.e. is truly harmonic. The curve continues to grow linearly at higher temperatures. The curves for the three different vibrational quantum energies scale, and really just show the same curve in different magnification

IR cross sections are usually much smaller than those associated with electronic transitions. This is a consequence of the much larger mass that enters in the denominator in (6.27), compared to the mass of the electron which is relevant for electronic transitions. The reason IR radiation is still observable and often dominant is a good example of the previously mentioned effect, viz. that the small matrix element is compensated by the relatively large Boltzmann-like factor $\rho(E - h\nu_0)/\rho(E)$.

Quantum mechanically calculated infrared cross sections are often given in the peculiar unit of km/mol. In order to translate a number from these units, multiply by the photon wavelength and divide by Avogadro's number, N_A. Hence a infrared intensity of 100 km/mol for a 0.1 eV photon is equivalent to a cross section of $2.06 \cdot 10^{-24}$ m^2.

6.4 Photon Emission from a Metal Particle

The most dominant low energy absorption feature in small metal particles is the surface plasmon resonance. This resonance can be understood in classical terms as the collective motion of the valence electrons in the approximately harmonic potential felt by these electrons when their center of mass is displaced a small distance from the equilibrium position where the negative charge of the electrons and the positive charges of the residual atoms almost cancel. The lifetime of the resonance is usually only a few times the period of oscillation. As any other resonance it therefore has a finite width and will extend both up and down in energy from the centroid energy, which is the name for the resonance peak position. A calculation of the classical photo-absorption cross section gives a low energy tail on the cross section that varies approximately as the frequency squared. The full classical form is ($\omega = 2\pi\nu$):

$$\sigma = \frac{e^2}{m_e c \varepsilon_0} \gamma N_e \frac{\omega^2}{(\omega^2 - \omega_s^2)^2 + (\omega\gamma)^2}, \tag{6.38}$$

where e is the electron charge, m_e the electron mass, N_e the number of valence electrons, ω_s the surface plasmon frequency, and γ the damping. The choice of constant in front of the pseudo-Lorentzian peak shape is made to have the resonance exhaust the TRK sum rule and (6.38) therefore represents the plasmon with the biggest radiative contribution possible. In real situations the contribution may be less, but it is not likely to be significantly less if the plasmon is present at all. It is a collective resonance, after all.

The two parameters that characterize the resonance are the width, γ, and the centroid of the resonance, ω_s. The width is a fraction of the resonance energy and may depend on temperature. The value of ω_s is proportional to the square root of the valence electron density in the free electron gas approximation. If the valence electrons are spread out over a sphere with radius r, the resonance frequency is

$$\omega_s^2 = \frac{e^2}{4\pi\varepsilon_0 m_e} \frac{N_e}{r^3}, \tag{6.39}$$

where m_e is the electron mass. Values are typically several eV, depending on the metal through the factor N_e/r^3 which essentially is the electron density in the particle. The bulk plasmon, for which the generic term 'plasmon frequency' is often used, is a factor $\sqrt{3}$ higher than the surface plasmon frequency.

Because this resonance is found to account for a large fraction of the low energy photon absorption cross section for a number of metals, it accounts for a correspondingly large part of the thermal photon emission. [1] When the resonance centroid energy is much larger than the temperature, $\hbar\omega_s \gg k_B T$, and the width is not too large, $\gamma < \omega_s$, the expression can be simplified in the low energy regime to

$$\sigma = \frac{e^2}{m_e c \varepsilon_0} \gamma N_e \frac{\omega^2}{\omega_s^4} = \pi r^2 \frac{4r\gamma}{c} \frac{\omega^2}{\omega_s^2} \tag{6.40}$$

In these cases, the emitted power varies with the 6th power of the temperature, vs. the 4th power for black body radiation.

6.5 Recurrent Fluorescence

Recently, radiation has been observed from thermally populated electronic states in molecules and small clusters. The process is known as recurrent fluorescence. The word 'recurrent' refers to the fact that the theory originally developed for the process assumed that the excitation energy was introduced into the particle by a photon absorbed into a specific electronically excited state.

Such a photo-excited state will often survive for some time, with the original geometry intact or more often having moved on the excited state Born-Oppenheimer surface into a geometry distorted relative to that of the ground state. Irrespective of the precise geometry, the system can then relax to the ground state or to some other low lying state by emission of a photon.

This relaxation is a common occurrence and is probably the one you hear about. But it competes with other channels known as radiationless transitions, in which the particle goes through a configuration where BO surfaces cross and the particle transfers to the electronic ground state, directly or in several steps. A molecule undergoing recurrent fluorescence follows this path but in addition also returns to the original electronically excited state. After reaching thermal equilibrium, a certain fraction of the particles will be in the absorbing state, and from these a photon can be emitted, in a process that is outwardly identical to a normal fluorescent photon emission.

Figure 6.3 illustrates the situation when the initial excitation is by photon absorption. In addition to the processes indicated in the figure, vibrational radiative transitions may occur, in particular from the electronic ground state, because it has the highest vibrational excitation energy. Since the emission is from a thermally equilibrated system, any way of exciting the system to the same energy will produce the

[1] Also for carbon, which is metallic in the sense of astronomers: Anything but hydrogen and helium.

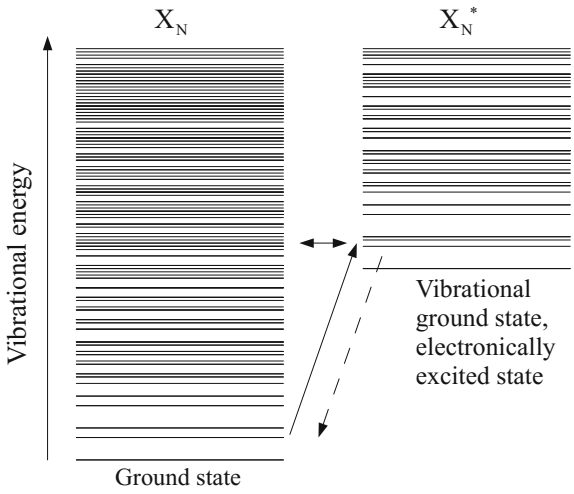


Fig. 6.3 Schematics of the recurrent fluorescence phenomenon. The electronic ground state particle absorbs a photon that promotes it to the electronically excited state indicated by an asterisk. Repeated inter-system crossings, indicated by the horizontal double arrow, populates the electronic ground state and the excited state according to their statistical weight. The excited state may emit a photon, shown by the dashed arrow, from this equilibrated system. A number of possible and competing decay channels have been left out of the figure for clarity

same light, and there is no fundamental difference between this process and photon emission connected with, say, infrared photons from vibrational transitions. It does exhibit the thermal nature of the process explicitly, however, and it is a good example of a thermal photon emission which proceeds via an activation energy, because the photons can have energies far above the temperature of the emitting particles.

The recurrent fluorescence mechanism is not expected to produce sharply defined photon energies. The energies of the emitted photons are determined by a few factors that will contribute to the smearing of the photon energy. One is the Franck-Condon factor, which spreads the emission over a number of vibrational states and therefore smears it energetically. This factor is also at work in the absorption, although the effect may have different magnitudes in the two situations, because the transitions are rarely taking place with the same nuclear coordinates on the two BO surfaces. The shift from absorbing, ground state geometry to an excited state emitting geometry also shifts the peak position.

Another contribution to the smearing is related to this shift in geometry. Relaxing to a lower point on the BO surface after excitation releases vibrational energy which will cause the surroundings of the excited state minimum geometry to be explored thermally. When the two BO surfaces are not aligned, also the energy difference, i.e. the photon energy, will therefore be sampled from a wider distribution.

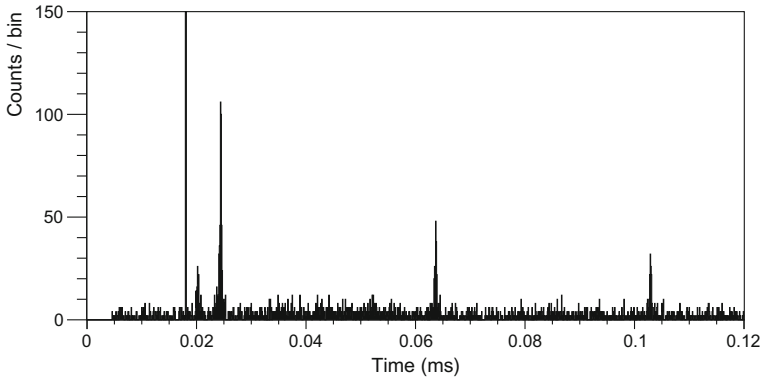


Fig. 6.4 The first detected recurrent fluorescent photons, emitted from C_6^- clusters in an electrostatic storage ring. The photons are emitted by a cluster bunch that pass a photo-multiplier detector at the times 24, 63, and 103 μs, where they give rise to the observed peaks. The photon energy is 2.04 eV

If one measures the photon emission rate constant, for example by the means outlined in Chap. 7, it is possible to get information about the energy of the emitted photon even without measuring it directly. This is possible because the TRK sum rule sets a rigorous upper limit on the possible absorption cross section for all absorbing species, not just for metallic particles with quasi-free electrons. The sum rule can be converted into a constraint on the cross section in (6.18), expressed in terms of the oscillator strength, f, which conforms to the inequality

$$f \leq N_e, \tag{6.41}$$

where N_e is the number of valence electrons. With the emitted frequency ν_0 the expression becomes

$$k_p = 7.43 \cdot 10^{-22}\text{Hz}^{-1} f \nu_0^2 \frac{\frac{\rho(E-h\nu_0)}{\rho(E)}}{1 - \frac{\rho(E-2h\nu_0)}{\rho(E-h\nu_0)}}. \tag{6.42}$$

Figure 6.5 shows the behaviour of (6.42) for cases with the simple level density $\rho \propto E^{s-1}$ for some values of s and energies per degree of freedom, E/s, around 0.2 eV. The difference between different sizes is minor, whereas the temperature is clearly more important. The shapes of the curves is defined essentially by the increasing factor ν_0^2, and the suppressing factor of the level densities, effectively the Boltzmann factor which prevents high energy excitations from radiating by reducing the thermal population of the emitting state. Photon emission rate constants can be determined experimentally with different methods. One is given in Chap. 7. An experimental value fixes the ordinate value in Fig. 6.5. It is clear from the figure that this determines both an upper and a lower limit for the energy of the emitting state, and just the upper limit expected from the Boltzmann suppression, the lower limit caused by the photon phase space factor ν^2.

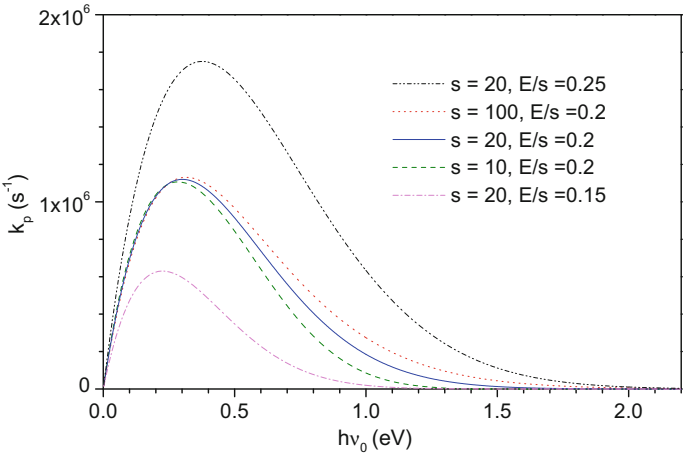


Fig. 6.5 The photon emission rate constants of (6.42) for $f = 1$ and the parameters indicated

The determination of the best limits by this procedure requires that the oscillator strength of the transition is known. The oscillator strength enters as a multiplicative factor on the curves and they will shift up or down with the value. Instead of best limits, one can derive the most conservative limits by assuming the maximum oscillator strength. This will still be useful, at least for small particles with a limited number of valence electrons. Another interesting number one can extract from (6.42) is the minimum oscillator strength required to produce an observed transition. By inspection of (6.42) we see that for sufficiently small oscillator strengths there will not be any photon energy with sufficient phase space and large enough Boltzmann factor to produce the observed rate constant. Both of these limits will provide guides for further experimental search for the emitting state.

Exercises

6.1 Use dimensional analysis to show that the combined level density of two photons with total energy $h\nu$ in a large volume V has the form

$$\frac{V^2(h\nu)^5}{h^6c^6}. \tag{6.43}$$

In addition to the dimensional analysis, you need to know that the result is proportional to the square of the volume, V^2. A large volume ensures that you can ignore the cases where two photons of the same energy are in the same state, which would give rise to double counting. Use the level density of (6.7) to calculate the value of the numerical constant missing in the dimensional analysis, and show that it is

$32\pi^2/15$. Compare the result with (6.7) with $V = 10^{-3}$ m^3 and $h\nu = 1$ eV. Calculate the hypothetical volume where the ratio of the two would be unity, and compare the result with the wavelength of a single photon with energy $h\nu$.

6.2 Use (6.11) to calculate an estimate of the ratio of the frequency factors for photon and atomic emission for your favourite atoms and a 0.1 eV photon.

6.3 Use (6.40) and the assumption leading to this equation to calculate an approximate expression for the emitted spectrum, the emitted power, the spectrally integrated emission rate constant and the mean photon energy. Compare results with Planck radiation for the same system.

6.4 Calculate the formula for the IR photon emission rate constant from the first vibrationally excited state by dimensional arguments. Use that the emission is proportional to the same matrix element squared that is responsible for absorption. It is proportional to the derivative of the dipole moment with respect to nuclear separation. Other relevant parameters for the problem are the coefficient in front of the raising and lowering operators, $\hbar$, the transition energy and the speed of light. Calculate a number with typical values of the mass of the oscillator and the transition energy.

6.5 Calculate the absorption cross section that gives rise to emission of IR radiation centered at $\hbar\omega = 0.1$ eV, with a small spectral width and with a rate constant of 100 s^{-1} from a particle with a microcanonical temperature temperature of 1000 K. Ignore the finite heat bath effect. Repeat the calculation for the same parameters but include the finite heat correction for a heat capacity of 30.

6.6 The photon emission rate constants can be approximated by an expression that is similar to the Planck radiation formula if the microcanonical temperature is used and the heat capacity is sufficiently large, as described in the text. Evaluate the correction due to the finite heat capacity to this Planck expression numerically for a real system of your own choice with a small to moderate heat capacity.

6.7 Use (6.37) and the canonical ensemble population density

$$P(E) = \frac{1}{Z}\rho(E)\mathrm{e}^{-\beta E} \tag{6.44}$$

of the of states with energy E to calculate the IR emission rate constant of a single harmonic vibrational mode in a particle. You can assume it does not couple to other modes and the canonical partition function therefore factorizes.

6.8 Show that the photon emission rate constant varies linearly with photon energy in (6.42) for small photon energies.

6.9 Is it possible, in principle at least, to use photon emission rates as a thermometer for the microcanonical temperature?

6.10 The emission of photons from electronically excited states seems to violate the principle of detailed balance, because the emitted photon energy is usually less than the absorbed. Explain why this breakdown is only apparent.

Chapter 7
The Evaporative Ensemble

Free and supported particles are generally different, and properties such as energy and structure are best studied in their pure form in free particles to avoid the effects of the interaction with a substrate, unless this interaction is the object of study. A large number of studies have therefore been performed on free clusters, either in molecular beams or in devices such as storage rings or ion traps, which are devices where ions can be stored for extended periods of time. If the particles are produced sufficiently cold and the experiments do not involve a significant energy transfer to the particle, no reactions that change the mass and/or the charge will occur during the experimental time. This means that one can identify the precursor of the products one is measuring in an experiment, which is not a trivial matter for neutral particles given the usually broad mass distributions produced in most particle sources. For these situations one can describe the systems with the microcanonical ensemble, apart from the short times during which the particles are exposed to some external manipulation (laser light, collisions), and possibly exchange of thermal radiation with the surroundings. Or more correctly, by a collection of microcanonical ensembles, very likely with different but individually conserved energies between the scientists' probing.

At the other extreme, particles that are prepared as hot as possible 'under the circumstances' will have certain properties that derive from this specific experimental condition. Some of these properties, energy distributions and decay rates in particular, will either be independent of the precise system studied or have scalable values. Knowing the systematics associated with this setting will allow us to extract properties of the systems. This is the subject of this chapter. Stated negatively, it will also help us to avoid some commonly seen errors. It turns out that a quantitative formulation of the vaguely sounding expression 'under the circumstances' is the key to understanding the properties of ensembles of hot systems.

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_7

7.1 Decay of Isolated Particles

The special features of ensembles of freely decaying particles appear as the consequences of two properties of evaporating particles. The first is that the energy is conserved between two decays. By decay we will here understand emission of an atom, an electron or a photon; a unimolecular reaction, in short. Absorption of a photon from the ambient radiation may also occur. If the absorption and emission of radiation is sufficiently rapid, the particle is in thermal equilibrium with a heat bath. We will restrict our cases to an absorption which is so weak it can be considered absent. Likewise, there is no continuous flux of atoms or electrons onto the particle. All changes in the composition of the particle are losses. When mass and energy is lost, we cannot describe the system as a microcanonical ensemble. The situation where a collection of particles lose energy and mass by evaporation in vacuum in the absence of equilibration with external energy sources is instead the situation described by the evaporative ensemble.

The second important feature is the enormous separation between the experimental time scales and any measured lifetimes on one hand, and on the other hand the intrinsic time scales for the unimolecular decay of the particles. The lifetime is typically given by flight times for ions over centimeters to meters or more, at thermal or keV energies, corresponding to times of several tens of nanoseconds and longer, or by the storage times in a storage device. Those times can exceed hours, and for Penning traps even months.

The decay time scales defined by the internal properties of the particle can potentially vary over a very large range. The Arrhenius form captures the essential features of an activated process rate constant;

$$k = \omega \mathrm{e}^{-E_a/T}, \tag{7.1}$$

where ω is a frequency which is energy independent or at least has only a weak dependence on energy compared to the energy dependence of the exponential function. The properties of ω, the value of T as a function of the excitation energy and the proper interpretation of E_a are all worked out in detail in Chap. 5. The results from that chapter are implicit here, although simplifications will be made. For the purpose of molecular beam experiments, the rate constant can take essentially all values between zero and 10^{12} s^{-1}, corresponding to decay times of a picosecond to essentially infinity, which in these situations is defined by the ion beam retention time.

When the observation time is much longer than $1/\omega$, the highest microcanonical temperature that can be sustained by the particle is significantly less than the evaporative activation energy, E_a, which in the following will be assumed identical to the dissociation energy with the modifications introduced in Chap. 5. The highest possible temperature the particles can have on this time scale is be estimated with the relation

$$k_{max} \sim 1/t \Rightarrow T_{max} \approx \frac{E_a}{\ln(\omega t)}. \tag{7.2}$$

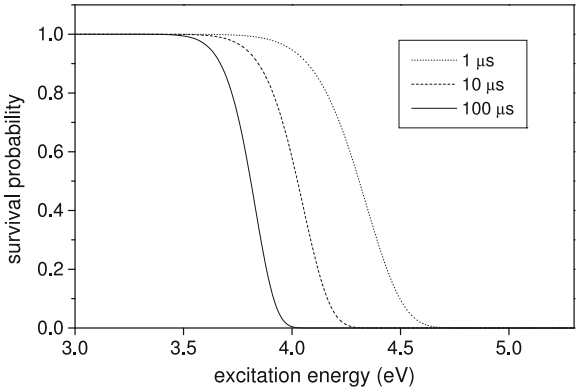


Fig. 7.1 The fraction of particles with decay rate constant given by (7.4) that have not decayed at the times indicated. The crossover from no decay to an entirely decayed population is not completely sharp, but the shape of the curve does not change much with time, the most significant change being a minor increase in the numerical value of the slope at the crossover energies with time

A more rigorous derivation of this relation will be made below. The situation is illustrated in Fig. 7.1, where the fraction, P, of surviving particles is shown as function of excitation energy, E, for three different times.

$$P(E) = \mathrm{e}^{-k(E)t}, \tag{7.3}$$

The example is calculated with the rate constant

$$k = 10^{16}\ \mathrm{s}^{-1}\mathrm{eV}^3 \frac{(E - 3\ \mathrm{eV} + 18 \cdot 0.01\ \mathrm{eV})^{17}}{(E + 21 \cdot 0.01\ \mathrm{eV})^{20}}. \tag{7.4}$$

This is the rate constant for a 9-mer, described as a collection of harmonic oscillators with an average frequency of 0.02 eV in the high energy limit. The non-exponential form of the rate constant is used to emphasize that the results that will be derived below do not depend on the exact functional form for the rate constant, nor do they require that a microcanonical temperature is defined.

The curves in Fig. 7.1 are the survival probabilities, but represent equally well the surviving population up to a constant factor, provided the initial population density is sufficiently flat. Precisely how flat is a question that will be investigated below. Unless otherwise mentioned, we will make the assumption that the survival probability also represents the population density of the surviving population.

The quantity $\ln(\omega t)$, which we have already encountered in Chap. 3, appears so often that it merits a name. It is called the Gspann parameter after J. Gspann who first suggested that there would be a highest temperature in an ensemble of freely evaporating particles. We define

$$G \equiv \ln(\omega t). \tag{7.5}$$

It is also denoted by γ in the literature. For $\omega = 10^{16}$ s^{-1} and $t = 10$ μs the value is $G = 25$. The Gspann parameter connects the highest microcanonical temperature in the ensemble to the activation energy. In the leading order finite heat bath approximation of the microcanonical temperature it is

$$T_{max} = \frac{E_a}{G} + \frac{E_a}{2C_N}, \tag{7.6}$$

where C_N is the heat capacity and the last term is the first order finite heat bath correction.

Let's now understand Fig. 7.1 in detail. The fact that the cutoff value of the energy depends on time, and also that it is not completely sharp is easily understood by considering the properties of (7.3). Consider first the diffuseness of the cutoff. The derivative with respect to energy (7.3) is

$$\frac{\mathrm{d}P}{\mathrm{d}E} = -\frac{E_a}{C_N T^2} kt \mathrm{e}^{-kt}. \tag{7.7}$$

We will ignore the T^{-2} dependence because it varies slowly with energy compared with the energy dependence of the rate constants. Then the numerical value of the derivative peaks at $kt = 1$, which is what we intuited in (7.2). Inserting this solution into (7.7) gives us a numerically largest slope of

$$\left.\frac{\mathrm{d}P}{\mathrm{d}E}\right|_{max} = -\frac{E_a}{C_N T^2}\mathrm{e}^{-1} = -\frac{G^2}{C_N E_a}\mathrm{e}^{-1}, \tag{7.8}$$

where the definition of G was used. Thus the slope changes only slowly with time, as we already saw in the numerical example given in Fig. 7.1, through the logarithmic factor $G^2 = \ln(\omega t)^2$.

The slope at the cut-off depends on particle size. Obviously the activation energies, E_a, will have size dependences in general, but the most important one is due to the size dependence of the heat capacity in the denominator. This dependence is absent if we consider the slope in the P vs. T curve instead,

$$\left.\frac{\mathrm{d}P}{\mathrm{d}T}\right|_{max} = -\frac{G^2}{E_a}\mathrm{e}^{-1}. \tag{7.9}$$

With $kt = 1$ at the crossover point, we can solve for the crossover temperature, T_{max}. The result is (7.2) without any qualifications on the equality sign. This result is exact for an Arrhenius rate constant, but is also a good approximation for other expressions that give better descriptions of the rate constants, and with the results in Sect. 5.2 we can translate the results of (7.3) to the properties of more realistic rate constants, if desired.

The approximate constancy of the slope of this survival curve allows us to find not only the time dependence of the crossover energy but also the decay rate, R. The decay rate is defined as the number of particles decaying per unit time, not to be confused with the rate constant. For a flat initial energy distribution with amplitude 1 it is

$$R \equiv -\int_0^\infty \frac{\mathrm{d}P}{\mathrm{d}t}\mathrm{d}E. \tag{7.10}$$

One way of calculating this is, instead of attempting a direct assault on the integral, to look at the time dependence of the crossover energy. The decay rate is proportional to the time derivative of this energy:

$$R = -g\frac{\mathrm{d}E_m}{\mathrm{d}t} = -gC_N\frac{\mathrm{d}T_m}{\mathrm{d}t} = g\frac{C_N E_a}{t(\ln(\omega t))^2} = g\frac{C_N E_a}{G^2 t}. \tag{7.11}$$

The main time dependence in this expression is the reciprocal of time. The constant of proportionality g, is the density of excitation energies, defined such that the number of particles that have an excitation energy between E and $E + \mathrm{d}E$ is equal to $g\mathrm{d}E$. Equation (7.11) tells us that an ensemble of mass-selected, isolated particles will decay essentially with a power law in time as long as g is constant.

The leading order power law decay is modified by the time dependence of the Gspann parameter. As a rule of thumb, for decays that occur on time scales of nano- to microseconds, a change in temperature of 10% corresponds to a factor of 20 in decay time, or conversely, that the factor G^2 changes 15% over a time interval that spans a factor of 10. The modification can be expressed as a correction to the power -1. In a power law decay we have a linear relation between logarithms of rates and times:

$$R \propto t^{-p} \rightarrow \ln(R) = -p\ln(t) + c \tag{7.12}$$

and p can be found as the derivative of the logarithm of R with respect to the logarithm of t:

$$p = -\frac{\mathrm{d}\ln(R)}{\mathrm{d}\ln(t)}. \tag{7.13}$$

Calculating the corrected power gives

$$\frac{\mathrm{d}\ln(R)}{\mathrm{d}\ln(t)} = \frac{\mathrm{d}}{\mathrm{d}\ln(t)}\left(-\ln(t) - 2\ln(\ln(\omega t))\right) = -1 - 2/G. \tag{7.14}$$

The reduction of the power with $-2/\ln(\omega t)$ is on the order of 5–10%, depending a little on the precise value of ω, i.e. the nature of the process, and the measurement time.

There are a few other small corrections to the 1/t decay in addition to the one generated by the Gspann parameter time dependence. One is the effect of finite heat capacities. We can find an expression for this and the previously calculated correction simultaneously for rate constants of the more general form

$$k \approx \omega \left(\frac{E + E_p - E_a}{E + E_r} \right)^{C_N - 1}, \tag{7.15}$$

where E_p and E_r are constants, and as usual E_a is the activation energy for the decay. Simplifications were made by setting ω constant and using the same heat capacity for both product and precursor. As before the decay rate is found by setting $k(E_{max})t = 1$, solving for E_{max} with (7.15) and taking the time derivative. We have, with the notation $\delta \equiv 1/(C_N - 1)$:

$$R \propto -\frac{\mathrm{d}E_{max}}{\mathrm{d}t} = \frac{E_a + E_r - E_p}{\left(1 - (\omega t)^{-\delta}\right)^2} \delta \omega^{-\delta} t^{-1-\delta}, \tag{7.16}$$

which gives

$$\begin{aligned} -p &= \frac{\mathrm{d}\ln(-\frac{\mathrm{d}E_{max}}{\mathrm{d}t})}{\mathrm{d}\ln(t)} = t \frac{\mathrm{d}\ln(-\frac{\mathrm{d}E_{max}}{\mathrm{d}t})}{\mathrm{d}t} \\ &= -1 - \delta - 2\delta \frac{(\omega t)^{-\delta}}{1 - (\omega t)^{-\delta}}. \end{aligned} \tag{7.17}$$

We introduce $G = \ln(\omega t)$ and write the power as

$$p = 1 + \delta + 2\delta \frac{\mathrm{e}^{-G\delta}}{1 - \mathrm{e}^{-G\delta}}. \tag{7.18}$$

The first correction is appreciable for very small particles, composed of on the order of several atoms. The second correction generalizes the $2/G$ correction calculated above. It is always less than the limit $2/G$ which is reached at large heat capacities, $C_N \gg G$. The behavior of (7.18) is shown in Fig. 7.2.

The power law decay is important enough to warrant an alternative derivation. The decay rate at time t is generally given by the integral of the surviving fraction times the decay rate constant and the density of particles with energy E, $g(E)$:

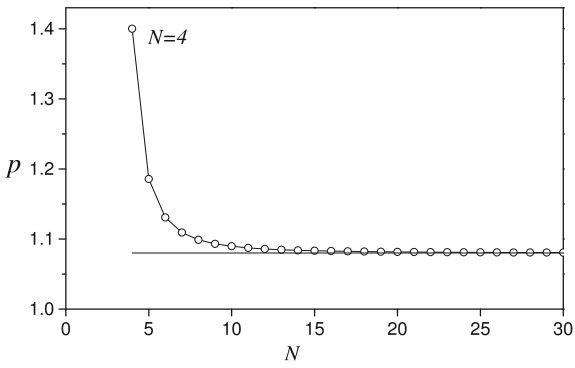


Fig. 7.2 The behavior of p calculated in (7.18) for $G = 25$ and a heat capacity of $C_N = 3N - 8.5$, corresponding to the average of the reactant and the product, both described as high temperature ideal harmonic oscillators

$$R = \int_0^\infty g(E)k(E)\mathrm{e}^{-k(E)t}\mathrm{d}E. \tag{7.19}$$

If, as before, we assume a constant density, $g(E) = g$, we can take that factor outside the integral. Next, substitute the rate constant for the energy in the integral:

$$R = g\int_0^\infty k\mathrm{e}^{-kt}\left(\frac{\mathrm{d}k}{\mathrm{d}E}\right)^{-1}\mathrm{d}k = g\int_0^\infty \mathrm{e}^{-kt}\left(\frac{\mathrm{d}\ln k}{\mathrm{d}E}\right)^{-1}\mathrm{d}k. \tag{7.20}$$

The integral's upper limit is extended to infinity whereas the formally correct value is $k(E = \infty)$. The substitution is justified because observation times are so long that the highest rate constant is much greater than the value $1/t$, which is where the integrand peaks. Furthermore, under these conditions we expect that the logarithmic derivative of k is a slowly varying function compared with the argument of the exponential, $-kt$, and can therefore be taken outside the integral. Evaluated at $k = 1/t$ it all gives:

$$R \approx g\left(\frac{\mathrm{d}\ln k}{\mathrm{d}E}\right)^{-1}\Big|_{k=1/t}\int_0^\infty \mathrm{e}^{-kt}\mathrm{d}k = g\left(\frac{\mathrm{d}\ln k}{\mathrm{d}E}\right)^{-1}\Big|_{k=1/t}\frac{1}{t}\int_0^\infty \mathrm{e}^{-kt}\mathrm{d}kt. \tag{7.21}$$

The integral is unity and we end up with

$$R \approx g\left(\frac{\mathrm{d}\ln k}{\mathrm{d}E}\right)^{-1}\Big|_{k=1/t}\frac{1}{t}. \tag{7.22}$$

To check if the approximation of the logarithmic derivative of k that appears in this equation with a constant is permissible, we need to specify an expression for k. We consider the by now familiar expression $k = \omega\rho_p(E - E_a)/\rho_r(E)$, where ρ is the level density of the product state and reactant (parent) particle, and calculate

$$\frac{\mathrm{d}\ln k}{\mathrm{d}E} = \frac{\mathrm{d}\ln\left[\rho_p(E - E_a)/\rho_r(E)\right]}{\mathrm{d}E} = \frac{1}{T_p} - \frac{1}{T_r} \approx \frac{E_a}{C_N T_p T_r}, \tag{7.23}$$

for identical heat capacities. The product $T_p T_r$ can be approximated with the square of the emission temperature, T, in the Arrhenius expression; $T_p T_r = T^2$. T has been defined previously (see Chap. 3) where the quality of the approximation was discussed. With $E_a/T = \ln(\omega/k)$ we have

$$\frac{\mathrm{d}\ln k}{\mathrm{d}E} \approx \frac{\ln(\omega/k)^2}{E_a C_N}. \tag{7.24}$$

The condition for approximating the right hand side with a constant is that $k \ll \omega$ or $1 \ll \omega t$. With respect to energies and temperatures this translates to $T \ll E_a$. Combining (7.22, 7.23) gives (7.11) again.

A third derivation of the power law decay also starts from the integral in (7.19) and uses a saddle point expansion to calculate it. This is left as an exercise.

At first glance it may seem that we have just derived the same time dependence in (7.11) that we put into the derivation from the outset, (7.2). Although the two equations look identical, they really describe different things: (7.2) gives the highest *rate constant* (k) one can have in the ensemble, whereas (7.11) gives an expression for the ensemble-averaged *decay rate* (R) one observes in an experiment. The $1/t$ decay rate is derived with only (7.2) and the fact that the rate constant is a rapidly varying function of energy. Exercise 7.1 illustrates the argument with an example of a hypothetical highest rate constant that does not vary with time as $1/t$ but still produces a $1/t$ decay rate.

It is worth emphasizing that the non-exponential decay appears even though the individual decays are completely statistical and each have the well-known exponential form. The power law is a result of adding many independent decays of this type with different decay constants. An integration over decay constants means that there are no characteristic time scales left in the problem, which is reflected in the fact that a power law in time does not contain any characteristic time. It is therefore not possible to extract any activation energy associated with a specific rate constant from a power law decay, at least under the condition that, as assumed so far, only one decay channel is activated. When two competing decays occur in parallel, the situation is much more favourable for determining system-specific parameters (see Sect. 7.8).

An experimental example of the power law decay is given in Fig. 7.3 which shows decay rates for a small anionic silver cluster measured in an ion storage ring. This storage device is capable of holding on to the ions for seconds by accelerating them to a fixed energy and then circulating them in a racetrack-shaped flight tube in vacuum, guiding their motion with static electric fields. The decay rate is measured as the rate of production of neutral particles in this experiment.

The results we have derived for the decay rates so far require that the energy distribution is relatively flat in the energy interval over which the decay takes place. As claimed, it is usually a good approximation to set g to a constant. To have a quantitative criterion for the applicability of this approximation, and also to be able

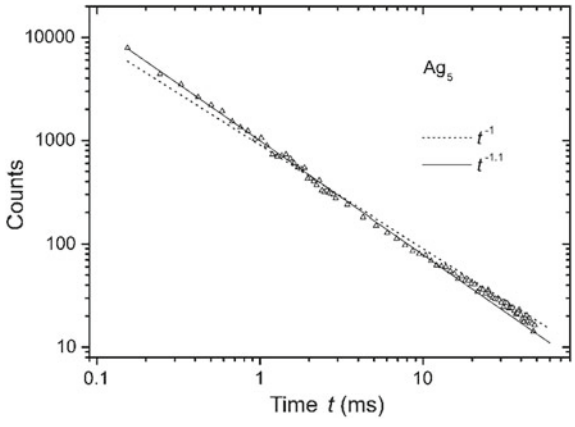


Fig. 7.3 The experimentally measured decay rate of Ag_5^-. The data are measured in the electrostatic storage ring ELISA. Reprinted from K. Hansen et al. *Phys. Rev. Lett.* **87** (2001) 123401, http://prl.aps.org/abstract/PRL/v87/i12/e123401. Copyright (2001) by the American Physical Society

to apply the corrections that may be present, we need to calculate decay rates including the variation of the distribution. A simple parametrization of the energy distribution is

$$g(E) = a\mathrm{e}^{bE}, \tag{7.25}$$

where b can be either positive or negative. The decay rate is then calculated as

$$R\mathrm{d}t \propto \int_0^{E_{max}(t)} g(E)\mathrm{d}E - \int_0^{E_{max}(t+\mathrm{d}t)} g(E)\mathrm{d}E = -g(E_{max})\frac{\mathrm{d}E_{max}}{\mathrm{d}t}\mathrm{d}t, \tag{7.26}$$

or

$$R \propto -g(E_{max})\frac{\mathrm{d}E_{max}}{\mathrm{d}t} = -a\mathrm{e}^{bE_{max}}\frac{\mathrm{d}E_{max}}{\mathrm{d}t}. \tag{7.27}$$

Taking the double-logarithmic derivative of this rate, the effect of the non-constant energy distribution is to change the power of the decay into

$$p \to p + b\frac{E_a C_N}{G^2}. \tag{7.28}$$

The nature of this correction differs from the other small corrections we have calculated above because it depends on the experimental conditions used to prepare the ensemble, and it will change if they are changed, e.g. if the power and/or the color of laser light used to excite the particles are changed.

We can find a good estimate for b from observed abundance spectra, provided the particle abundance distribution is measured as produced, without any later externally enforced cooling of their internal degrees of freedom. The width of the energy distribution for a specific size is then close to E_a for particles with heat capacities below a few hundred, as we will see below. An exponentially varying energy distribution as (7.25) will therefore translate into an exponentially varying mass distribution,

$$I_N = a' \exp(b'N), \tag{7.29}$$

where the two coefficients b and b' are related as

$$b' = -bE_a. \tag{7.30}$$

The minus sign appears because high energies appear as low masses. Hence, for a mass abundance spectrum that varies approximately exponentially, $I_N = a' \exp(b'N)$ over a not too narrow size range, we have the correction

$$p \to p - b'\frac{C_N}{G^2}, \tag{7.31}$$

instead of (7.28).

If one prepares the ensemble by laser excitation, there will potentially also be an effect on the decay from the non-smooth energy distribution caused by the discreteness of the energies absorbed. This discreteness will be partly washed out by the width of the initial internal energy distributions generated by the particle source and from whatever other smearing there may appear, such as the stochastic kinetic energy release that will appear during evaporation. We will calculate a conservative estimate of the effect of the discrete excitation energy caused by photon absorption by ignoring this smearing.

When the photon energies are sufficiently small, it is clear that the distributions of absorbed energy can be considered continuous. Quantitatively, if the rate constants of a particle having absorbed n photons is denoted k_n, the ratio of two particles having absorbed consecutive numbers of photons, $r \equiv k_{n+1}/k_n$, must be sufficiently close to unity for the discreteness of the photons to be unobservable. Hence the method of excitation by multiple photon absorption does not change the power law decay provided

$$\frac{\mathrm{d}\ln(k)}{\mathrm{d}E} h\nu \lesssim \chi \Rightarrow h\nu \lesssim \chi \frac{C_N E_a}{G^2}, \tag{7.32}$$

where χ is a pure number. As will be calculated, it turns out to be around $\ln(7)$, depending on the quality and statistics of the experiment. Figure 7.4 shows the decay rates vs. time in a number of ensembles where the decay constants differ by a factor r, i.e the figure shows the curves

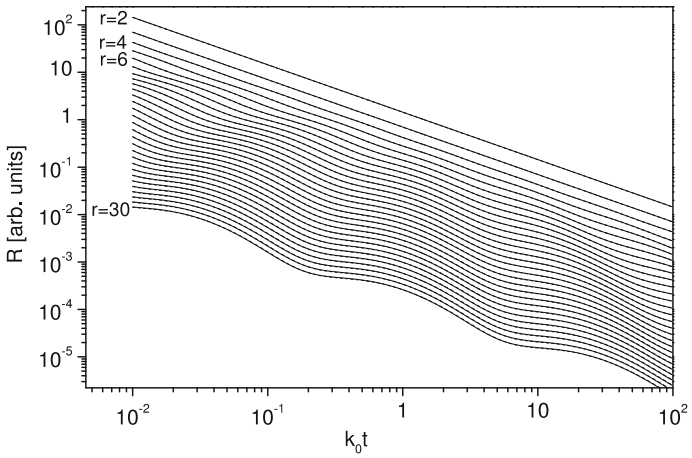


Fig. 7.4 The curves show the decay rates for series of geometrically spaced rate constants given by (7.33) for different values of r (full lines). No thermal smearing is included. The curves are offset for display purposes by a factor 1.3 between consecutive values of r. The curves calculated with the approximate expression derived in the text, (7.39), are also shown (dashed), but are hardly discernible on the scale of the figure. The good agreement between the two is striking considering that only the two lowest order terms are included in the Fourier series resulting in (7.39)

$$R = \sum_{n=-\infty}^{\infty} k_0 r^n \mathrm{e}^{-tk_0 r^n}, \tag{7.33}$$

for different values of r. On inspection of the figure one sees that the curves are well represented by a simple $1/t$ functional form up to at least $r = 4$.

We calculate the correction to the power law decay for the situation described in (7.33) by observing that the product of rate and time is periodic in the logarithm of the time with the period $\ln(r)$. This suggests the use of a Fourier series for the curve Rt in $\ln(k_0 t)$,

$$Rt = \sum_{j=-\infty}^{\infty} C_j \mathrm{e}^{\mathrm{i}2\pi \frac{j \ln(k_0 t)}{\ln(r)}} \tag{7.34}$$

The coefficients in this expansion are

$$C_j = \frac{1}{\ln(r)} \int_0^{\ln(r)} \sum_{n=-\infty}^{\infty} k_0 t \mathrm{e}^{n \ln(r)} \mathrm{e}^{-k_0 t \mathrm{e}^{n \ln(r)}} \mathrm{e}^{-\mathrm{i}2\pi \frac{j \ln(k_0 t)}{\ln(r)}} \mathrm{d} \ln(k_0 t). \tag{7.35}$$

We substitute $x \equiv \ln(k_0 t)$ and note that the sum of integrals adds up to the integral over all real values of $\ln(k_0 t)$. This gives us

$$C_j = \frac{1}{\ln(r)} \int_{-\infty}^{\infty} \mathrm{e}^x \mathrm{e}^{-\mathrm{e}^x} \mathrm{e}^{-\mathrm{i}2\pi \frac{jx}{\ln(r)}} \mathrm{d}x. \tag{7.36}$$

The coefficient of the constant term, $j = 0$, is calculated to

$$C_0 = \frac{1}{\ln(r)}. \tag{7.37}$$

We can calculate the higher order terms with a saddlepoint expansion. With the notation $\alpha \equiv j2\pi/\ln(r)$, the saddle is located at $x = \ln(-1 + \mathrm{i}\alpha)$. Thus, the value of that integral is approximated with

$$C_j \approx \frac{1}{\ln(r)} \int_{-\infty}^{\infty} (1 - \mathrm{i}\alpha) \mathrm{e}^{-1 + \mathrm{i}\alpha - \mathrm{i}\alpha \ln(1 - \mathrm{i}\alpha)} \mathrm{e}^{\frac{1}{2}(\mathrm{i}\alpha - 1)(x - \ln(1 - \mathrm{i}\alpha))^2} \mathrm{d}x. \tag{7.38}$$

After performing the Gaussian integral, the result can be simplified by noting that $|\alpha|$ is usually large compared to unity, and that therefore factors $\ln(1 - \mathrm{i}\alpha)$ can be expanded as $\ln(-\mathrm{i}\alpha) - 1/(\mathrm{i}\alpha)$. After a little algebra and adding the $j = 1, -1$ terms, this gives

$$Rt \approx \frac{1}{\ln(r)} + \frac{4\pi}{(\ln(r))^{3/2}} \mathrm{e}^{-\frac{\pi^2}{\ln(r)}} \cos\left(2\pi \frac{\ln(k_0 t)}{\ln(r)} + \phi\right). \tag{7.39}$$

The phase is:

$$\phi = 2\pi \left(-\frac{1}{8} + \frac{1}{\ln(r)} - \frac{1}{\ln(r)} \ln\left(\frac{2\pi}{\ln(r)} \right) \right). \tag{7.40}$$

From Fig. 7.4 the zeroth and the first order terms alone are seen to give a very good approximation to the amplitude, even for values of r that at first sight seem very large.

The ratio of the first and the zeroth Fourier coefficient in (7.39), which gives the magnitude of the deviation from a $1/t$ decay, oscillates between ± 0.05 when $r = 7.8$. If this is taken to be the highest acceptable oscillation, the photon energies that produce power law decays are limited by (7.32) to

$$h\nu \leqslant \ln(7.8) E_a \frac{C_N}{G^2} = 0.01 N E_a, \tag{7.41}$$

when $C_N = 3N$ and $G = 25$ is used.

One way of understanding the result qualitatively is that a power law decay appears when enough different rate constants contribute to the decay at any given time. This is effectively a smearing of the energy of the decaying particles. As mentioned, the above result is a conservative estimate. Any other smearing of the energy, and there are more reasons for this to occur than you may like to think, will increase the right hand side of (7.41). If the power law decay monitored is that of a particle which is the product of several evaporations, another and very efficient smearing effect comes into play and should be added to the picture (independent sources of smearing add, they don't replace each other). Because dissociation energies are rarely constant wrt. size and also rarely an exact multiple of a photon energy, the equidistantly spaced δ functions of the photon energy distributions will, upon evaporation of a number of monomers, be placed practically randomly on the energy axis between $Nh\nu$ and $(N+1)h\nu$, say. This means that the criterion on photon energies in (7.41) changes to

$$h\nu \leqslant 0.01 N \Delta n E_a, \tag{7.42}$$

when Δn denote the number of different initial, $t = 0$, particle sizes that contribute to a particle observed at time t. Although Δn is rarely known, it is easy to see that this more relaxed condition can be a lot easier to fulfil than (7.41), in particular considering both photon absorption statistics and laser beam inhomogeneities.

7.2 Abundances, Small Particles

One of the first quantitative measurements to perform when a cluster apparatus is commissioned is to record an abundance spectrum. Shapes of abundance spectra reflect a lot of factors, such as carrier gas pressures and temperature in the source, the geometry of the flow channel leading from the source to the vacuum, laser pulse

energies, degree of focusing, photon energies etc. These are the factors that will determine the mean sizes and widths of particle size distributions and the value of the parameter b of the energy in (7.25) in the nascent distribution.

In addition, spectra often reflect special stabilities of particles of certain sizes that have become known as magic numbers. The size-to-size variations in such abundance spectra at the magic numbers are usually generated in a way that makes an analysis in terms of the evaporative ensemble ideal.

The most important requirement for the analysis is that the particles have undergone at least one decay before measurement. This is not a necessary condition for every result in this chapter, but for abundances is easy to understand; if part of the beam consists of cold particles that have never evaporated, there is no way evaporative processes will give any information about this fraction, and there is no way of knowing how large this fraction is when detecting the particles. In most of the remainder of this chapter we will assume that this condition is fulfilled and remind the reader that this can and should be verified experimentally in each case by variations of the source and/or excitation conditions.

The existence of a highest temperature for a given particle size in the ensemble, discussed in detail above, automatically ensures that there is also a lowest temperature for the decay product of this size under these conditions. This lowest temperature, or energy, is determined by energy conservation:

$$E_{min,N-1} = E_{max,N} - E_{a,N} - \varepsilon_{kin,v,rot,el} \tag{7.43}$$

where for simplicity we have assumed that the decay is monomer loss, and we have added a size dependence to E_a. The term with $E_{a,N}$ accounts for energy conservation and it is therefore the true activation energy that enters this equation, not the one corrected with the energy offsets in (3.17) which is used in the rate constant. The last term on the right hand side is the sum of the kinetic energy release and the internal degrees of freedom of the small fragment. It is usually small enough to be left out in this context without serious problems. The relation says implicitly that no particles of size N with energy less than $E_{max,N}$ have decayed to produce particles of size $N-1$, just as no particle of N with energy above $E_{max,N}$ has survived.

Provided the lowest and the highest energies are sufficiently well separated, the energy distribution of size N is therefore confined to a distribution between the two approximate limits $E_{max,N+1} - E_{a,N+1}$ (lower limit) and $E_{max,N}$ (upper limit). For particles with small heat capacities, both of these cutoffs in the energy distribution are steep and the distribution will be approximately square. The low energy cutoff which includes a negative term, must both be non-negative and less than the high energy cutoff for the theory to be applicable. The non-negativity is required by the fact that the excitation energy is zero at the ground state (recall that energies here are excitation energies, i.e. calculated with the ground state as zero). Approximations that produce negative excitation energies must be refined or abandoned. In sufficiently realistic calculations, the energy will only very rarely be negative.

Particles usually contain a finite amount of thermal excitation energy, also after decay. This residual excitation energy is known as the kinetic shift and is the reason

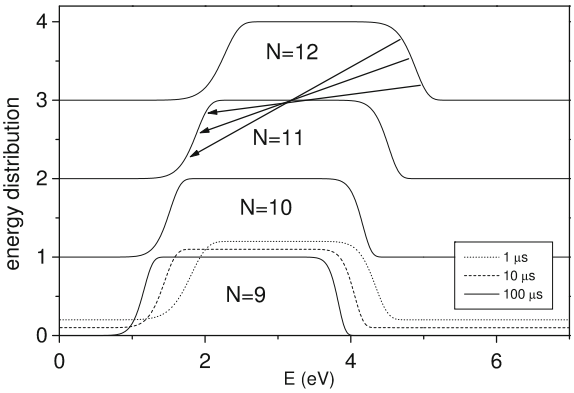


Fig. 7.5 The energy distributions at 100 μs of the 9-mer through the 12-mer of our generic cluster. The ordinate is shifted with unity for each mass for display purposes. The arrows from $N = 12$ to $N = 11$ point from different positions on the high energy edge of $N = 12$ to the corresponding low energy edge at $N = 11$. If drawn correctly the arrows have identical length on the energy axis. The three curves for $N = 9$ with ordinates shifted by 0.1 are the curves for 1, 10, and 100 μs

appearance energies, which are the lowest excitation energies of the reactant needed for a decay on a specific time scale, are not equal to the dissociation energies.

These considerations can be formulated quantitatively by multiplying the survival fraction in (7.3) with the complement of the survival fraction for the reactant;

$$P(E)\mathrm{d}E = \mathrm{e}^{-k_N(E)t}\left(1 - \mathrm{e}^{-k_{N+1}(E+E_{a,N+1})t}\right)\mathrm{d}E. \tag{7.44}$$

The first exponential gives the high energy cutoff we have seen before, and the last factor the low energy cutoff. As illustrated in Fig. 7.5 for different times, the equation can be approximated with a more or less square box between E_{min} and E_{max}, for judicious choices of these energies:

$$P(E) \approx \Theta(E_{max} - E)\Theta(E - E_{min}), \tag{7.45}$$

where $\Theta(x)$ is the step function which is zero for $x < 0$ and unity for $x > 0$. The generation of the lowest temperature for particle size $N - 1$ from the highest temperature of particle size N is also illustrated in Fig. 7.5. The rate constants used to produce the figures are the prototypical high temperature expressions also used with $N = 9$ in Fig. 7.3,

$$k_N(E) = \omega\frac{(E - E_a + (s-3)\hbar\overline{\omega}/2)^{s-4}}{(E + s\hbar\overline{\omega}/2)^{s-1}}, \tag{7.46}$$

with $s = 3N - 6$, $\hbar\overline{\omega} = 0.02$ eV, $E_a = 3$ eV and $\omega = 10^{16}\,\mathrm{s}^{-1}$. We note from Fig. 7.5 that the width of the energy distribution is close to E_a. This will be shown to be a

general feature of the evaporative ensemble applied to small particles as long as the E_a's are independent of size.

Abundances are proportional to the width of the internal energy distribution, $I_N \propto E_{N,max} - E_{N,min}$, provided the distribution of excitation energies is flat in the sense that it does not vary appreciably for a given particle size between the low and the high energy limits. An equivalent and more experimentally accessible measure is whether or not the excitation energy density varies from one particle size to the next. This is a well-defined question: 'How many particles are there per eV for size N and how many per eV for size M'. We have touched upon this question in connection with experimental-specific corrections to the power law decay above, where the correction was derived from a linear fit to the logarithms of abundances around a specific size. We need to expand on the analysis of that situation to cover the effect of variations in densities on abundances.

In the general case we can write the abundance as the integral of the energy distribution, $g(E)$, multiplied by the distribution in (7.44):

$$I_N = \int_0^\infty g(E)P(E)\,\mathrm{d}E \approx \int_{E_{min}}^{E_{max}} g(E)\,\mathrm{d}E. \tag{7.47}$$

The abundances involve the function $g(E)$ that includes instrumental functions, such as molecular beam transmission coefficients and detection efficiency, functions of the particle source parameters etc. This all sums up to a g which is very difficult to determine directly. However, we need to eliminate it from the experimentally measured abundances to have a chance to extract the two energies in (7.47), so g is an integral part of the problem we need to solve.

The relation between g and abundances can perhaps be understood easier if we use that abundances are the projection of the distribution of initial sizes and energies on final sizes, as illustrated in Fig. 7.6. This projection is almost unique; a definite initial N' and E' will be projected onto a definite E and N at some given measurement time. The only ambiguous cases are the energies located at the boundaries of the energy distributions of two neighboring cluster sizes, where the flip of a coin decides if it ends up as N or $N-1$. The stochastic kinetic energy releases will in principle also give rise to a stochastic final energy, but this just amounts to a minor smearing of the initial values of N' and E' that contribute to the intensity at E and N.

With the one-to-one correspondence between initial and final pairs of (E, N), we can understand g as the density of points in the (E, N) plane that are projected onto a given final energy. In Fig. 7.6 this is proportional to the number of dots enclosed between the two sloping lines. The mean value of g for a given size is denoted by $\tilde{g}_N$ and is formally defined as

$$\tilde{g}_N \equiv \frac{\int_{E_{min}}^{E_{max}} g(E)\mathrm{d}E}{E_{max} - E_{min}}. \tag{7.48}$$

The amount of particles that ends up as size N is proportional to this density times the width of the distribution, i.e. the value of $E_{max} - E_{min}$ for the specific size, or

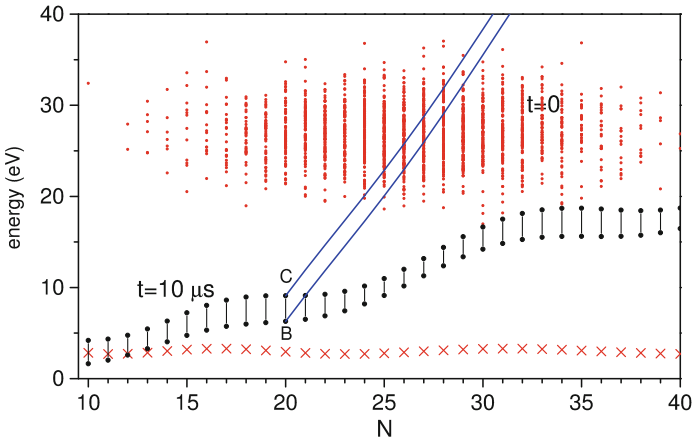


Fig. 7.6 An illustration of the projection of initial distributions on the final energies and sizes. The parameters of the decay constant are our generic ones and the measurement time is 10 μs. A periodic variation (originating from an equally generic shell structure) is added on top of the 3 eV average value to give the dissociation energy shown by crosses. The shell structure is observable as the wiggles on the energy distributions, shown for each size as a line terminated with filled circles. The two lines terminating at B and C define the region that is projected on size $N = 20$ from the cloud of dots which represent the initial particle sizes and energies. The finite excitation energies at the time of measurement in this Gedanken experiment is the kinetic shift

$$I_N = \tilde{g}_N(E_{max} - E_{min}). \tag{7.49}$$

This is not a formalistic definition. If the initial conditions (the density in the cloud in Fig. 7.6) are smooth, which is usually the case, $\tilde{g}_N$ is also a smooth function of N, and can be divided out.

The widths of the energy distributions are about an activation energy, as mentioned earlier and shown later. If we express the E_a's as a smooth function of N, $\tilde{E}_{a,N}$, plus a more rapidly varying function of size, we can find $\tilde{g}_N$ as

$$\tilde{I}_N = \tilde{g}_N \tilde{E}_{a,N}, \tag{7.50}$$

where $\tilde{I}_N$ is a smoothened abundance.
This connects the observable $\tilde{I}_N$ to the unknown $\tilde{g}_N$ and we have from the two preceding equations

$$\frac{I_N}{\tilde{I}_N} = \frac{1}{\tilde{E}_{a,N}}(E_{max,N} - E_{min,N}). \tag{7.51}$$

The factor $I_N/\tilde{I}_N$, which we may call the cluster's stability factor, is by construction independent of the details of the production process and the subsequent decay before detection if the assumption of at least one prior evaporation is fulfilled. This is an experimentally testable statement. The requirement for passing the test is that the stability factor does not change when source conditions etc. are changed to give

different mass distributions and hence different $\tilde{I}_N$'s. If the stability factors for all relevant particle sizes are indeed invariant under these changes, they can be used to determine the values of $E_{max} - E_{min}$ from the data.

We must have a prescription to determine $\tilde{I}_N$ from the data. There is some freedom in the choice of this function and some constraints. On one hand it should not vary too rapidly with N, because then the interesting abundance variations will disappear in the ratio $I_N/\tilde{I}_N$. On the other hand it should still vary sufficiently rapidly to remove the slow variations of I_N associated with the envelope of the abundance distribution. One possibility is to use an average produced with a Gaussian weight function with a running width:

$$\tilde{I}_N \equiv \frac{\sum_{N'} I_{N'} \exp\left(-\frac{(N'-N)^2}{\alpha \cdot N^2}\right)}{\sum_{N'} \exp\left(-\frac{(N'-N)^2}{\alpha \cdot N^2}\right)}, \tag{7.52}$$

where the sum runs over all masses in the spectrum and α is a suitable number. There are alternatives to this, for example spline functions. These functions are constructed so that they go to zero exactly at a certain point, as do their derivatives to whatever order one prefers to prescribe. The optimal width of the smoothing region, the value of α in (7.52), can be determined by application of the averaging function to several experimental spectra and varying α, starting with a high value, until, ideally, all stability factors in all spectra agree for any given particle size.

Figure 7.7 shows an application of the procedures described above to mass spectra of protonated water clusters, $(H_2O)_N H^+$. The four different abundance spectra that correspond to different source conditions are given as counts per cluster size.

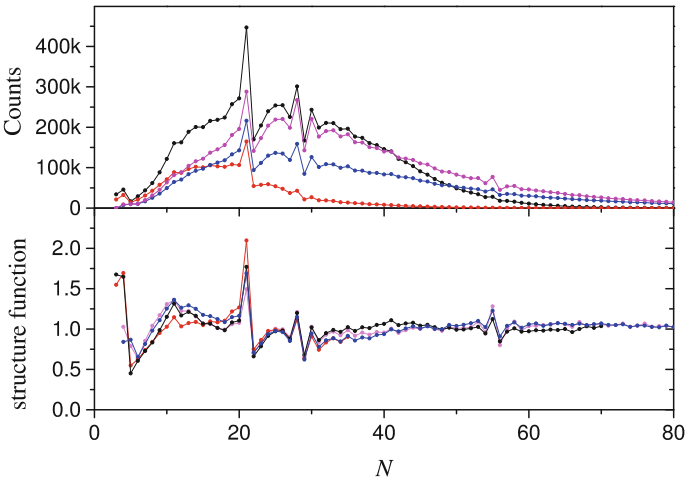


Fig. 7.7 Abundances of protonated water clusters produced in an electrospray source (top frame), and their stability function calculated with the methods given in the main text. The data are published in K. Hansen et al., *J. Chem. Phys.* **131** (2009) 124303

After $\tilde{I}_N$ is extracted from the experimental data, we have accounted for the instrumental effects on the abundance distributions. The stage is almost set to solve the job of relating the extracted structure function to the binding energies. To do so we need to include one more ingredient, though. Decays constants are determined by the temperatures, but the abundances are related to energies. We therefore need a relation between the energies and the temperature. This is called the caloric curve. We will expand the caloric curves of the particles in temperature to linear order,

$$E_{max,N} = C_N T_{max,N} - E_{0,N}. \tag{7.53}$$

As we have seen before, $E_{0,N}$ is in general not zero because the heat capacity for a certain particle is temperature dependent, and energy is therefore not simply proportional to temperature. Together with the Arrhenius expression for the rate constant and the leading order finite heat bath correction we have

$$E_{max,N} = E'_{a,N}\left(\frac{C_N}{G} + \frac{1}{2}\right) - E_{0,N}, \tag{7.54}$$

and

$$E_{min,N} = E'_{a,N+1}\left(\frac{C_{N+1}}{G} + \frac{1}{2}\right) - E_{0,N+1} - E_{a,N+1}. \tag{7.55}$$

The primed activation energies are the effective values derived in Chap. 3. With the notation here they are

$$E'_{a,N} \equiv E_{a,N} + E_{0,N} - E_{0,N-1}. \tag{7.56}$$

They are the ones that determine the maximum temperature. For the last term in (7.55) the unmodified value is used because it is determined by energy conservation.

The differences between these two energies involve the heat capacities and potentially also the energy offsets. We will treat the situations where the heat capacities are similar, although not necessarily identical. This should cover a fair number of situations (not a lot of heat capacities have been measured experimentally, so this is still a conjecture). Typically, the heat capacities of the two particles will differ by a relative amount of order $1/N$. We therefore define the difference and average of the heat capacities

$$\begin{aligned} \Delta C_N &\equiv C_{N+1} - C_N, \\ \overline{C}_N &\equiv \frac{1}{2}(C_{N+1} + C_N). \end{aligned} \tag{7.57}$$

Using (7.56) on the combination of the unprimed $E_{a,N+1}$ and the energy offsets $E_{0,N}$, $E_{0,N+1}$ gives the simple form for the width of the energy interval:

$$E_{max,N} - E_{min,N} = \frac{1}{2}(E'_{a,N} + E'_{a,N+1})\left(1 - \frac{\Delta C_N}{2G}\right) + \frac{\overline{C}_N}{G}(E'_{a,N} - E'_{a,N+1}). \tag{7.58}$$

For identical dissociation energies, the expression contains a negative correction. In the example shown in example in Fig. 7.5, a reduction in the width of the distributions relative to E_a can in fact be seen; The width of the distributions are all slightly less than the $E_a = 3$ eV used as input.

We want to use (7.58) with (7.51) iteratively to solve for the $E_{a,N}$'s from one end of the experimental spectrum to the other. That procedure is equivalent to solving a difference equation with the source term $I_N/\tilde{I}_N$. In this process, the small errors will accumulate and cause the procedure to develop exponentially increasing or decreasing solutions. The calculated activation energies will then be exponentially increasing or decreasing with N.

There are different ways to avoid this unphysical situation. One is simply to normalize the corrections away with an overall multiplicative factor. An alternative, which is the one that will be used here, is to approximate the equation by leaving out the offending term in (7.58), $\Delta C_N/2G$, which is on the order of $3/2G$, or 0.05–0.10. We will therefore use the approximate relation

$$E_{max,N} - E_{min,N} \approx \frac{1}{2}(E'_{a,N} + E'_{a,N+1}) + \frac{\overline{C}_N}{G}(E'_{a,N} - E'_{a,N+1}). \tag{7.59}$$

Together with (7.51) this gives the relative abundances as

$$\frac{I_N}{\tilde{I}_N} = \frac{E'_{a,N} + E'_{a,N+1}}{2\tilde{E}_{a,N}} + \frac{\overline{C}_N}{G}\frac{E'_{a,N} - E'_{a,N+1}}{\tilde{E}_{a,N}}. \tag{7.60}$$

As required this reproduces $I_N/\tilde{I}_N = 1$ when $E'_{a,N+1} = E'_{a,N}$. Solving for $E'_{a,N}$ gives

$$\frac{E'_{a,N}}{\tilde{E}_{a,N}} = \frac{1}{\frac{\overline{C}_N}{G} + \frac{1}{2}}\left[\frac{I_N}{\tilde{I}_N} + \frac{E'_{a,N+1}}{\tilde{E}_{a,N}}\left(\frac{\overline{C}_N}{G} - \frac{1}{2}\right)\right]. \tag{7.61}$$

This is the form of the equation that is solved iteratively. The ratio $E'_{a,N}/\tilde{E}_{a,n}$ for the highest experimental N is set to unity, and the iteration converges very rapidly to values that do not depend on this choice.

As the numerical recipe goes, this is the end of the road. When it comes to the interpretation of the results, it should be recalled that the activation energies found from this expression are the ones defined by (7.56), and the values of the two energy offsets must be substituted into the numbers to find the physical values.

Usually this is not a major correction that can be estimated from ground state properties, but there are situations where the two offsets can be very different. One

case is associated with a melting/freezing type phase transition, or in general any transition involving a latent heat. Such a situation will obviously also have implications for the approximations of the heat capacity differences used. These situations are best treated on a case by case basis.

Equation (7.60) invites several other remarks. It is obviously only valid for non-negative values of the right hand side of the equation. This sets an upper limit to the contrasts in dissociation energies the theory can handle. Exercise 7.3 asks you to calculate this limit. The limit is rarely exceed in practise. The experimental signature of a problem of this kind is a very small ratio of abundances, I_N/I_{N+1}. Large ratios do not pose any problem. This is fortunate because they often appear in mass spectra of clusters with shell structure. In fact, the analysis here has been developed with such systems in mind.

Secondly, the width of the energy distribution is close to the dissociation energy, as mentioned earlier. But, as we have also observed previously, see (7.8), the slope of both the high energy end and the low energy end of the distribution will decrease when the particle size is increased, because they are both inversely proportional to the heat capacity. At some sufficiently large heat capacity, the two sides will therefore begin to overlap. At that point we need to reconsider the analysis critically. From (7.8) we see that slopes merge at heat capacities around $C_N = G^2/2e$. Beyond this size, the question requires a separate treatment. That will be the subject of a good deal of the remainder of this chapter.

Thirdly, the binding energies derived from abundances can only be interpreted relative to neighboring particle sizes. This means that only relative activation energies can be found.

Finally, we note that a knowledge of the heat capacity is required in order to perform the inversion from abundances. Heat capacities can be determined experimentally from metastable decay fractions, and if such data are not available, one must make a educated guess.

One can draw some important conclusions from (7.60) even before using it on real data. First of all, one sees that abundance anomalies are related less to binding energies than to *changes* in binding energies. The second term in (7.60) is the product of changes in the E_a's and the numerical factor C_N/G. For particle sizes above ca. 10, this factor will for most materials be higher than unity and will cause the difference term to dominate the abundance variations variations. Another interesting observation is that for identical relative binding energy variations at different sizes, the abundance anomalies will be largest for the particles with the largest heat capacity.

It is also interesting to note what does *not* appear in the relation (7.60). First of all, one does not see any trace of a rate constant or a Boltzmann factor. They have been effectively removed by integrating over the energy. Similarly, temperatures appear only in a very hidden form.

There is very little time dependence in these abundances. Somewhat counter-intuitively, abundance variations that are due to variations in dissociation energies are most pronounced at short times if the E_a's are temperature independent. The end of a sequence of decays is reached when G is very large which, given the definition of the quantity, is at an exponentially long, practically infinite time.

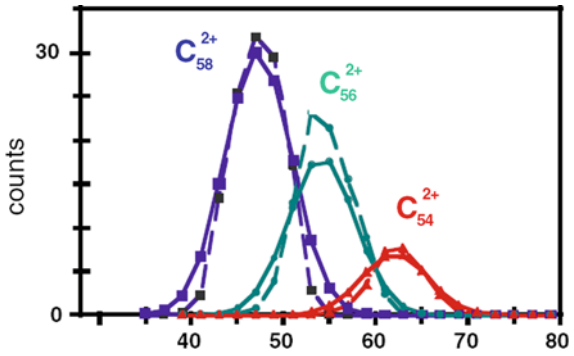


Fig. 7.8 The measured internal energy distributions of fragmenting fullerenes in eV. The all-carbon molecules called fullerenes fragment by C_2 loss, hence the appearance of only the even numbered fragments of C_{60}^+. The full line is the measured values and the dotted line the distribution after deconvolution of the instrumental resolution. Reprinted from Chen et al., *Phys. Rev. Lett.* **98** (2007) 193401, http://prl.aps.org/abstract/PRL/v98/i19/e193401. Copyright (2007) by the American Physical Society

No trace of the previous history of the particle appears in the equation. It is not possible to see if the final product results from a particle that has lost three, five or ten monomers. The reason for this is the very strong separation of time scales between two consecutive decay constants in a decay chain. Unless there are very good arguments for assigning a specific particle size as the unobserved precursor, claims to that effect should be taken cum grano salis.

Figure 7.8 shows a rare example of measured internal energy distributions. The example is from a work on C_{60} fragments. The technique used was to measure the kinetic energy of the H^- produced in a collision between a H^+ with a known kinetic energy and a relatively cold neutral C_{60}. Application of energy conservation yields the internal energy of the ionized molecule directly. The hydrogen anions were measured in coincidence with the ions which enabled to distinguish between the different ions in overlapping regions. The abscissa is the energy deposited in the neutral C_{60} fullerene and the ordinate the ions counts. The curves therefore do not give the energy distributions directly; the consumed energy for the formation of the ions should be subtracted. For C_{58}^+ this amounts to the sum of the ionization and dissociation energies of C_{60}, $D_{60} + \phi_{60}$. For C_{56}^+ a further D_{58} needs to be subtracted, etc.

For these fullerene molecules the value of G is about 30, and the heat capacities of $C_N \approx 3N - 7 \approx 140{-}160$ is close to the limit of $G^2/2e = 160 - 170$ of the validity of the theory presented in this chapter so far, hence the strong deviations from the box-shape distributions valid for small particle sizes. The FWHM of the C_{58} energy distribution is a little below 10 eV, in reasonably agreement with the dissociation energy of 8 eV. The steepest slopes on either side of the peak are, with the peak intensity normalized to unity, equal to 0.2 eV^{-1}. This also agrees very well with the value $G^2/eC_N E_a = 0.2$ eV calculated with the above numbers.

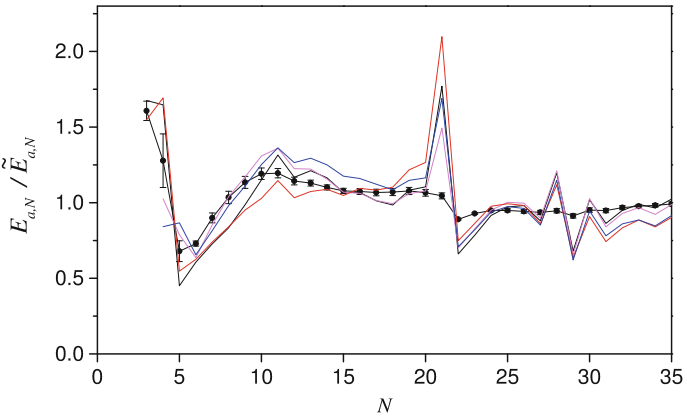


Fig. 7.9 Relative evaporative activation energy for protonated water clusters (filled circles), calculated from the data shown in Fig. 7.7. Also the stability function in Fig. 7.7 are shown as colored lines for the four spectra. The very strong abundance variations result from relatively small variations in the activation energies and the effect of the derivative term in (7.60) on the abundances is clearly seen at $N = 21$ and $N = 28, 29$. The plotted activation energies have not been corrected for the product and reactant caloric curve energy offsets

Figure 7.9 shows the relative evaporative activation energies obtained by inversion of (7.60) with the experimental data for protonated water clusters shown in Fig. 7.7. Four different spectra were analyzed separately. The uncertainties give the standard deviation on the mean value from these four spectra.

7.3 Evaporation of Large Standard Particles

When (7.3) is used to calculate the survival probability for a certain particle size at time t, it is implicitly assumed that t is the time that has elapsed after some experimentally well-defined excitation time. In principle it is really the time that has elapsed since the particle was created from the decay of the precursor. These two times are different and equating them is a good approximation only if the precursor itself is produced practically instantaneously after the excitation. In decay chains, defined as a sequence of decays for a specific particle starting as size N_0;

$$N \leftarrow N+1 \leftarrow \cdots \leftarrow N_0 - 1 \leftarrow N_0, \tag{7.62}$$

the criterion on the rate constants is that they obey the inequality

$$k_{N+n}(E + E_{a,N+n}) \gg k_N(E), \quad n \geqslant 1. \tag{7.63}$$

This holds for small particles, but not for large particles.

How large particles need to be for (7.63) to fail can be seen by expressing the rate constant with the microcanonical temperature and, for the sake of argument, keeping all parameters constant. A calculation of two consecutive rate constants gives, with $C_N \geq E_a/T$:

$$k_{N+1}(E+E_a) = \omega \mathrm{e}^{-\frac{E_a}{T(E+E_a)}} = \omega \mathrm{e}^{-\frac{E_a}{T(E)+E_a/C_N}} \approx \omega \mathrm{e}^{\frac{-E_a}{T(E)}} \mathrm{e}^{\frac{E_a^2}{C_N T(E)^2}} = k_N(E) \mathrm{e}^{\frac{E_a^2}{C_N T(E)^2}}. \tag{7.64}$$

This is not a good approximation for the rate constant when the argument of the last exponential is comparable to or larger than unity and should not be used in these situations, but it is sufficient to tell us that the two rate constants will be comparable when the heat capacity becomes so large that the argument of the exponential becomes less than some value which we can set to unity:

$$1 \gtrsim \frac{E_a^2}{C_N T(E)^2} \Rightarrow C_N \gtrsim G^2, \quad \text{(large particle limit)}. \tag{7.65}$$

The last equality follows from $G = \ln(\omega t) = E_a/T$.[1] In connection with energy distributions in evaporative processes, the magnitude of C_N/G^2 is thus the measure for whether the particles are small or large. Using the estimate $C_N = 3N$ and the value $G = 25$ gives the size

$$C_N \geqslant G^2 \Rightarrow N \gtrsim 200. \tag{7.66}$$

In (7.8) the slopes of the energy distributions were calculated to $G^2/eC_N E_a$. Equating twice the reciprocal of this value to the width of the total width of the distributions of E_a gives an alternative value of the crossover heat capacity of

$$C_N \gtrsim \frac{G^2}{2e} \simeq 110 - 120 \Rightarrow N \gtrsim 40. \tag{7.67}$$

This is the most realistic estimate when the issue is the shape of energy distributions, but otherwise (7.66) is the most relevant.

The effective evaporation time is calculated with a summation of the decay times of the previous decays in the chain, represented by the reciprocal rate constants. We will initially write the rate constants for size independent dissociation energies with heat capacities that are proportional to the number of atoms in the particle. Those are our standard particles. For the rate constants in a decay chain, the consecutive rate constants for these standard particles are

[1] In this section we will dispense with the primed activation energies and finite heat bath corrections to keep the notation manageable.

$$k_{N+n} \approx k_N \exp\left(n \frac{\mathrm{d}\ln(k_N)}{\mathrm{d}N}\right) = k_N \exp\left(n \frac{G^2}{C_N}\right). \tag{7.68}$$

The approximation in this equation consists of truncating the expansion of $\ln(k)$ in N at first order. The average time, t_a, elapsed between the excitation and the decay of size $N+1$ into size N is then calculated as the sum:

$$t_a = \sum_{n=1}^{\infty} k_{N+n}^{-1} = k_N^{-1} \sum_{n=1}^{\infty} \mathrm{e}^{-nG^2/C_N} = k_N^{-1} \frac{\mathrm{e}^{-G^2/C_N}}{1 - \mathrm{e}^{-G^2/C_N}}. \tag{7.69}$$

The rate constants of size N particles that decay at time t will be on the order of

$$k_N \sim \frac{1}{t - t_a}, \tag{7.70}$$

which, combined with (7.69), gives

$$k_N \sim \frac{1}{t}\left(1 - \mathrm{e}^{-G^2/C_N}\right)^{-1}. \tag{7.71}$$

For small heat capacities this expression reduces to $k_N \sim 1/t$, as required. For large heat capacities it becomes

$$k_N \approx \frac{1}{t}\frac{C_N}{G^2} \quad \text{(large particles)}. \tag{7.72}$$

As for small particles, no quantity of dimension energy/temperature appears in the equation. The order of magnitude sign in (7.71) was converted into an approximate sign in (7.72) because the rate for large clusters turns out to be fairly well determined by (7.72), as will become clear later.

The time in (7.72) is the average time it takes before a N-mer (the term used for the size N particle) decays when the decay chain starts with a sufficiently hot and large initial cluster, in the absence of any intrinsic size-specific features in the rate constants. This time will have a non-zero variance because the lifetimes of a number of the slowest decaying precursors in the decay chain are similar. By the Central Limit Theorem, the distribution of decay times will approach a Gaussian distribution when C_N/G^2 approaches infinity. Infinity is never reached but the gross features of a Gaussian distribution is already reached even when only a limited number of terms contribute to the sum.

An example involving sums of exponential decays is shown in Fig. 7.10, for the sum of decay times of 1, 2, 5 and 10 identical exponential decays. The convergence is rather good, considering the strong asymmetry of an exponential decaying function compared to a Gaussian shape. Do note, however, that the tails of the distribution converge significantly slower to a Gaussian than the central region. This is only a problem when you need a precise expression for these tails which is not the case here.

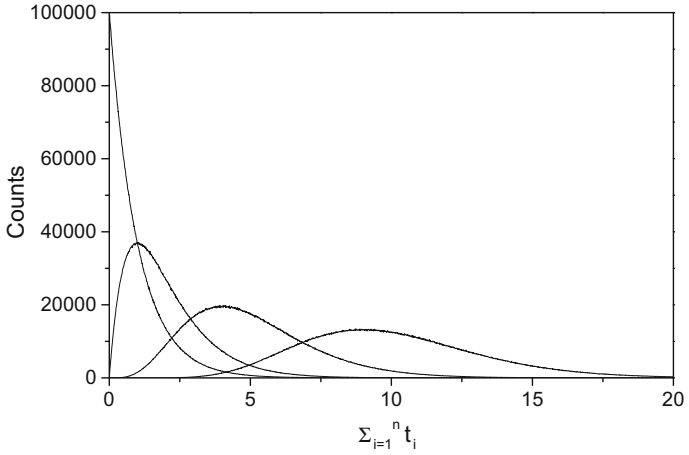


Fig. 7.10 The distributions of sums of decay times, t_i, for $n = 1, 2, 5, 10$ exponential decays, $\sum_{i=1}^{n} t_i$. The distribution is already reasonably close to a Gaussian distribution for five decays, although some asymmetry is apparent. Only the central part is well approximated by a Gaussian function; the convergence of the wings is much slower. The curves are simulated numerically with methods explained in Chap. 9, but could equally well be calculated analytically

Equipped with this information we proceed. The variance of the sum of individual decay times is the sum of the variances, and the variance of the decay time of a single exponential decay with decay constant k is k^{-2}. We therefore get the variance of the total time

$$\sigma_{t_a}^2 = \sum_{n=1}^{\infty} k_{N+n}^{-2} = \sum_{n=1}^{\infty} k_N^{-2} \mathrm{e}^{-2nG^2/C_N} = t^2 \frac{\left(1 - \mathrm{e}^{-G^2/C_N}\right)^2}{1 - \mathrm{e}^{-2G^2/C_N}} \mathrm{e}^{-G^2/C_N}. \tag{7.73}$$

For large particles the standard deviation becomes

$$\sigma_{t_a} = t \left(\frac{G^2}{2C_N} \right)^{1/2}. \tag{7.74}$$

This limit is reached if the number of evaporated monomers exceeds approximately $\Delta n = C_N/G^2$. With $C_N \approx 3N$, this corresponds to a relative loss of mass of $3/G^2$, or a relative loss of excitation energy of $\Delta n E_a/(E_a C_N/G) \approx 1/G$, which is a small fraction, somewhere between 0.03 and 0.05. This criterion is quite easily fulfilled. In a laser excitation experiment, for example, all they translate to is that the particles absorb a single photon with energy $h\nu$ for each $Gh\nu/E_a$ atoms, in addition to a possible starting fee used to heat the clusters to the evaporation limit.

The standard deviation on the time given in (7.74) can be translated into more experimentally accessible quantities. The number of evaporated particles in a decay

chain starting at high excitation energies is also approximately a Gaussian with the standard deviation

$$\sigma_N = \sigma_t k_N = \left(t^2 \frac{G^2}{2C_N}\right)^{1/2} \frac{1}{t} \frac{C_N}{G^2} = \frac{\sqrt{C_N/2}}{G}. \tag{7.75}$$

This is the standard deviation of the size distribution after time t when a size-selected hot and large particle evaporates freely.

The energy and temperature distributions of the final products is also approximately Gaussian with the standard deviations

$$\sigma_E = E_a \sigma_N = \frac{E_a \sqrt{C_N/2}}{G}, \tag{7.76}$$

$$\sigma_T = \sigma_E / C_N = \frac{E_a}{G\sqrt{2C_N}}. \tag{7.77}$$

Finally, the distribution of rate constants has the standard deviation

$$\sigma_k \approx \frac{\partial k}{\partial T} \sigma_T = kG/\sqrt{2C_N} = \frac{1}{t}\sqrt{\frac{C_N}{2G^2}}, \tag{7.78}$$

where we have approximated the derivative of the rate constant with that of an Arrhenius expression. It should be kept in mind that all these calculations of the standard deviation assume a size independent activation energy.

7.4 Rates for Large Particles; General Case

We will now calculate the rate constants with size dependent activation energies. We have seen that for large particles we can approximate ensemble rate constants and temperatures for a specific size N at a given time by a single value, k_N and T_N. When averaging over the excitation energy distribution this means we can use the approximation

$$\langle k_N(E) \rangle \approx k_N(\langle E \rangle), \tag{7.79}$$

because the width of the distribution of k's is small compared to the mean value of k (see (7.78)). This is an approximation which is absolutely and utterly hopeless for small particles and should never be made for these, but which is quite reasonable for large particles.

In general, k_N is size dependent, not only because heat capacities are size dependent, as discussed in the previous section, but more interestingly because also the activation energies are. To find how ensemble average rate constants depend on activation energies, we will first present an alternative derivation of the dependence of the decay rates on time and size for size-independent separation energies. This will be the starting point for the calculation of the general case.

Since the particles are large, we can approximate finite differences by derivatives and vice versa. Then the Boltzmann factor that enters the Arrhenius rate constant, $B_N \equiv \exp(-E_a/T_N)$, develops with time according to

$$\frac{\mathrm{d}B_N}{\mathrm{d}t} \approx \frac{\partial B_N}{\partial T_N}\frac{\partial T_N}{\partial E}\frac{\mathrm{d}E}{\mathrm{d}t} = B_N \frac{E_a}{{T_N}^2}\frac{E_a}{C_N}(-k_N) = -\frac{1}{C_N}\left(\frac{E_a}{T_N}\right)^2 B_N k_N \tag{7.80}$$

where we have used that $\frac{\mathrm{d}T_N}{\mathrm{d}N} = \frac{E_a}{C_N}$, $\frac{\mathrm{d}N}{\mathrm{d}t} = -k_N$. This differential equation also holds approximately when the monomeric unit has internal degrees of freedom, as e.g. water, methane and SF_6 clusters, which carry away thermal excitation energy, in addition to the basic consumption of the separation energy during the process.

One caveat is required here concerning the applicability of (7.80). The cluster needs to be equilibrated, i.e. the temperature decrease after an evaporation must be given by E_a/C_N. For sufficiently large droplets this condition may not be fulfilled, because the monomers always evaporate from the surface and cool it, and the thermal conductance between the interior and the surface is not infinitely high. When the finite thermal conductance plays a role, the surface is colder than the average and evaporation is strongly suppressed. One can understand it heuristically as a reduction in the effective heat capacity. The interior of a droplet tends to be a spectator and the heat capacity that enters the equation is limited to a surface layer of some effective thickness. The effect has been observed in the reduced evaporation rate of liquid water surfaces. Temperature measurements give a measurable and even significant reduction in surface temperature when compared to measured interior temperatures. Exercise 7.12 discusses this question from the point of view of a spherical particle cooled by the rates derived in this section, possessing a finite thermal conductance.

Returning to the situations where this effect can be ignored, the equation is rewritten using the scaled dimensionless time τ, defined as

$$\tau \equiv \frac{\omega t}{C_N}. \tag{7.81}$$

With $\ln(B_N) = -E_a/T_N$ and dropping the subscripts we have

$$\frac{\mathrm{d}B}{\mathrm{d}\tau} = -B^2(\ln B)^2 \tag{7.82}$$

The equation is solved by an asymptotic series with the Ansatz

$$\frac{1}{B(\ln B)^2}\left(1 + \frac{a}{\ln B} + \frac{b}{(\ln B)^2} + \cdots\right) = \tau + \tau_0, \tag{7.83}$$

where τ_0 is the left hand side of (7.83), evaluated at $t = 0$. By inserting the Ansatz in both sides of (7.82), one finds the coefficients $a = -2$ and $b = 6$. In the relevant limit where $-\ln B \gg 1$, this can be approximated by

$$\frac{1}{B}\frac{1}{(\ln B)^2} = \tau + \tau_0 \tag{7.84}$$

The quantitative condition that the particles are initially hot derived in the previous section, can be stated quantitatively as $B(0) \gg B(\tau)$. As was also made clear in that section, this strong inequality is fulfilled even when the initial temperature is only slightly higher than the final. This means that τ_0/τ is small and terms involving this ratio can usually be ignored without any problems. Assuming this, (7.84) becomes

$$(\ln B)^2 B = \tau^{-1} \tag{7.85}$$

From this and the definition of τ it follows immediately that

$$k = \frac{C_N}{(\ln B)^2} t^{-1}. \tag{7.86}$$

This is also the expression for the evaporation *rate* because these two are identical for large particles. Furthermore, it is also the same as the rate for *small* particles found previously, although the origin of the power law is different in the two cases.

A solution to an equation only represents the state or dynamics of a system if it is stable. We therefore need to examine the stability of our solution in (7.85). This is done by calculating how small deviations from the solution develop with time. It turns out that deviations are best parametrized in the form of a factor, H, that divides B:

$$\frac{\mathrm{d}\left(\frac{B}{H}\right)}{\mathrm{d}\tau} = \frac{1}{H}\frac{\mathrm{d}B}{\mathrm{d}\tau} - \frac{B}{H^2}\frac{\mathrm{d}H}{\mathrm{d}\tau} = -\left(\frac{B}{H}\right)^2 \left(\ln\left(\frac{B}{H}\right)\right)^2. \tag{7.87}$$

If we we isolate the derivative of H and use the equation for the time derivative of B we get

$$\frac{\mathrm{d}H}{\mathrm{d}\tau} = -HB\,(\ln)^2 + B\left(\ln\left(\frac{B}{H}\right)\right)^2 \approx (-H+1)\,\tau. \tag{7.88}$$

H is unity for the unperturbed solution and we are therefore interested in the behavior of deviations from unity. To this end we write this equation as

$$\frac{\mathrm{d}(H-1)}{\mathrm{d}\ln\tau} = -(H-1). \tag{7.89}$$

We can convert the differential in $\ln\tau$ into a differential in N by use of the definition of τ, the relation $\mathrm{d}N/\mathrm{d}t = -k$, and the unperturbed solution $kt = C_N/B^2$:

$$\mathrm{d}\ln\tau = \mathrm{d}t/t = -\mathrm{d}N/kt = -\frac{(\ln B)^2}{C_N}\mathrm{d}N. \tag{7.90}$$

The differential equation that determines the dynamics of deviations from the unperturbed solution with loss of monomers is therefore

$$\frac{\mathrm{d}(H-1)}{\mathrm{d}N} = \frac{(\ln B)^2}{C_N}(H-1) \equiv \frac{N_\kappa}{N}(H-1). \tag{7.91}$$

The last equality sign defines the crossover particle size N_κ as

$$N_\kappa \equiv G^2/c_v, \tag{7.92}$$

where c_v is the heat capacity per monomer. The differential equation for $H-1$ describes an exponentially damped function with the damping rate N_κ/N. In spite of the positive sign on the right hand side, it is damped because the value of N decreases with time, and any perturbation of the solution in (7.85) will disappear after roughly N/N_κ evaporation events.

The solution to the unperturbed equation for B and the effect of perturbations are shown in Fig. 7.11. All curves are calculated by solving the differential equation (7.82) numerically. The main frame shows the product τB which is approximately $(\ln B)^2$. The branches are simulation trajectories started at different times with values of B that are intentionally off relative to the asymptotic solution. The convergence is clearly monotonic for all these cases. The inset shows the value of B on a double logarithmic scale. The time scales are in both cases from the shortest relevant (tens of fs) to tens of seconds.[2]

With the solution of (7.82) and the lesson about the stability of the solution, we are in a position to find the effect of size dependent E_a's on the rate constants. The solution of that problem is easiest found if we restrict ourselves to the physically interesting cases where the variations of the E_a's are relatively small. We calculate it by returning to (7.80) and retain the term that accounts for the variation of the activation energies with size:

$$\frac{\mathrm{d}B_N}{\mathrm{d}t} = -\frac{1}{C_N}\left(\frac{E_{a,N}}{T_N}\right)^2 B_N k_N - \frac{\Delta_1 E_{a,N}}{T_N} B_N(-k_N), \tag{7.93}$$

where we used the finite difference $\Delta_1 E_{a,N} \equiv E_{a,N+1} - E_{a,N}$ for the derivative with respect to size. In terms of the scaled time τ the equation reads:

$$\frac{\mathrm{d}B_N}{\mathrm{d}\tau} = -B_N^2(\ln B_N)^2 + B_N^2 \ln(B_N) C_N \frac{\Delta_1 E_a}{E_a}, \tag{7.94}$$

where $T_N = E_{a,N}/B_N$ was used. As before, the variations of the activation energies are most efficiently handled by introducing the dividing function H_N, with the N-dependence shown explicitly:

[2]The latter time scale has recently been exceeded experimentally in cryogenic electrostatic storage rings, and the experimentally relevant upper limit is now above 10^3 s. The curves can safely be extrapolated to these values.

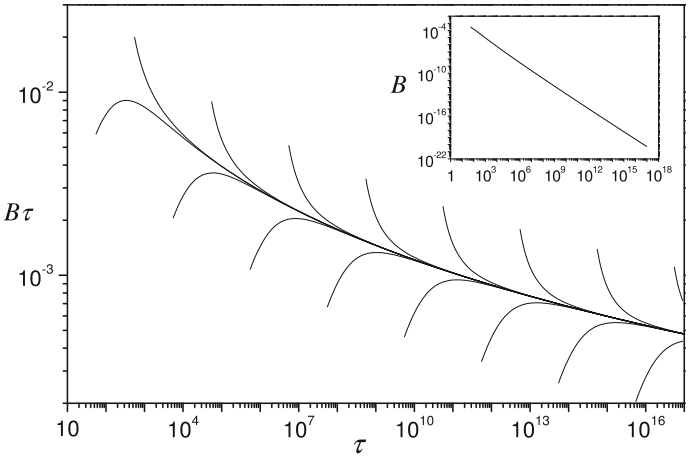


Fig. 7.11 The solution of the equation of motion for B, represented as $B\tau$. The branches converging on the asymptotic solution demonstrate the convergence of B when initialized off that solution. The inset shows the value of B

$$B_N \equiv B/H_N, \tag{7.95}$$

where B is the solution in (7.83) to the constant E_a situation. Substituting (7.95) into (7.94) and assuming that $|\ln B| \gg |\ln H_N|$ (i.e. small variations in $E_{a,N}$), and $|\ln B| \gg 1$ (i.e. long times compared with ω^{-1}), yields an equation for the temporal development of H_N:

$$\frac{\mathrm{d}H_N}{\mathrm{d}\ln\,\tau} = -(H_N - 1) + \frac{\Delta_1 E_{a,N}}{E_{a,N}} \frac{C_N}{\ln B}. \tag{7.96}$$

This is solved to give

$$H_N - 1 = \int^{\ln\,\tau} \frac{\tau'}{\tau} \frac{\Delta_1 E_{a,N'}}{E_{a,N'}} \frac{C_{N'}}{\ln\,B'} \mathrm{d}\ln\,\tau' \tag{7.97}$$

Since, as we have used before, $\mathrm{d}\ln\,\tau = -\mathrm{d}N(\ln B)^2/C_N$, (7.97) can be transformed into

$$H_N - 1 = \int^{N} \frac{\tau'}{\tau} \frac{\Delta_1 E_{a,N'}}{E_{a,N'}} \ln B' \mathrm{d}N' \tag{7.98}$$

With the linear expansion

$$\ln\,\tau' = \ln\,\tau - (N' - N)\frac{(\ln B)^2}{C_N} \tag{7.99}$$

and ignoring the N-dependence of $\ln B'$, this yields

$$H_N - 1 = \ln B \int^N \frac{\Delta_1 E_{a,N'}}{E_{a,N'}} \exp\left(\frac{N - N'}{C_N}(\ln B)^2\right) \mathrm{d}N' \tag{7.100}$$

or, in the discrete version,

$$H_N = 1 + \ln B \sum_n^{N_0 - N} \frac{\Delta_1 E_{a,N+n}}{E_{a,N+n}} \mathrm{e}^{-n\frac{N_\kappa}{N}}, \tag{7.101}$$

where the definition of $N_\kappa \equiv G^2 N / C_N$ has been used. N_0 is the initial size of the particle. With this expression for H_N the result for k_N is

$$k_N = \omega B \left[1 + \ln B \sum_n^{N_0 - N} \frac{\Delta_1 E_{a,N+n}}{E_{a,N+n}} \mathrm{e}^{-n\frac{N_\kappa}{N}}\right]^{-1}. \tag{7.102}$$

The result can be understood with reference to the curves shown in Fig. 7.11. The change from the unperturbed value of $k = \omega B$ to k_N is induced by a source term in (7.96) proportional to $-\Delta_1 E_a / T_N$; A change in the separation energy which occurs upon evaporation will change the separation energy-to-temperature ratio away from the solution given by (7.86) by this amount. A positive value of $\Delta_1 E_a$ will lower the value of H_N and increase the decay constant because $\ln B$ is negative. With this higher rate the particle will cool faster. This faster cooling acts as a negative feedback and provides an efficient focusing mechanism for the rate constant to bring it back to the universal attractive curve for k. Obviously an analogous reasoning holds for negative $\Delta_1 E_a$'s.

It takes, however, a finite number of evaporative cooling steps to return to the universal curve, and that number is $C_N/(\ln B)^2 \equiv N/N_\kappa$. This is also the number of evaporated monomers required to obtain a fully developed evaporative spectrum and, as we have seen, it is sufficient that the particles lose one or two percent of their mass in evaporative processes to fulfill this criterion. If this is realized physically, one can set N_0 in (7.102) equal to infinity.

The expression for H_N can quite often be approximated further. When the variations of E_a with N are sufficiently smooth, the first differences of separation energies can be set constant in the summation, which then reduces to the factor $N/N_\kappa = C_N/(\ln B)^2$, or

$$H_N = 1 + \frac{\Delta_1 E_{a,N}}{E_{a,N}} \frac{C_N}{\ln B} \tag{7.103}$$

and thus

$$k_N = \omega B \left[1 + \frac{\Delta_1 E_{a,N}}{E_{a,N}} \frac{C_N}{\ln B}\right]^{-1}. \tag{7.104}$$

It is necessary that variations in the dissociation energies do not induce too large size-to-size changes in the associated rate constants for these results to hold. Quantitatively, we are restricted to describing cases for which $\left|\frac{\Delta_1 E_a}{T_N}\right| \lesssim 1$, equivalent to a change in separation energy which is less than about five percent or so for consecutive particles in the decay chain, depending on the precise value of G. For an experimentally operational test of this criterion, one can use that a change in dissociation energy by T_N corresponds to a relative change of the abundance by a factor e^{-1}. For experimental spectra where the abundances of two consecutive, large ($N > N_\kappa$) particles vary by more than this factor, the rate constants calculated here would need further development.

It is intuitively clear that the rate constants calculated in (7.102, 7.104) play an important role in determining the variations in the particle abundances for large particles. To find the precise role, we will first establish the relation between abundances and rate constants for a single decay chain. Specializing to a single decay chain greatly simplifies the problem if we ignore the small stochastic kinetic energy carried away in the evaporation processes, because then all decay constants are fixed when the initial temperature is given and the only stochastic element in a decay chain is the time at which the evaporations occur.

Consider the simultaneous distribution of decay times in a chain starting at size N_0 and reaching size N:

$$P(t_N, t_{N+1}, \ldots, t_{N_0}) \prod_{j=N}^{N_0} \mathrm{d}t_j = \prod_{j=N}^{N_0} k_j \mathrm{e}^{-k_j t_j} \mathrm{d}t_j. \tag{7.105}$$

The t_j's are the times since the size j particle was produced by evaporation from the precursor particle of size $j+1$. The probability that size N decays into size $N-1$ at time t is the integral of this distribution with respect to all the intermediate lifetimes, with the restriction that these are positive and sum up to t:

$$k_N I_N = \int_0^\infty \delta\left(\sum_{j=N}^{N_0} t_j - t\right) \prod_{j=N}^{N_0} k_j \mathrm{e}^{-k_j t_j} \mathrm{d}t_j. \tag{7.106}$$

This is an $N_0 - N + 1$ fold integral and the integral sign is a shorthand for all integrals over the t_j's, which all have the limits 0 and ∞. We can decouple the integrals if we express the δ function as the Fourier transform of 1,

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kt} \mathrm{d}k. \tag{7.107}$$

After insertion of this into (7.106) and interchanging the order of integration, the integrals over the times are easily performed:

$$k_N I_N = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kt} \prod_{j=N}^{N_0} \frac{k_j}{k_j + \mathrm{i}k}\mathrm{d}k. \tag{7.108}$$

The integrand in this expression is the Fourier transform of the desired probability. It can be calculated by the method of residues with the result (assuming all rate constants are different)

$$k_N I_N = \sum_{j=N}^{N_0} k_j \, \mathrm{e}^{-k_j t} \prod_{n\neq j} \frac{k_n}{k_n - k_j}, \tag{7.109}$$

which can be useful in numerical simulations.

Returning to (7.108), we can use it to find the relation between the abundances and the rate constants. The derivative of I_N with respect to k_N is

$$\frac{\partial I_N}{\partial k_N} = -\frac{1}{2\pi k_N}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kt} \left(\prod_{j=N}^{N_0} \frac{k_j}{k_j + \mathrm{i}k}\right) \frac{1}{k_N + \mathrm{i}k}\mathrm{d}k. \tag{7.110}$$

There are effectively C_N/G^2 rate constants of magnitude k_N in this equation. The product of ratios of rate constants in the integral, as in (7.108), is therefore suppressed when k exceeds the value given by

$$\left|\frac{k_N}{k_N + \mathrm{i}k}\right|^{C_N/G^2} \simeq \frac{1}{\mathrm{e}}. \tag{7.111}$$

With a leading order expansion in k, this limit is seen to be reached for the value

$$k \simeq k_N \left(\frac{2G^2}{C_N}\right)^{1/2}, \tag{7.112}$$

which does not exceed k_N by the assumption that C_N is large. We can therefore calculate (7.110) by expanding the last factor of the integrand in k;

$$\frac{\partial I_N}{\partial k_N} = -\frac{1}{2\pi k_N}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kt} \left(\prod_{j=N}^{N_0} \frac{k_j}{k_j + \mathrm{i}k}\right) \left(\frac{1}{k_N} - \mathrm{i}\frac{k}{k_N^2} - \frac{k^2}{k_N^3} + \cdots\right)\mathrm{d}k \tag{7.113}$$

By comparison with (7.108) we see that the first term integrates to $-I_N/k_N$. Taking the time derivative of I_N in (7.108) we see that this is identical to the integral of the second term, apart from a factor k_N^{-2}. The third and higher terms correspond similarly to higher order derivatives. We therefore end with

$$\frac{\partial I_N}{\partial k_N} = -\frac{I_N}{k_N} + \frac{1}{k_N^2}\frac{\partial I_N}{\partial t} + O\left(\left(\frac{G^2}{C_N}\right)^{3/2}\right), \tag{7.114}$$

where (7.112) was used for the estimate by the third and higher order terms.

When we are looking at the specific decay chain with an initial energy such that

$$t = \sum_{j=N}^{N_0} k_j^{-1}, \tag{7.115}$$

I_N will be maximal and its time derivative therefore zero. For the decay chains that have higher (lower) initial energies, the rate constants are larger (smaller), and the decay chains have progressed further (shorter), which makes the time derivative in (7.114) negative (positive). Averaging over the whole distribution will therefore cancel the contributions from the high and low energy sides to a very good approximation. After this averaging we therefore have for (7.110)

$$\frac{\partial I_N}{\partial k_N} \approx -\frac{I_N}{k_N}. \tag{7.116}$$

This is solved to give

$$I_N \propto k_N^{-1}. \tag{7.117}$$

The relation was tested in Monte Carlo simulations (see Chap. 9 for more on that technique) with a generic species of particles, decaying with a rate constant of Arrhenius type and a heat capacity of $3N - 7$. The particle sizes were around $N = 900$ and had identical dissociation energies equal to unity, except for N = 925 for which the value was set to 0.975 through 1.025 in steps of 0.005 in the 11 separate simulations performed.

Figure 7.12 shows the result of these simulations. The main plot contains the values of k_N, I_N^{-1} for all sizes and values of $E_{a,925}$. The linear relation in (7.117) is seen to hold very well. The simulations also gave data for the energy distributions for the $N = 925$ cluster. They are shown in Fig. 7.13.

We have now found that abundances for a given particle size in a decay chain are related to rate constants and given a formula for calculation of the rate constants. We need to sum decay chains over all possible initial conditions, which effectively is a summation over energies. This is done by covering the (N, T)-plane densely with decay chains and summing them up, just as for the small cluster calculation visualized in Fig. 7.6. After some inspection of that figure and contemplation, we realize that the density of decay chains that terminate at size N is proportional to the separation energy, because the slope of the curves in the (N, T)-plane has this proportionality. Figure 7.6 represents small heat capacity clusters but this conclusion holds also for large particles. Also the amount of the initial distribution that is projected into the final

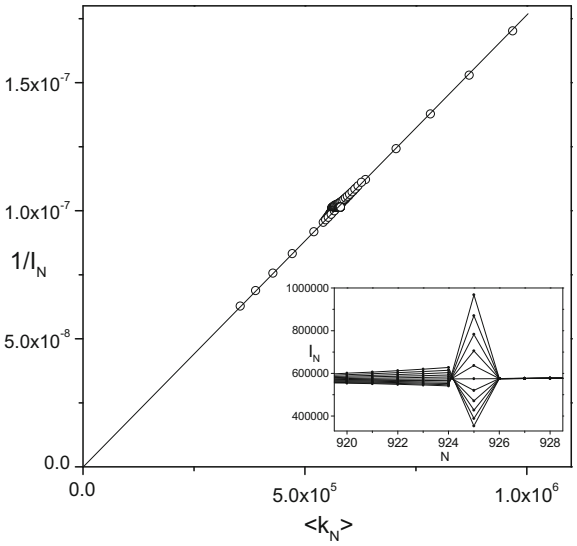


Fig. 7.12 Abundances and rate constants for large clusters for different values of the $N = 925$ dissociation energy, generated in MC simulations. The inset shows some of the abundances from the simulations

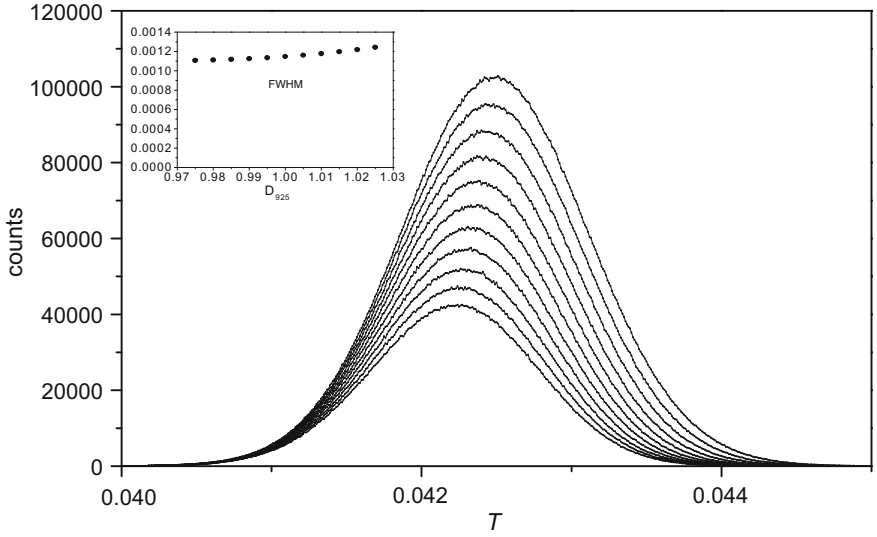


Fig. 7.13 Energy distributions from the same simulations as used in Fig. 7.12, expressed in terms of the temperature. The values of $E_{a,925}$ are 0.975 to 1.025 in steps of 0.05, from smallest to highest intensities. The shapes are very well fitted with Gaussian shapes. The inset shows the Full Width at Half Maximum (FWHM) of the distributions as a function of $E_{a,925}$. The variation is slow compared to the variation of the intensities with $E_{a,925}$, in spite of the apparent broadening seen in the figure

temperature is inversely proportional to the separation energy of the initial particle size. Hence

$$I_N \propto \frac{E_{a,N}}{E_{a,N_0}}. \tag{7.118}$$

In experiments one usually measures the final size and rarely the initial size, in particular if one works with broad energy and size distributions, and initial sizes will more often than not be averaged over a broad mass range. This means that one can substitute the value of E_{a,N_0} with a mean value E_a. Then (7.118) reduces to

$$I_N \propto E_{a,N}. \tag{7.119}$$

7.5 Large Particle Abundances

All the ingredients are now in place to express abundances as functions of separation energies. With the results from the previous section, (7.117, 7.119), we have:

$$I_N \propto E_{a,N} \frac{1}{k_N} = E_{a,N} \frac{H_N}{k}. \tag{7.120}$$

Introducing the expression for H_N and leaving out the smooth function k, (7.120) yields:

$$I_N \propto E_{a,N} \left(1 + \ln\, B \sum_{n=0}^{\infty} \frac{\Delta_1 E_{a,N+n}}{E_{a,N+n}} \mathrm{e}^{-n\frac{N_K}{N}} \right). \tag{7.121}$$

For $\Delta_1 E_a / E_a \ll 1$ this expression can to a good approximation be rewritten as

$$I_N \propto E_{a,N} + \ln\, B \sum_{n=0}^{\infty} \Delta_1 E_{a,N+n}\, \mathrm{e}^{-n\frac{N_K}{N}}. \tag{7.122}$$

If $\Delta_1 E_{a,N}$ varies sufficiently slowly we can approximate $\Delta_1 E_{a,N+n}$ with the constant value $\Delta_1 E_{a,N}$ in the sum in (7.122), as already noted when discussing the H function. Quantitatively the approximation requires that $\Delta_2 E_{a,N} \equiv E_{a,N+2} - 2E_{a,N+1} + E_{a,N} \ll \Delta_1 E_{a,N}((\ln B)^2/C_N)^2$. If this is the case, the abundances become

$$I_N \propto E_{a,N} + \frac{C_N}{\ln B} \Delta_1 E_{a,N}. \tag{7.123}$$

When analyzing experimental data the constant of proportionality can be found in complete analogy with the small particle case by dividing out a smooth abundance distribution on the left hand side and $\overline{E}_{a,N}$ on the right hand side.

One should note that this is identical to the result for small clusters some sixty equations ago, (7.60), at this level of approximation ($G = -\ln B$), apart from the

average $(E_{a,N} + E_{a,N+1})/2$ that appears in that equation, as opposed to the $E_{a,N}$ here. But it is possible to replace the $E_{a,N}$ with the average of (7.60) under the same assumptions as we calculated the sum in (7.121). The two equations are therefore identical. Since the physical situation and the mathematical derivations are different for the two cases, this is a non-trivial fact. We will not look a gift horse in the mouth and just accept the strong suggestion that the interpolation between the two regimes can be made with good confidence. This is supported by numerical simulations which have corroborated that the intermediate region with $N \sim N_\kappa$ behave exactly as the low and high N limits.

7.6 Kinetic Energy Release Revisited

Kinetic energy release distributions depend essentially on two quantities, viz. the capture cross section and the temperature of the product. The details of this for some interaction potentials have been given in Chap. 5. When particles evaporate freely, the systematics of the processes described in this chapter add a constraint on the temperature that makes it possible to extract information on the binding energies from measured kinetic energies. First and foremost because absolute energies are measured and this provides a much needed energy scale.

The procedure involves four steps. First a kinetic energy distribution is measured for the hot particles. Secondly, this distribution is converted to a product temperature. The third step is to convert the product temperature to an effective emission temperature, expressed in terms of the Gspann parameter and the binding energy. The final step is to solve for the binding energy.

Once the measured kinetic energy release distribution has been delivered from the lab, the analysis starts with relating the kinetic energy to the microcanonical temperature of the product particle, T_d. As described in Chap. 5, this amounts to a measurement or calculation of the energy dependence of the capture cross section in the reverse process. In the absence of a barrier to attachment, the average kinetic energies, $\langle\varepsilon\rangle$, vary between T_d to $2T_d$, or parametrized with $1 \leq \alpha \leq 2$, as

$$\varepsilon = \alpha T_d. \tag{7.124}$$

Any absolute binding energy determination based on kinetic energy measurements will not be more reliable than the value of α. In sufficiently accurate experiments one can determine not just the average kinetic energy release but the whole distribution. For a spherical distribution it is

$$P(\varepsilon) \propto \sigma(\varepsilon)\varepsilon \mathrm{e}^{-\varepsilon/T_d}, \tag{7.125}$$

as shown in Chap. 5. This will further narrow down the cross section and allow one to extract reasonable values for T_d. In principle one can use this expression to extract the temperature from the large ε behavior of the kinetic energy distribution. This is

rarely experimentally feasible, however, and one must resort to more approximate but experimentally more robust schemes.

This can be achieved by parameterizing the cross section. A popular form due to C.E. Klots is to use a power dependence for the pre-exponential. This is obviously a spherical horse approximation[3] because capture cross sections rarely have such simple energy dependences. It will, however, be a reasonable and useful approximation as long as one has a realistic attitude to its accuracy and, importantly, it will provide the desired robust experimental numbers. We will therefore proceed with the cross section

$$\sigma(\varepsilon) \propto \varepsilon^l. \tag{7.126}$$

The functional form indicates that there is no reverse activation barrier for the process. The majority of cases show no such barrier, but the experimental data allow this to be checked (see Chap. 5 for the signature of a non-vanishing barrier). Likewise, the potential that gives rise to the cross section is implicitly assumed to be spherical symmetric.

With (7.126) we find the average kinetic energy as

$$\langle\varepsilon\rangle = \frac{\int \varepsilon^{l+2} \mathrm{e}^{-\varepsilon/T_d} \mathrm{d}\varepsilon}{\int \varepsilon^{l+1} \mathrm{e}^{-\varepsilon/T_d} \mathrm{d}\varepsilon} = (l+2)T_d, \tag{7.127}$$

and

$$\langle\varepsilon^2\rangle = \frac{\int \varepsilon^{l+3} \mathrm{e}^{-\varepsilon/T_d} \mathrm{d}\varepsilon}{\int \varepsilon^{l+1} \mathrm{e}^{-\varepsilon/T_d} \mathrm{d}\varepsilon} = (l+3)(l+2)T_d^2. \tag{7.128}$$

The experimentally measured $\langle\varepsilon\rangle$ and $\langle\varepsilon^2\rangle$ can be used to find the value of l as

$$\frac{\langle\varepsilon^2\rangle}{\langle\varepsilon\rangle^2} = \frac{l+3}{l+2} \Rightarrow l = \frac{3 - 2\frac{\langle\varepsilon^2\rangle}{\langle\varepsilon\rangle^2}}{\frac{\langle\varepsilon^2\rangle}{\langle\varepsilon\rangle^2} - 1}. \tag{7.129}$$

With this l the temperature is easily found to

$$T_d = \frac{\langle\varepsilon\rangle}{l+2}. \tag{7.130}$$

This solves the second problem; How to relate the measured energies to microcanonical temperatures. The third task is to relate the product temperature to the binding energy. For this purpose, one must find the relation between the product temperature and the effective microcanonical temperature for the emission process because the latter is given by the Gspann parameter and we know that relation.

This question has basically already been answered in Chap. 3. Recapitulating the notation, the rate constant is represented as

[3] The physicist's starting point when explaining how to breed the fastest horse in the world.

$$k \approx \omega \left(\frac{E + E_r - E'_a}{E + E_r} \right)^{\overline{s}}, \tag{7.131}$$

where ω is a frequency factor, E is the excitation energy, E_r is the energy offset in the caloric curve for the reactant (the zero point energy if vibrations are harmonic oscillators), $\overline{s}$ is the average microcanonical heat capacity of product and reactant, and $E'_a \equiv E_a + E_r - E_p$, where E_a is the evaporative activation energy and E_p the energy corresponding to E_r but for the product particle.

The kinetic energies are sampled from the decaying particles for which $k = 1/t$. With this condition and $\omega t = \exp(G)$ (7.131) can be solved for the energy of the reactant:

$$E = \frac{E'_a}{1 - \mathrm{e}^{-G/\overline{s}}} - E_r. \tag{7.132}$$

The product energy is $E - E_a$ and with the caloric curve for the product we have

$$E - E_a = s_p T_d - E_p \Rightarrow T_d = \frac{E - E_a + E_p}{s_p}, \tag{7.133}$$

where s_p is the microcanonical heat capacity of the product. Inserting the expression for E we get the product microcanonical temperature, which is also the temperature of the measured kinetic energy distribution.

$$T_d = \frac{1}{s_p} \frac{E'_a}{\mathrm{e}^{G/\overline{s}} - 1}, \tag{7.134}$$

or

$$E'_a = s_p T_d \left(\mathrm{e}^{G/\overline{s}} - 1 \right). \tag{7.135}$$

We should check if this has the correct limit for large $\overline{s}$. It does. An expansion of the exponential to second order gives $E'_a \approx T_d s_p (G/\overline{s} + G^2/2\overline{s}^2)$. If we for the purpose of this expansion ignore the difference between s_p and $\overline{s}$ this reduces to

$$E'_a \approx T_d G \left(1 + \frac{G}{2\overline{s}} \right). \tag{7.136}$$

This is the relation between activation energy and temperature in the leading order finite heat bath approximation (see Exercise 7.8).

When applying (7.135) the reader should keep in mind, as always, the interpretation of E'_a as the true activation energy plus a correction from the difference between the offsets of the two caloric curves involved. A second reminder is that the heat capacities are the microcanonical values, which for well behaved systems is one (k_B) less than the canonical heat capacities. Finally, because of the functional dependence on G, it is important to make a good estimate of its value, in particular for small particles.

7.7 Metastable Decay Fractions

Metastable decay is another useful tool in molecular beam experiments involving clusters. This section deals with the systematics of the number of particles that undergo metastable decay in an experiment on hot particles. The word metastable decay refers to the fact that the decay occurs some time after the creation or excitation of the particles, and that they must therefore be stable on some finite time scale.

Experiments typically consist of a mass selection of the particles at time t_1 and a measurement of the amount of particles that decay between times t_1 and t_2, or possibly with a later starting time, t_1', with $t_1 < t_1' < t_2$. The latter situation involves a few more factors in the calculations but they should all be easy to write down after reading this section. In the interest of simplicity we stick with the former scenario. The ratio of the particles that have undergone a decay between t_1 and t_2 to the total number of particles at t_1 is the metastable fraction, f_N. It will be denoted f_N, and can by definition take values $0 \leqslant f_N \leqslant 1$.

Figure 7.14 gives a typical example from experiments on gold-erbium alloy clusters. In the experiments the cluster are accelerated and mass selected in a hot[4] internal state and the metastable fraction at t_2 measured with a device called a reflectron. This particular alloy does not have a competing dimer evaporation channel ($X_N \rightarrow X_{N-2} + X_2$) which would complicate the picture a little. For $N = 17, 19$, $f_N > 0.5$ and the intensity of the metastable peak in the mass spectrometer exceeds the intensity of the remaining intact reactant peak intensity. The increase of the metastable fraction with size is a consequence of the fact that the heat capacity increases with size, as we will see below, and is very frequently observed in experiments where metastable decay is measured.

If the metastable decay is thermal, which is a fairly safe bet, the metastable fraction can be calculated with the formalism already developed to calculate the abundance variations. For large particles this is very simple because distributions of rate constants can be approximated with single exponentials and the decay is therefore just exponential. We will leave it to the reader to work out the details for these cases.

For particles smaller than N_κ the problem is more complicated. In addition, the results depend critically on the presence or absence of radiative cooling. The non-radiative situation will be treated first. The general expression for the metastable fractions is most easily obtained by stepping back to consider the derivation of the integral of the energy distribution that gives rise to the abundances, I_N. The energy distribution was given in (7.44) which reads

$$P(t_1) = \mathrm{e}^{-k_N(E)t_1}\left(1 - \mathrm{e}^{-k_{N+1}(E+E_{a,N+1})t_1}\right), \tag{7.137}$$

[4] We remind the reader that 'hot' does not mean above human body temperature. It means as hot or hotter than the particles can survive in the experiments in question. Hence 1 K is hot for a helium droplet (see Chap. 12) but 2000 K is cold for a fullerene molecule.

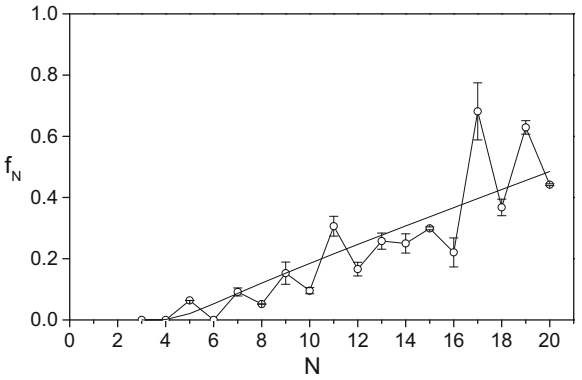


Fig. 7.14 The metastable fraction of $Au_N Er^+$ clusters that decay by evaporation of a single gold atom in a Time-of-Flight mass spectrometer. Evaporated and non-evaporated species are separated experimentally with a device called a reflectron in the spectrometer. The metastable fraction is found as the ratio of evaporated to the sum of evaporated and intact clusters of the same initial size. The line is the function calculated with (7.149) with effective heat capacities given by (3.32), assuming harmonic oscillator values for the d.o.f.'s, but disregarding the size-to-size variations of abundances. The data are published in N. Veldeman et al., *Faraday Discussions* **138** (2008) 147

where the first factor gives the upper limit on the energy distribution and the second factor the lower energy. Mass selecting the particle at time t_1 makes the energy distribution of the un-decayed particle at t_2 equal to

$$P(t_2) = \mathrm{e}^{-k_N(E)(t_2-t_1)} P(t_1). \tag{7.138}$$

The decayed part is therefore

$$\begin{aligned} P(t_1) - P(t_2) = {} & \exp\left(-k_N(E)t_1\right) - \exp\left(-k_N(E)t_2\right) \\ & - \exp\left(-[k_N(E) + k_{N+1}(E + E_{a,N+1})]t_1\right) \\ & + \exp\left(-k_N(E)t_2 - k_{N+1}(E + E_{a,N+1})t_1\right). \end{aligned} \tag{7.139}$$

The only time that multiplies k_{N+1} is t_1 and we will use this fact. Denote the ratio of the two decay constants by $\alpha(E) \equiv k_{N+1}(E + E_{a,N+1})/k_N(E)$, evaluated at the energy where $k_{N+1}(E + E_{a,N+1})$ is equal to $1/t_1$. With this definition the decayed part can be expressed as the integral

$$\Delta E = \int_0^\infty \left[\mathrm{e}^{-k_N(E)t_1} - \mathrm{e}^{-k_N(E)t_2} - \mathrm{e}^{-k_N(E)t_1(1+\alpha)} + \mathrm{e}^{-k_N(E)(t_2+t_1\alpha)}\right] \mathrm{d}E. \tag{7.140}$$

This is the width of the energy distribution that has decayed at t_2. It is not yet a proper decay probability that will give you the metastable fraction because we have not normalized it to the initial width at t_1. Before we do that, we simplify the integral in (7.140) by noting that the exponentials provide the cutoff energies $E_{N,max}(t_i)$, for

whichever of the times t_i may be. Taking differences between these cutoff energies cancel constants that appear in the caloric curves (see (3.6)), and gives us

$$\Delta E(t_2, t_1) = C_N E_{a,N} \left[G(t_1)^{-1} - G(t_2)^{-1} + G(t_2 + t_1\alpha)^{-1} - G(t_1(1+\alpha))^{-1} \right], \tag{7.141}$$

where $G(t_i) = \ln(\omega t_i)$. Usually, one can take the Gspann parameter to be constant, but here we must include its time variation because it is precisely this time dependence that describes the metastable decay in these equations.

The normalization is easily found as $\Delta E(\infty, t_1)^{-1}$, which gives

$$f_N = \frac{\Delta E(t_2, t_1)}{\Delta E(\infty, t_1)} \tag{7.142}$$

$$= \frac{G(t_1)^{-1} - G(t_2)^{-1} + G(t_2 + t_1\alpha)^{-1} - G(t_1(1+\alpha))^{-1}}{G(t_1)^{-1} - G(t_1(1+\alpha))^{-1}}. \tag{7.143}$$

Given the experimentally determined metastable fractions on the left hand side, we can then solve (7.142) for α numerically.

If the structure function is determined, it is possible to extract more information from the data. Use the relation between the energy interval and the abundance variations found previously:

$$\Delta E(\infty, t_1) = \tilde{E}_{a,N} \frac{I_N}{\tilde{I}_N}. \tag{7.144}$$

Replacing the denominator of (7.142) with this value gives

$$f_N \frac{I_N}{\tilde{I}_N} = \frac{E_{a,N}}{\tilde{E}_{a,N}} C_N \left(G(t_1)^{-1} - G(t_2)^{-1} + G(t_2 + t_1\alpha)^{-1} - G(t_1(1+\alpha))^{-1} \right). \tag{7.145}$$

If the structure function is not just known but also has been inverted as described previously in this chapter (which is obviously a must if it is known), the dimensionless fraction $E_{a,N}/\tilde{E}_{a,N}$ is also known, and because also α has been calculated from the data, we can extract the heat capacity.

This is the rigorous approach. Real life may not provide all the information you need for this. If the structure function is not known, for example, one has to resort to using average values, $E_{a,N}/\tilde{E}_{a,N} \approx 1$ and $I_N/\tilde{I}_N \approx 1$, and derive a trend in the heat capacity with size, with the risks that entails. Alternatively, one can guesstimate a heat capacity and extract ratios of dissociation energies from the relation

$$f_N = \frac{C_N E_{a,N} \left[G(t_1)^{-1} - G(t_2)^{-1} + G(t_2 + t_1\alpha)^{-1} - G(t_1(1+\alpha))^{-1} \right]}{\frac{1}{2}(E_{a,N} + E_{a,N+1}) + \frac{C_N}{G}(E_{a,N} - E_{a,N+1})}, \tag{7.146}$$

where (7.59) was used. Use of that equation reminds us that the heat capacity that should be used in the equation is a modified value approximately equal to the average of reactant and product heat capacities, at least for not very small particles.

It is of interest to have a short time approximation for the full expression in (7.145). In this connection short time means that

$$t_2 \ll t_1\alpha. \tag{7.147}$$

Physically, the demand on t_2 corresponds to a t_2 where the decay has not reached close to the lower energy edge of the initial distribution that was created at t_1. This is not necessarily a short time compared to t_1 and can in fact easily be orders of magnitude longer. The approximation below is actually likely to cover a good part of the situations encountered in practical applications. In this regime the last two terms in the numerator in (7.142) cancel to a good approximation and the metastable fraction becomes

$$f_N \approx \frac{G(t_1)^{-1} - G(t_2)^{-1}}{G(t_1)^{-1} - G(t_1(1+\alpha))^{-1}}. \tag{7.148}$$

In the numerator we use $G(t_2) = G(t_1) + \ln(t_2/t_1)$ and expand the reciprocal of this in $\ln(t_2/t_1)/G(t_1)$. The result is the short time metastable fraction

$$f_N \frac{I_N}{\tilde{I}_N} = \frac{E_{a,N}}{\tilde{E}_{a,N}} \frac{C_N}{G(t_1)^2} \ln\left(\frac{t_2}{t_1}\right)\left(1 - \frac{\ln(\frac{t_2}{t_1})}{G(t_1)^2} + \cdots\right). \tag{7.149}$$

We note (again) that the heat capacities determined from the data are not the canonical values. As shown in Chap. 3 with (3.2), the value extracted needs to be converted to an equivalent number of harmonic oscillator degrees of freedom, $\bar{s}$, with the help of a reasonable estimate of G. For values of $\bar{s}$ above 20 or so the correction is small, however.

The quality of (7.142) can be judged from the simulation in Fig. 7.15 which compare the numerical integration of the exact equation and our approximate relation for a $N = 20$ particle. The rate constants for $N = 20, 21$ were given by the expression

$$k_N = 10^{15}\,\mathrm{s}^{-1}\left(\frac{E - E_{a,N} + 0.01 \cdot (3N-9)\,\mathrm{eV}}{E + 0.01 \cdot (3N-6)\,\mathrm{eV}}\right)^{3N-7}. \tag{7.150}$$

The activation energy of $N = 21$ is 3 eV in all cases, whereas that of the $N = 20$ particle was set to the values 2.7, 3, and 3.3 eV. It is a slight distortion of our standard system to use identical powers on energy in numerator and denominator but this makes the definition of the frequency factor simple. The fastest decay occurs not surprisingly for the least bound 20-mer and the slowest for the strongest bound. Equation (7.60), including the corrections on the activation energy from (3.17), was used to calculate α.

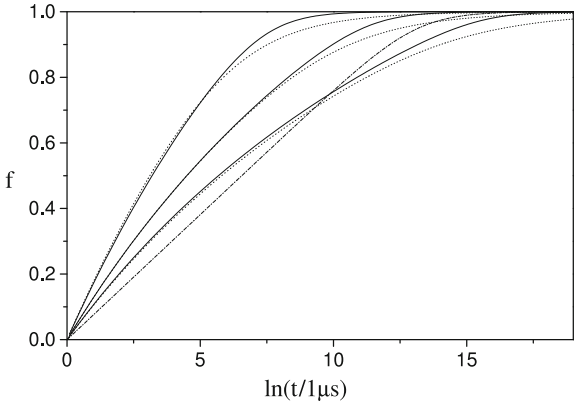


Fig. 7.15 Metastable decay calculated with (7.142) (full line) and numerical integration of (7.150) (dotted line). The dashed-dotted line is Klots' prediction for $E_{a,20} = E_{a,21} = 3$ eV, which should be compared with the middle dotted line. Other parameters are given in the text

The approximate analytical result agrees quite well with the exact numerical, although discrepancies are seen at the longest times. These are a consequence of treating the energy distributions as sharply defined at the lower edge, which is an approximation made between (7.140) and (7.141). Experiments where the metastable fraction approaches unity will be of very long duration. In these situations it is likely that radiative cooling will interfere with the spontaneous decay and quench the decay at some time and render the remaining fraction stable. More accurate expressions for this limit are therefore usually of minor interest.

C. Klots has derived a complicated formula for the metastable fraction that has been widely used by the cluster community in the interpretation of experimental decays. For the special case where $E_{a,N} = E_{a,N+1}$ it reduces to

$$1 - f = \ln\left[1 + \left(\mathrm{e}^{\alpha W_N} - 1\right) t_1/t_2\right] / \alpha W_N, \tag{7.151}$$

where, for this case,

$$\alpha W_N = \frac{G^2}{3N-7}\left[1 - G/2(3N-7) + (G/(3N-7))^2/12\right]^{-1}. \tag{7.152}$$

This is also shown in Fig. 7.15. The curve should be compared with the middle simulated curve. We will not go into details with the more general case of this theory.

7.8 Radiative Cooling

At sufficiently long times, such as those experienced in ion traps or storage rings, radiative cooling will be important. This is readily understood from a comparison of frequency factors for photon emission on one hand and those for emission of massive

particles on the other. The photon emission rate constant frequency factor is limited by the dipole sum rule. The competing massive particle emission frequency factors are several orders of magnitude higher. At high excitation energies, the energy emitted is therefore predominantly emitted as particles, not light; under these high temperature conditions, photon emission is reduced to an extremely marginal occurrence. At low temperatures the picture is reversed, because the difference in activation energy for photon and massive particle emission, $h\nu$ and E_a, suppresses the latter strongly, and for systems with $E_a > h\nu$ the dominant cooling channel will always be radiative at sufficiently low temperatures. For a freely cooling ensemble of particles, where temperature decreases monotonically with time, a cross-over therefore exists from particle emission at short times to radiative cooling at long times.

As for unimolecular decays, size also matters for photon emission. If a particle has a sufficiently small heat capacity, the emission of a single photon will quench any further unimolecular decay for a long time and effectively prevent any further of these, because the next radiative decay will occur before the next unimolecular decay. For quenching to occur, the photon energies need to have a certain magnitude

$$h\overline{\nu}\frac{\mathrm{d}\ln k_a}{\mathrm{d}E} \gtrsim 1, \tag{7.153}$$

where $h\overline{\nu}$ is the mean energy of the radiated photons and k_a is the unimolecular rate constant for the competing process, which can be loss of an atom, a larger fragment or an electron. With the Arrhenius expression for k_a with activation energy E_a and the Gspann parameter G, one gets

$$h\overline{\nu}\frac{E_a}{T^2C} = \frac{h\overline{\nu}}{E_a}\frac{G^2}{C} \gtrsim 1, \tag{7.154}$$

for the hottest particles in the ensemble. The heat capacity C is an effective value which is somewhere between the precursor and the product value, as discussed in Chap. 3. For simplicity we can here use the reactant value. This gives the condition for one-photon quenching

$$C \lesssim G^2\frac{h\overline{\nu}}{E_a}. \tag{7.155}$$

Radiative cooling is relevant only if the (reciprocal) time constant of that decay exceeds the unimolecular ditto. For the high energy quenching photon emission, the radiative time constant is simply the photon emission rate constant. For emission of photons with the average energy $h\nu$, this implies the constraint

$$\frac{\omega_{ph}\mathrm{e}^{-h\nu/T}}{\omega_a\mathrm{e}^{-E_a/T}} \gtrsim 1 \Rightarrow \frac{h\nu}{E_a} \lesssim 1 + G^{-1}\ln\left(\frac{\omega_{ph}}{\omega_a}\right). \tag{7.156}$$

The equation was calculated disregarding the stimulated emission contribution, which is usually only a minor correction for the relatively high photon energies involved.

When photon energies are below the limit in (7.155), the relevant time constant can be found by considering the effect of radiation on a particle with a δ function initial energy distribution, and calculation of its subsequent unimolecular decay. The time dependence of the rate constant is calculated with an expansion of its logarithm with respect to time. To first order in the emitted power and the time we get:

$$k_a(E,t) = k_a(E,0)e^{-wt}, \tag{7.157}$$

where the radiative constant w is defined in terms of the radiated power $P_r = h\overline{\nu}k_{ph}$ as

$$\tau^{-1} = w \equiv \frac{\mathrm{d}\ln k_a(E,0)}{\mathrm{d}t} = \frac{E_a}{CT^2}P_r = \frac{G^2}{CE_a}h\overline{\nu}k_{ph}. \tag{7.158}$$

This describes the continuous cooling situation to leading order in the expansion of the rate constant. This will be a poor approximation at longer times. Usually the decay is so strongly suppressed that the error will not have serious consequences for the interpretation of data, but for good statistics data higher order terms need to be included. For calculations involving the cross-over time, the approximation is more than sufficient, though. The criterion for radiative cooling to be dominant (for single photon emission) is that the rate constant for massive particle loss, say atomic evaporation, is less than the $w \equiv 1/\tau$ defined in (7.157). This gives

$$k_a\tau \lesssim 1 \Rightarrow \frac{\omega_a}{\omega_{ph}}\frac{E_a}{h\overline{\nu}}\frac{C}{G^2}\mathrm{e}^{-G}\left(\exp\left(G\frac{h\overline{\nu}}{E_a}\right)-1\right) \lesssim 1. \tag{7.159}$$

Here the stimulated emission is included. It gives rise to the last term in the bracket, which is important for small photon energies.

If we for the sake of illustration use the specific values $\omega_{ph} = 10^6\ \mathrm{s}^{-1}$, $\omega_a = 10^{16}\ \mathrm{s}^{-1}$, we have the regions shown in Fig. 7.16 for the time corresponding to $G = 25$, i.e. $t = 7\ \mu\mathrm{s}$. Note that the figure only describes situations where the emitted photon spectrum is effectively a δ function and the emission therefore can be described as an activated process. For the use of Fig. 7.16 as a reference map, please keep in mind that it involves several parameters, viz. time and two frequency factors. Maps for other values of these parameters are readily calculated from the equations.

After having mapped out the regions of the different types of processes, we now look at the effect of radiation for the unimolecular decay. This, and not the direct detection of photons, is still the most important method for detection of thermal radiation of free nanoparticles. Consider first the continuous cooling. As in the beginning of the chapter we look at the survival probability. As the rate constants are not really constant in the presence of radiation, we need to solve a generalized equation for the survival probability P:

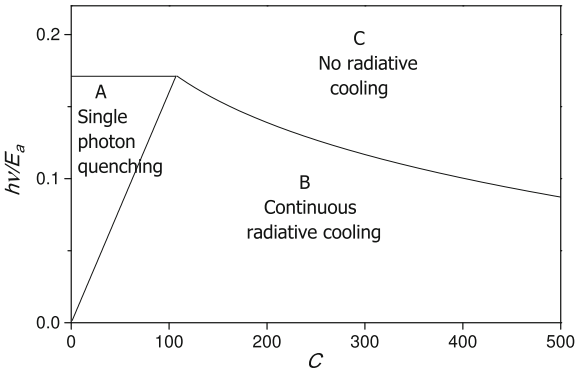


Fig. 7.16 The different regions for photon emission vs. particle heat capacity, labelled according to the predominant type of cooling. The time corresponds to $G = 25$. Regions A and C are separated by the line given by (7.156) with a ratio of frequencies of 10^{10}, region A and B by (7.155), and B, C by (7.159). These boundaries change with frequency factors and the experimental time scale

$$\frac{\mathrm{d}P}{\mathrm{d}t} = k_a(E,t)P, \tag{7.160}$$

where E is the initial energy of the particle. This integrates to

$$P(E,t) = \exp\left(-\int_0^t k_a(E,t')\mathrm{d}t'\right). \tag{7.161}$$

With the parametrization of the rate constant already introduced, this becomes

$$P(E,t) = \exp\left(-\frac{k_a(E,0)}{w}\left(1-\mathrm{e}^{-wt}\right)\right). \tag{7.162}$$

The point in the initial energy distribution where radiation quenches the unimolecular decay is then given by the relation $w \approx k_a(E,0)$. For lower energies than this dividing point, w is larger than $k_a(E,0)$ and the ratio approaches zero and P consequently unity. The survival probability vs. energy at infinite time is therefore identical to the survival probability at time $1/w$ for an otherwise identical particle that does not cool radiatively. The decay rate can then be calculated with substitution of $w^{-1}\left(1-\mathrm{e}^{-wt}\right)$ for time to give

$$R(t) = \int_0^\infty g(E)k_a(E,0)\mathrm{e}^{-wt}\exp\left(-\frac{k_a(E,0)}{w}\left(1-\mathrm{e}^{-wt}\right)\right)\mathrm{d}E. \tag{7.163}$$

With the substitution $x \equiv k_a(E,0)$ and noting, as in the derivation of (7.21), that the logarithmic derivative of this is much slower varying than the quantity itself and can be set constant in the calculation, the integral is easily performed and gives

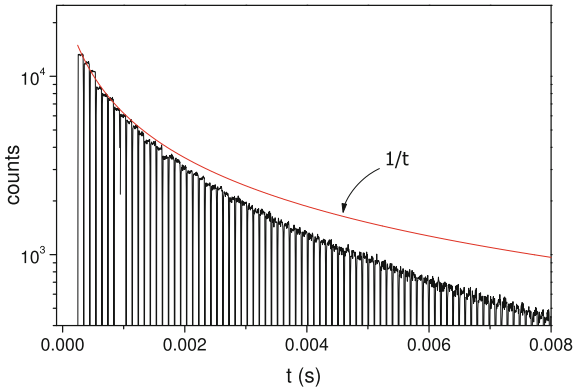


Fig. 7.17 The rate of spontaneous decay of C_{60}^- by thermal electron emission, measured in the electrostatic storage ring ELISA. The power law fitting the short time decay rate is given with the red line. The vertical lines correspond to a short part of the ring that does not contain ions and the flat pieces in between are the rates for the process $C_{60}^- \rightarrow C_{60} + e^-$. The ions are created in a plasma source called the Nielsen source

$$R(t) \propto \frac{1}{\mathrm{e}^{wt} - 1}. \tag{7.164}$$

This reduces to a power law at short times and is approximately an exponential for long times. Figure 7.17 shows the measured electron emission rates for C_{60}^- which is well described by this situation. The relatively large heat capacity makes the photon emission a continuous cooling process, even if the average photon energy is about one eV.

The survival rate given in (7.162) is of course also valid at very long times, where it then gives the probability, as a function of initial energy, that the particle survives to arbitrarily long times. When $t \gg w^{-1}$ the expression reduces to

$$P(E) = \exp\left(-\frac{k_a(E,0)}{w}\right), \tag{7.165}$$

and we see that $1/w$ takes the role of some fictitious time at which the decays are frozen, although of course the physical radiative energy loss continuous after that time. Figure 7.18 shows a comparison of two radiative cooling situations with the calculated high energy end of the 10-mer energy distribution shown in Fig. (7.5) for a cooling time of 100 μs and identical parameters. The two radiative spectra are calculated numerically with Monte Carlo simulations (see Chap. 9) of each 10^8 decay chains. The photon emission rate constants were set to the constants 1×10^5 and 2×10^5 s^{-1}, and the photon energies to 0.01 and 0.005 eV to have identical emitted powers. The spectra were sampled at the exit time of 1 s, where practically all clusters were either decays or quenched. As the figure shows, there is a very good agreement between the two types of curves. The emission rate constants used to

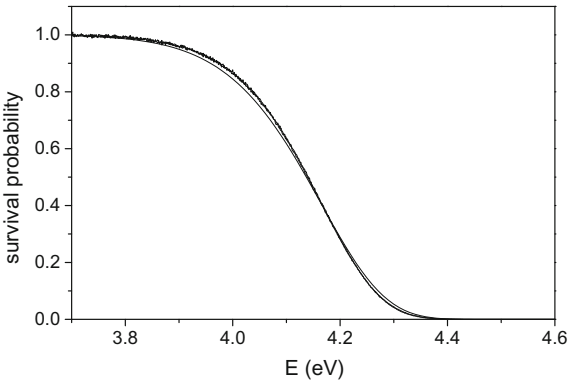


Fig. 7.18 Part of the survival probability curve for a hot 10-mer generic cluster at 100 μs (smooth line), and two radiatively cooled clusters with the same thermal properties, emitting low energy photons at a rate of 1 keV/s (slightly wiggly curves)

calculate Fig. 7.18 are unrealistic for such small photon energies that must originate from slow, vibrational transitions. If this grates the eye, all times can be rescaled a factor 100 to produce realistic IR emission rate constants with no change in the figures.

The case of high photon energies is simpler than the small photon energy case. The unimolecular decay rate is, with the initial population $P(0)$ set to one:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -k_a \mathrm{e}^{-(k_a+k_{ph})t}. \tag{7.166}$$

The amount of particles that have not undergone unimolecular decay at asymptotically long times is

$$P(E < E(t=0)) = \frac{k_{ph}}{k_{ph}+k_a}, \tag{7.167}$$

and the unimolecular decay rate is

$$R(t) = \int_0^\infty g(E) k_a \mathrm{e}^{-(k_a+k_{ph})t} \mathrm{d}E. \tag{7.168}$$

Apart from the photon emission rate constant in the argument of the exponential, this is identical to a calculation of the radiation-free decay rate. It often happens that k_{ph} varies much slower with energy than the unimolecular rate constant. In fact, as we have discussed above, this is almost unavoidable given the smaller frequency factor and the smaller activation energy, if any, for the photon emission. We can therefore to a good approximation extract the photon emission rate constant from the integral, which then gives

$$R(t) \approx \mathrm{e}^{-k_{ph}t} \int_0^\infty g(E) k_a \mathrm{e}^{-k_a t} \mathrm{d}E \propto \frac{\mathrm{e}^{-k_{ph}t}}{t}. \tag{7.169}$$

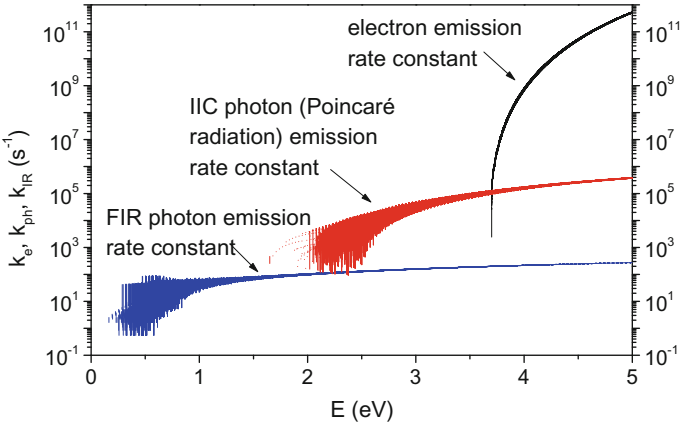


Fig. 7.19 The three decay channels that are active for C_4^-. The ion is very well characterized and all parameters relevant for the figure are known, except the electron attachment cross section to the neutral which needs to be estimated. The high energy decay channel is the thermal electron emission, the next lower is the recurrent fluorescence, i.e. the emission of a photon from a thermally excited electronic state. The low energy rate is the radiation from vibrational transitions

The decay rate is a simple product of a power law and an exponential decay caused by the quenching effect of the radiation. Notably, the time constant one extracts is equal to the photon emission rate constant. Figure 7.19 illustrates this point for the specific example of C_4^-. Figure 7.20 shows the measured quenching of the unimolecular decay measured in a storage ring after photo-excitation. The size of the system combined with the observed very high quenching rate would make it virtually certain that the quenching occurs during emission of a single photon, even if we did not know the rate constants shown in Fig. 7.19.

The question of whether one observes w or k_{ph} in a given experiment is not always so clearcut. It may need to be answered by either by theoretical arguments or by direct measurements of the photon energy. The two functional forms, for small or big photon energies, are quite similar and it is practically impossible to tell them apart from a measurement over a limited time interval. A quick and dirty test is to compare an observed time constant with expected or typical values. IR cooling times are usually above milliseconds, whereas emission times for electronic excitations are much faster, at least the relatively few that have been measured at the time of writing. The comparison is helped by the fact that w is always less than the corresponding photon emission rate constant. A sub-ms cooling rate is very likely to be electronic in nature, although the opposite can not be concluded.

Another way of stating the difference between the two types of cooling regimes, i.e. whether an experiment measures k_{ph} or w, is that for the single photon quenching regime, a exponentially decreasing unimolecular decay is caused by the exponentially decreasing *population* of the available decaying particles, somewhat similar to the standard exponential decay observed for decay out of a single state, whereas the continuous cooling is exponential because the *rate constant* decays exponentially.

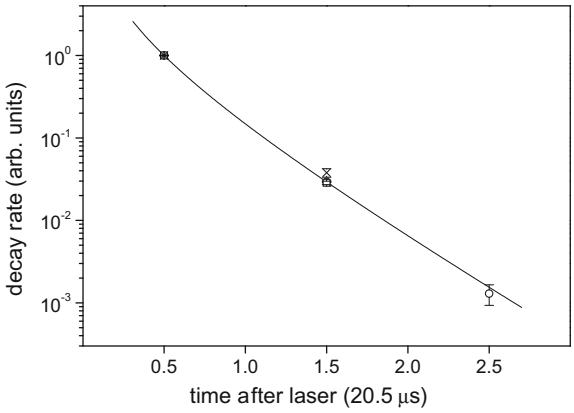


Fig. 7.20 The unimolecular decay of C_4^- after laser excitation, measured in the Tokyo Metropolitan University electrostatic storage ring, TMUe-ring. The curve is a fit with (7.169). Different symbols represent measurements for different delay between creation of the ions and injection into the ring and laser pulse, and with different photon energies. The time unit is the ion circulation time in the ring

Note that the ensemble averaging, which originally converted a perfectly normal exponential unimolecular decay rate constant to a power law decay, with the regrettable loss of the information associated with the rate constants, now turns out to be an advantage. Decays from a sufficiently narrow energy distribution will proceed exponentially, irrespective of the presence of absence of radiation, and it is impossible to reliably assign any numerical value to a radiative cooling constant from experimental data in such a case. A power law, however, modified by an exponential, signals the presence of radiation with an easily extractable radiation time constant.

In Sect. 7.7 radiation was implicitly ignored when the amount of metastable decay was calculated. But radiative cooling obviously influences this number. The effect can be calculated by integrating either of (7.164, 7.169). Here we will just be concerned with a relatively small quenching and calculate it explicitly for the continuous cooling case.

For this purpose, reconsider the non-radiative case. For this, the metastable decay between the times t_1 and t_2 was calculated as the (properly normalized) energy interval $E_{max}(t_1) - E_{max}(t_2)$ of the particle. Restricting the considerations to the low photon energy situation, we have that in the radiative case the corresponding times are $w^{-1}\left(1 - \mathrm{e}^{-wt_1}\right)$ and $w^{-1}\left(1 - \mathrm{e}^{-wt_2}\right)$. We saw in the example in Fig. 7.18 that these values give a very good representation of the effective times in the problem. With these times the metastable fraction becomes

$$\begin{aligned} f_N &\propto E_m(t_1) - E_m(t_2) \\ &= C E_a \left(\frac{1}{\ln(\omega(1 - e^{-wt_1})/w)} - \frac{1}{\ln(\omega(1 - e^{-wt_2})/w)} \right). \end{aligned} \tag{7.170}$$

The two terms in this expression can be expanded in $wt_i, i = 1, 2$. Under the condition that the effect is small, an expansion to second order is sufficient. It gives us

$$f_N \propto \frac{CE_a}{\ln(\omega/w)^2}\left(\ln(t_2/t_1) - \frac{1}{2}w(t_2 - t_1) + \frac{1}{24}w^2(t_2^2 - t_1^2)\right), \tag{7.171}$$

where higher order terms in $\ln(\omega/w)$ have been left out. This equation can be used in several ways with a reflectron Time-of-Flight setup, similar to the one used for the data in Fig. 7.14. One way is to create the ensemble with a pump laser pulse and accelerate the created ions with an adjustable delay after this laser pulse. This procedure makes the difference $t_2 - t_1$ constant and, if $t_2 \gg t_1$ which is often the case, the sum $t_2 + t_1 \approx t_2$ in the last term can be considered constant:

$$f_N \propto \ln(t_2/t_1) - \left(\frac{1}{2} - \frac{1}{24}w(t_2 + t_1)\right)w(t_2 - t_1). \tag{7.172}$$

A plot of f_N vs. the experimentally known variable $\ln(t_2/t_1)$ gives a line with a negative intercept from which one finds w.

Equation (7.172) should be used only for $w(t_2 - t_1) \lesssim 1$ to extract reliable values for w from experimental data. But although the precise value of w is not reliable determined for larger values, the fact remains that a non-zero second term on the right hand side of the equation will signal the presence of radiative cooling.

7.9 Action Spectroscopy

Spectroscopy is one of the main tools for the elucidation of the structure of particles and molecules. But nanoparticles produced in a beam are most often so dilute that absorption spectroscopy is not feasible. The solution to do spectroscopy is then to use a technique called action spectroscopy. The idea is to observe a consequence of the photon absorption, the action, which is a photon-induced, observable decay.

To convert measured yields to cross sections, it is necessary to know the quantum efficiency of the process, i.e. how probable it is that an absorbed photon causes the measurable action. When laser fluences are such that only single photon absorption occurs with any appreciable probability, a first approximation is to use the expression

$$\sigma \propto \frac{I}{\Phi F} \tag{7.173}$$

with a wavelength independent quantum yield efficiency, Φ, defined as the number of reactions per absorbed photon. I is the measured photo-induced count or count rate, and F is the photon flux or fluence, depending on whether a continuous wave (CW) or a pulsed light source is used.

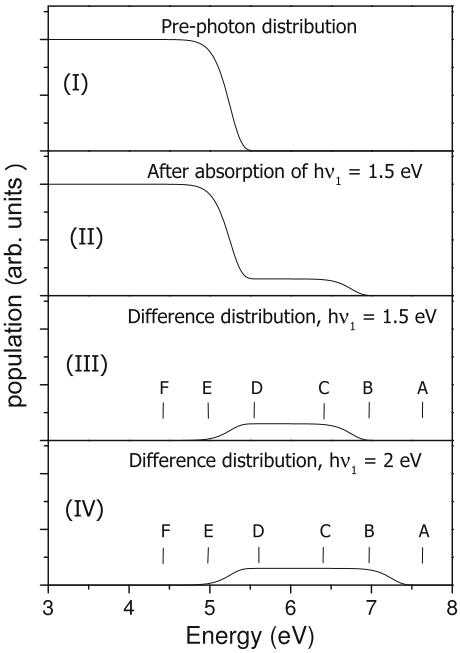


Fig. 7.21 The energy distributions immediately before (I) and after (II) absorption of a 2 eV photon by some (unspecified) system. The initial, broad distribution in frame (I), approximated by a constant value up to the energy corresponding to the decay time at the laser firing time, has some fraction shifted up by the photon energy, as shown in frame (II). The markers indicate some possible positions of measurement times-equivalent energies discussed in the main text

There are a number of situations where a constant quantum yield is a reasonable approximation, in particular for very small particles or molecules. It will, however, fail when the action is a thermally induced process of an even moderately sized particle or when photon energies are small. These are the cases considered here. The frames (I–IV) in Fig. 7.21 illustrates the situation.

The enhanced decay caused by photon absorption is the decay of the shifted part of the energy distribution within the temporal measurement limits set by the experimental device. Let the photon be absorbed at time t_{las} and the measurement of the induced action take place in the time interval between t_1 and t_2, where $t_{las} \leq t_1 < t_2$, during which some constant fraction of the decays is recorded. We can express the total measured yield as the integral of the action yield specified wrt. excitation energy, photon energy and the two times t_1 and t_2 as

$$\Phi(h\nu, t_1, t_2) = \int_0^\infty \phi(E, h\nu, t_1, t_2)\rho(E - h\nu)\mathrm{d}E, \tag{7.174}$$

where $\rho(E - h\nu)$ is the energy distribution before photon absorption. Assuming this is flat up to the inevitable cross-over to zero caused by the factor $\exp(-kt_{las})$, we can set it to unity and write the distribution the distribution right after photon absorption as

$$P(t_{las}, E) = \mathrm{e}^{-k(E)t_{las}}\,(1 - \sigma F) + \sigma F \mathrm{e}^{-k(E-h\nu)t_{las}}. \tag{7.175}$$

Adding a time dependence, $\mathrm{e}^{-k(E)(t-t_{las})}$ to this and subtracting the distribution corresponding to no laser pulse gives the photo-induced signal:

$$P_p(t, E) = \sigma F \left(\mathrm{e}^{-k(E-h\nu)t_{las}-k(E)(t-t_{las})} - \mathrm{e}^{-k(E)t}\right). \tag{7.176}$$

The total action signal is found as the difference between these populations at t_1 and t_2, integrated over all energies:

$$\begin{aligned} I \;\propto & \int_0^\infty \left(P_p(E, t_1) - P_p(E, t_2)\right) \mathrm{d}E = \\ & \sigma F \int_0^\infty \left[\mathrm{e}^{-k(E-h\nu)t_{las}-k(E)(t_1-t_{las})} - \mathrm{e}^{-k(E)t_1}\right] \mathrm{d}E - \\ & \sigma F \int_0^\infty \left[\mathrm{e}^{-k(E-h\nu)t_{las}-k(E)(t_2-t_{las})} + \mathrm{e}^{-k(E)t_2}\right] \mathrm{d}E. \end{aligned} \tag{7.177}$$

The rate constant $k(E - h\nu)$ is usually much smaller than $k(E)$, and we can capture this by defining a back-shifted time t_0 as

$$k(E - h\nu)t_{las} \equiv k(E)t_0. \tag{7.178}$$

With this, the signal becomes

$$I \propto \sigma F \int_0^\infty \left[\mathrm{e}^{-k(E)(t_1-t_{las}-t_0)} - \mathrm{e}^{-k(E)(t_2-t_{las}-t_0)} + \mathrm{e}^{-k(E)t_2} - \mathrm{e}^{-k(E)t_1}\right] \mathrm{d}E. \tag{7.179}$$

Each exponential term in the integrand has the same form as we have seen before, and can be calculated with the same approximation. We can even use the fictitious time involving w if radiation is present. Any offset in energy in the caloric curve cancels, and the expression for the yield becomes

$$\begin{aligned} I \;\propto\; \sigma F \Big[& \frac{1}{\ln\left(\omega\left(1 - \mathrm{e}^{-(t_1-t_{las}+t_0)w}\right)/w\right)} - \frac{1}{\ln\left(\omega\left(1 - \mathrm{e}^{-(t_2-t_{las}+t_0)w}\right)/w\right)} \\ - & \frac{1}{\ln\left(\omega\left(1 - \mathrm{e}^{-t_1 w}\right)/w\right)} + \frac{1}{\ln\left(\omega\left(1 - \mathrm{e}^{-t_2 w}\right)/w\right)} \Big]. \end{aligned} \tag{7.180}$$

There are four times in this problem, giving rise to a lot of different situations for the different combinations of relative values of these. A discussion of all these cases is basically the job the equations are supposed to do, and we only need to look at a few cases to get a feeling for the results.

As before we note that each of the factors in (7.180) represents a cutoff energy. The main features of the results will then be well represented by considering the amount of the difference spectrum in frames (III) or (IV) in Fig. 7.21 that resides between the two cutoff energies representing t_1 and t_2. We can therefore identify the two times with the energies on the abscissa. The difference spectrum representing

the excess energy distribution caused by the photon absorption is shown in frames (III–IV) of Fig. 7.21 for the two photon energies 1.5 eV and 2.0 eV.

For the times corresponding to energies A (t_1) and B (t_2), the yield is exponentially suppressed with energy. If observable at all, it will therefore have an exponential dependence on the photon energy. Changing the photon energy shifts the onset of the exponential to a different energy, as seen when comparing frame (III) and (IV). The quantum yield therefore varies exponentially with photon energy.

If instead the times are represented by A and C, one will observe a signal which varies almost linearly with photon energy, apart from a negative offset. Consequently the quantum yield is also photon energy dependent in this case.

If the times are then both located within the enhanced region, as C and D, the quantum yield is independent of photon energy. In this region one can therefore identify yields with cross sections, up to a constant.

This is also the case when t_1 corresponds to C or D, and t_2 to either of E or F.

Finally, when the times refer to the energies at E and F, the signal is again exponentially suppressed.

Both the lower and the upper cutoff of the difference distribution in frames (III) and (IV) depend on the time the probe laser is fired. For fixed values of t_1 and t_2, a variation of the laser firing time may therefore shift the relative position of the corresponding two energies and move the quantum yield into another regime with the concomitant changes in the quantitative interpretation of the yields.

Exercises

7.1 Derive an approximate $1/t$ decay with (7.3) and a rate constant that depends on energy as $k \propto E^s$, where $s \gg 1$. Next, pretend that the highest rate constant for some unidentified and unphysical reason varies not as $1/t$ but as $1/t^2$, and derive the ensemble average $1/t$ decay again.

7.2 Pretend Fig. 7.5 represents experimentally measured curves. Extract all the parameters you can from the data. Those can include C_N, G, E_a, and the offset E_0 in the caloric curve $E = C_N T + E_0$. Assume $C_N = C_{N+1}$.

7.3 The validity of (7.60) is limited by the requirement that the abundances must be positive. Calculate what this limit translates into for the relative difference in activation energy of sizes N and $N + 1$.

7.4 Compare the reduction in the widths of the distributions in Fig. 7.5 relative to the input 3 eV dissociation energy to the values expected from (7.58).

7.5 Calculate the fluctuation of the energy content of a particle with constant heat capacity C_N in the canonical ensemble. Compare it with the evaporative ensemble result for particles with large heat capacities.

7.6 Verify the estimate that says that on time scales of nano- to microseconds, a change in temperature of 10% corresponds to an order of magnitude in decay time.

7.7 Calculate (7.19) with a saddle point expansion. You can set $g(E)$ to a constant.

7.8 Show that (7.136) is the relation between activation energy and temperature in the leading order finite heat bath approximation, as claimed in the text.

7.9 Use (7.19) to show that for an arbitrary distribution of decay constants, the decay rate $R(t)$ is always a decreasing function of time and the curvature of $R(t)$ is always positive. Show that the same holds for $\ln(R)$.

7.10 The energy distributions in Fig. 7.8 gets a contribution from the source, which sublimes the molecules with a temperature of 700–800 K. This distribution contributes to the width and slopes of the curves in the figure, because the coincidence technique used to measure them does not include measurements of the initial, pre-collision energy. Assume for simplicity that this distribution, as well as the fragmentation-induced, is a Gaussian. Calculate the correction of the source energy distribution to the width and the slopes of the experimentally observed curves.

7.11 Figure 7.10 shows the time of occurrence of the last decay in decay chains of a few different lengths, for the special case where all rate constants in the chain are identical. The data were calculated numerically and illustrate the approach to a normal distribution for this special situation.

It is possible to derive expressions for the curves with (7.108). With convenient time units, the probability for decay at time t for a chain of length n at time t is given by

$$P(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t} \prod_{j=0}^{n} \frac{1}{1+\mathrm{i}k}\mathrm{d}k. \tag{7.181}$$

The integral can be calculated by the methods of residues, as the corresponding integral in (7.108). The result is

$$P(t) = \frac{1}{(n-1)!}t^{n-1}\mathrm{e}^{-t}. \tag{7.182}$$

Approximate this with a saddle point expansion and compare it with the similar approximation of (7.181). Calculate the third moment of the exact distribution and give a criterion for applicability of the saddle point expansion in this special case, as a function of the distance to the mean value.

7.12 Use the macroscopic description of the thermal conductance,

$$P_t = -A\kappa\overline{\nabla}T \tag{7.183}$$

in combination with the emitted power from the surface

$$P_e = -kD, \tag{7.184}$$

to find the order of magnitude of the temperature difference between the interior and the surface of a freely evaporating particle. P is the energy flow through the surface A, κ is the material dependent thermal conductivity, $\overline{\nabla} T$ the temperature gradient, k the decay rate and D the dissociation energy. Assuming a stationary state, i.e. that the temperature drop in the surface due to the particle emission is compensated by the heat conducted from the interior, find a estimate for $\overline{\nabla} T$. With the requirement that the temperature gradient multiplied by the radius of the cluster is less than T/G, we want to find the largest size where the finite conductivity can be ignored. Show first that temperature differences occur if

$$\frac{C_v}{r} \gtrsim 4\pi\kappa t. \tag{7.185}$$

Use a value of $\kappa = 100$ W/Km and an r of 1 μm to find a first estimate of the cross-over heat capacity. Next assume a heat capacity of $3N$ and express a more accurate value of the critical size as

$$N \gtrsim (36\pi)^{2/3} \kappa \left(\frac{m}{\rho}\right) t, \tag{7.186}$$

where m is the mass of the monomer and ρ the mass density. Plug in numbers. Finally, discuss the criterion $\overline{\nabla} T \lesssim T/G$.

7.13 Radiative cooling will cause changes in the interpretation of the abundance spectra when one or more of the radiative time constants is shorter than the observation time. Use the expression for the effective decay time in (7.162) to re-derive the small particle limit abundances in terms of dissociation energies, i.e. the radiative equivalent of (7.61). Assume different cluster sizes can have different radiative time constants. Estimate the magnitude of the effect when these constants differ by factors of ten and times are standard Time-of-Flight mass spectrometric values.

7.14 For a particle that decays both radiatively and unimolecularly, consider the large photon energy case with a δ function initial energy distribution. Calculate, as a function of time, the number of particles that have not decayed at all, that have decayed by photon emission, and that have decayed unimolecularly. Assume that the unimolecular decay constant changes by a factor α on photon emission.

7.15 Calculate the order of magnitude of the photon energies for which stimulated emission should be included in (7.159).

Chapter 8
Abundance Distributions; Large Scale Features

The question of size is of overwhelming importance for small particles. To define the question we need to state precisely what we mean by the word size. It may mean some geometric size, which is relevant for obvious reasons, and which has given the name to the field of nanoscience (science where the studied objects are less than a micrometer in at least one dimension); it may mean the number of monomers in a particle, and it may mean the effective number of degrees of freedom.

We will consider the last aspect in this chapter, and disregard a number of effects that are much more important in other connections, such as quantum size effects where properties vary with the addition of a single monomer or a single electron to the particle. Size distributions will be calculated based on only the coarsest features in the size dependence of the binding energies, those that survive in the large particle size limit and that can be measured on bulk samples; the bulk cohesive energy, the surface tension and the electrostatic charging energy.

8.1 Liquid Drop Energies

When concerned with abundances of particles created in thermal processes, particle stabilities are of overwhelming importance. In Chap. 7 we studied how losses shaped the size-to-size variations of abundance spectra. Here we will look at abundances when also growth is important and usually dominant.

We will apply the schematic but usually reasonably accurate liquid drop model for the binding energies in the particles. The energies parametrized by this model are the ground state energies, unless otherwise stated. For charged particles the model is also known as the Thomson liquid drop model, and in nuclear physics as Weizsäcker's mass formula, when the terms specific to nuclear physics are added.

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_8

Although the name of the model contains the word 'liquid', its applicability is not limited to liquids. The term refers mainly to the spherical shape of the particle. It will not be the best description of binding energies of nanocrystals, i.e. crystalline nanoparticles that have their lattice faces defining the surface of the particles. But the liquid drop concept can be useful even for those when the proportions of the nanocrystals are size independent, as they are expected to be.

The most important energy term in the liquid drop model is proportional to the number of particles and to the bulk binding (or cohesive) energy per atom, or more generally per monomer. A monomer is defined here as the main component of the saturated vapor of the material, and can be an atom, a diatomic molecule or whatever. Usually a single of these components will dominate the composition of the vapor, although mixed composition vapors are not that rare, a case being sulphur with several molecular species in its vapor. Acetic acid vapor for example, is composed of dimers of the molecules, and dimers are then the monomers in the sense used in this chapter, irrespective of whether the bulk can be described as composed of such dimers or not. For simplicity we will use the word atom for the vapor monomers, even for molecular vapors.

If the zero of energy is chosen as the infinitely separated monomers, the leading order term of the ground state energy of an N atom particle as $E_N = -AN$, where A is the bulk binding energy per atom. It could be argued that one should use the substitution $N \to N-1$ because the energy of the monomer is set to zero (a monomer does not bind to itself) but we will not bother with this.

This simple proportionality with bulk properties obviously ignores the effects caused by the surface. In addition to the bulk binding energy, particles also have a surface tension. The surface tension increases the energy, and consequently reduces the binding energy, with an amount proportional to the surface area. If drops of the same material and different sizes have identical densities, the volumes of the particles are proportional to N, the radii proportional to $N^{1/3}$, and the surface areas therefore proportional to $N^{2/3}$. The ground state energy is then

$$E_N = -AN + BN^{2/3}. \tag{8.1}$$

The value of B is related to the macroscopically measurably surface tension, γ, by equating the macroscopic surface energy calculated with γ to the microscopic, expressed with B:

$$BN^{2/3} = 4\pi r_N^2 \gamma = 4\pi r_1^2 N^{2/3} \gamma \Rightarrow B = 4\pi r_1^2 \gamma. \tag{8.2}$$

The conversion requires a value for r_1. We can find a first and often final estimate for this by equating the bulk reciprocal number density, ρ, to the volume $4\pi r_1^3/3$,

$$\frac{1}{\rho} = \frac{4\pi}{3} r_1^3. \tag{8.3}$$

Surfaces of nanocrystals, in the sense of structures that are small pieces of regular lattices, do not conform exactly to this simple picture. They are better described by the so-called Wulff construction, by which the surface of a particle is constructed from lattice planes with sizes that are determined by the individual surface tensions of these planes. The procedure gives rise to systematic deviations from the spherical shape, beyond those imposed by the lattice structure itself. For use in (8.1) and the later, improved version, an effective surface tension for those particles can be defined as the properly weighted value for the contributing surfaces. Such a value will not, however, be directly related to the energy associated with the deformation of the particle from its ground state configuration, as the B used in (8.1) is expected to be.

Values of B are strongly correlated with A. The values of B based on macroscopically measured surface tensions are shown in Fig. 8.1 vs. the enthalpies of evaporation for several metallic elements and the rare gases. The surface tension depends on temperature and most data points in the figure refer to the surface tension at the melting temperature. Both the surface tension and the enthalpy of vaporization therefore differ from the ground state properties but they usually give good estimates of the ground state values of both A and B. From the figure it is clear that there is a strong correlation between the values of A and B, approximately as

$$B = {}^{2}\!/_{3}\, A. \tag{8.4}$$

This type of relation is understandable for systems with short range, two-body interactions. The origin of the 2/3-rule in (8.4) can then be derived, semi-quantitatively at least, by considering matter as a packing of atoms and cohesive energy as a measure of the number of nearest neighbor atoms. A bulk atom, i.e. an atom on the inside of the material, will be surrounded by a number of other atoms, and with short range interatomic potentials we can take the binding energy of an atom to be proportional to the number of these neighbors. For a surface atom this number is reduced relative to bulk atoms, and to find the reduction in binding energy, we simply need to calculate how many there are of these neighbors for an atom in the surface vs. in the bulk.

Let's simplify the problem and consider the atoms small cubes instead of the usual spheres. For a surface atom, one of the six surfaces does not touch another atom and the binding of a surface atom is therefore reduced by $A/6$. To find the total increase in energy, which is equal to the reduction in binding energy, we must multiply this energy with the number of surface atoms. This number is equal to the surface area divided by the area of a single atom,

$$N_s = \frac{4\pi r_N^2}{\pi r_1^2} = \frac{4\pi \left(r_1 N^{1/3}\right)^2}{\pi r_1^2} = 4N^{2/3}. \tag{8.5}$$

This gives a total surface energy of

$$E_{surf} = 4N^{2/3}\frac{A}{6} = \frac{2}{3}N^{2/3}A, \tag{8.6}$$

in agreement with the suggestion from the experimental data.

With all the reservations one may have to this calculation, the simple argument nevertheless shows that it is reasonable that B is proportional to A and that the constant of proportionality is on the order of unity. Surprisingly, this simple relation also holds reasonably well for metallic systems, where the whole concept of two-body atom-atom interactions is deeply suspect. Also the liquids ^{3}He and ^{4}He (the two lowest points in Fig. 8.1) follow the relation reasonable well, albeit not perfectly.

At this point we therefore feel confident that we have a expression for the total ground state liquid drop energy, or cohesive energy of the form

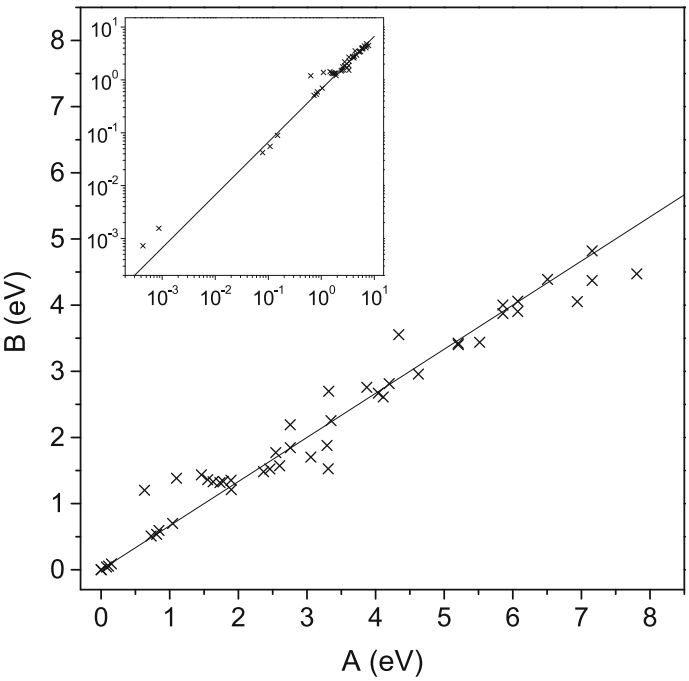


Fig. 8.1 Values of the B parameter, calculated from experimentally measured surface tension with (8.2), vs. the experimental bulk enthalpy of vaporization per atom at the boiling point at 1 atmosphere, summarily set equal to A. Values of A have been extrapolated to the boiling point for some elements. The surface tension depends on temperature and the melting temperature values have been used. Only elements where the vapor consists of monomers are included. The ^{3}He has a significant temperature dependence of the enthalpy of vaporization and the highest value was used. The inset shows the same data on a double-logarithmic scale. The lines are the relation $B = 2/3\,A$ in both cases

$$E_N = -AN + BN^{2/3}, \tag{8.7}$$

and we have related A and B to each other and argued that they are also related to macroscopic and measurable quantities.

Writing the ground state energies of small particles in terms of a few macroscopic parameters means extrapolating by something like 20 orders of magnitude. This may seem rather bold (foolhardy, nonsensical, whatever you prefer). It is actually not such a poor idea if judged by the few and limited size range cases where it has been tested, although whether or not it is sufficiently accurate depends critically on the applications. In nucleation contexts, for example, even small errors in surface tensions give rise to exponentially large errors (see Sect. 8.10).

Experimentally determined absolute binding energies are not easy to come by for an extended size range. As an alternative we compare the liquid drop parametrization with a few sets of theoretical data. Rewriting (8.7) as

$$\frac{E_N}{N} = -A + BN^{-1/3}, \tag{8.8}$$

provides a sensitive test of that equation when the left hand side is plotted vs. $N^{-1/3}$. Figure 8.2 shows the results for sodium chloride clusters from a theoretical calculation. A fit with a simple straight line gives a reasonable value for the ratio of surface tension to dissociation energy. The seemingly erratic size-to-size variations are due to the special stability or instability of certain sizes. These are not included in the liquid drop parametrization, by definition. Such types of deviations from liquid drop energies are frequently seen for small particles and are often due to shells structure, either of the geometric or of the electronic type, but can also, in particular for the smallest particles, have no obvious or easily rationalized systematics associated with them.

Even disregarding the shell structure and other terms varying around zero, the liquid drop parametrization has its limitations. One case where this is seen clearly is the theoretically calculated ground state energies of the toy element with atom-atom interactions given by the Lennard-Jones potential. This is a two-body interaction which mimics the interaction of rare gas atoms and which is discussed in more detail in Chap. 12. The ground states and their energies of a long range of sizes of clusters of atoms interacting with this potential, colloquially known as Lennard-Jonesium, have been found by minimizing total energies, with no small expenditure of time and manpower.

The energies per particle for Lennard-Jonesium is shown in Fig. 8.3. From the deviations from a straight line it is clear that the two terms in (8.7) are not sufficient to account quantitatively for the energies. A third term proportional to the radius of the particle, or $N^{1/3}$ with a negative coefficient is required to provide a good fit, as seen from the curvature in the data. If one insists on a liquid drop limited to the first two terms, this would correspond to a surface tension that decreases as the size decreases.

Figure 8.3 also shows that the surface term has a coefficient which is *bigger* than the bulk cohesive energy (the large cluster part of the curve extrapolates to

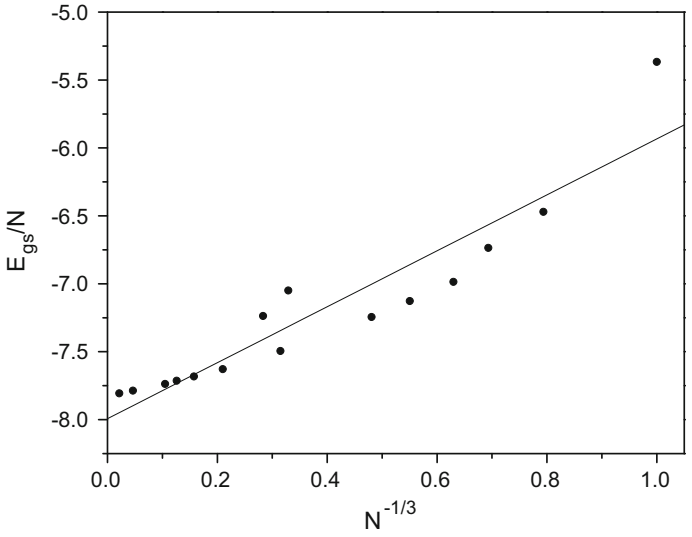


Fig. 8.2 The cohesive energy per monomer molecule for NaCl clusters. The straight line fit gives a value close to the predicted 2/3 of the monomer binding dissociation energy. The data are theoretical values from 'Clusters of alkali halide molecules', D.O. Welch, O.W. Lazareth, G.J. Dienes, and R.D. Hatcher, *J. Chem. Phys.* **68** (1978) 2159

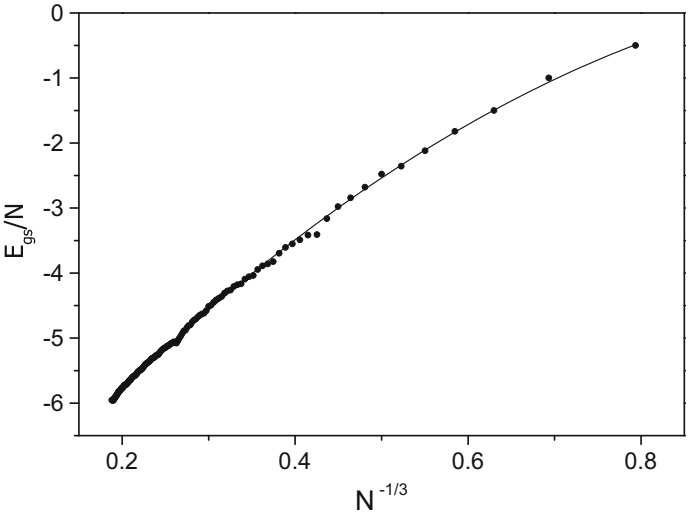


Fig. 8.3 The left hand side of (8.8) vs. $N^{-1/3}$ for the ground state energies of clusters of atoms interacting with the Lennard-Jones potential. The small deviations from the smooth curve are finite size effects caused by the variations of the packing efficiency of the atoms into the cluster, known as packing shell structure. The data are from the Cambridge Energy Landscape Database at www-wales.ch.cam.ac.uk/CCD.html

zero energy before reaching $N^{-1/3} = 1$), in a very un-pedagogical contradiction to our estimate above. One reason this can happen is that the structure of the clusters plotted is not that of the bulk and extrapolation to infinite size therefore fails, and this is the situation here. The structures have a five-fold symmetry, icosahedral to be specific, for certain sizes where the atoms pack into closed structures, which is an example of packing or geometric shell structure. For the icosahedral structures, shell closings appear for $N = \frac{10}{3}K - 5K^2 + \frac{11}{3}K - 1$, with K the (positive integer) shell number. The first four of these are $N = 1, 13, 55, 147, \ldots$. At other sizes, ground state structures of the clusters are often one of these closed shell structures decorated with an incomplete next shell.

The icosahedral structure is not compatible with bulk translational symmetry, and the particles are therefore not small pieces of bulk material. With growing particle sizes tension builds up on the inside, eventually reaching a limit where the crystalline structure becomes the most stable. This means that for this type of clusters, total energies get non-linear contributions from regions other than the surface. Then it is no surprise that the surface tension appears larger than our 2/3 rule predicts.

We could continue the expansion of the binding energy to higher orders of $N^{-1/3}$ by adding more terms to the series, as suggested by the fit of the Lennard-Jones clusters, with a procedure known as the leptodermous expansion ('thin skin' expansion, obviously a definite misnomer in the case of the bulk atom effects of Lennard-Jonesium). These higher order terms are rarely known because there are no macroscopic measurements of the coefficients, and for most purposes the first two terms are sufficient to account for the smooth size dependence of the binding energy.

An exception to this, both with respect to our knowledge and to the magnitude, is the charging energy. Addition of a charge, positive or negative, to a neutral particle causes the energy to change with an amount which will include a term proportional to $N^{-1/3}$, i.e. a size dependence which is a factor $N^{4/3}$ smaller than the bulk contribution. The coefficient of this contribution can nevertheless be large enough to justify the inclusion of the term.

The energy depends on the location of the charge. We will consider two charge distributions. In the first the charge, carried by an ion or an electron, is located in the center of the spherical particle. For this charge distribution, the classical electrostatic energy is calculated to

$$E_c = \frac{1}{2}\frac{q^2}{4\pi\varepsilon_0}\left(\frac{1}{Kr'} + \frac{1}{r_N}\left(1 - \frac{1}{K}\right)\right), \tag{8.9}$$

where K is the relative dielectric constant, equal to the ratio $\varepsilon/\varepsilon_0$, and r' is the radius of the central ion. The concept of the radius of an ion in a medium have limited applicability, and it is better to extract the term with r' and write it as an effective charge solvation energy, $E_{c,0}$:

$$E_c = E_{c,0} + \frac{1}{2}\frac{q^2}{4\pi\varepsilon_0}\left(\frac{1}{r_N}\left(1 - \frac{1}{K}\right)\right). \tag{8.10}$$

This has the added advantage that other effects can be collected into the empirical constant $E_{c,0}$. Depending on the zero of energy one chooses, these effects can include the energy gained or spent in creating the core charge. For the zero of energy used here, $E_{c,0}$ is the energy needed to insert the preformed charge into the center of the particle while, hypothetically, keeping the polarization energy in the second term zero. In general, we must expect that the effects collected into $E_{c,0}$ will have some size dependence unrelated to the polarization energies, if not for anything else, then because of the increased and size dependent surface area created upon the insertion. Equation (8.10) can be interpreted as the leading order term in an expansion in $1/r_N$ of the smooth part of this size dependence. We will include the term in the form of the bulk work function, W, and let the surface-like contribution be absorbed into the existing surface term.

The factor in front of the parenthesis in (8.9) is half the standard electrostatic constant. It takes the value 7.2 eVÅ, which justifies the inclusion of the effect in the liquid drop expression. The factor of one half appears as the charging energy for $K \to \infty$, corresponding to a metallic sphere. This gives us directly the electrostatic energy for the other charge distribution we will consider, which is the one where the charge is homogeneously distributed over the surface of the particle. The energy is

$$E_c = E_{c,0} + \frac{1}{2}\frac{q^2}{4\pi\varepsilon_0 r_N}. \tag{8.11}$$

$E_{c,0}$ can be interpreted as a surface attachment energy of quantum mechanical origin. Both this energy and the corresponding one in (8.10) will in general depend on q, and the equations are best used to represent the variations with N and not with q.

These two situations do not cover all possibilities. A centrally located charge, in particular, is not a good description for multiply charged particles if the charge is located on more than one atom or monomer. The electrostatic repulsion between these mobile charges will guarantee that not both can be located in the center of the particle, which in most cases is also not possible due to the fact that two atoms can usually not be at the same place. It may still be energetically favourable to have charges embedded into the interior, though. Then the problem includes some non-trivial geometrical optimization, even for the simple liquid drop description. But if care is exercised in the interpretation of the constants of the charging energy, one may still use the above reasoning as a convenient numerical template.

We then finally have the expression for the liquid drop (ground state) energy:

$$E_N = -AN + BN^{2/3} + CN^{-1/3} + qW, \tag{8.12}$$

where all numbers are related to macroscopic quantities according to the recipes given above.

If this were all, life would be dull. On top of these contributions, which vary smoothly with N, shell structure will generate terms that vary with N more or less systematically around zero, as we have already seen in plots of theoretical data. For metallic particles, an important contribution of this nature is the electronic shell

energy, $E_{shell,N}$, that represents the stability variations due to the electronic shell structure. For cold clusters, metallic or not, one may also observe a geometric shell structure due to the packing of the atoms, which has already been mentioned above. Another non-smooth term is the odd-even effect of metallic cluster, $E_{oe,N}$, which is similar to shell energies in the sense that it oscillates around zero, albeit with a period of only two. We will meet these in Chap. 10 on valence electrons. The full expression for the liquid drop ground state energy of particles that show these variations is then

$$E_N = -AN + BN^{2/3} + CN^{-1/3} + qW + E_{shell,N} + E_{oe,N}, \tag{8.13}$$

where any of the last three terms may be zero for a specific system. Common for the shell energy and the odd-even energy is that their contribution to the total energy is zero on the average, and for this reason we will disregard them in the rest of this chapter.

8.2 The Partition Functions

The first goal of this chapter is to find abundances in thermal and chemical equilibrium. This is done with the canonical partition functions because these provide the free energies and chemical potentials we need to calculate particle numbers (see Chap. 1). To calculate these partition functions, some schematic features must be introduced into the description of the excitation spectra of the particles, just as the picture of the ground state energies was painted with a broad brush. Rigorously speaking, all degrees of freedom except the translational motion couple directly or indirectly. We will ignore these couplings and, as in Chap. 5, write the total partition function of a single particle as the product of the translational, vibrational, electronic and rotational partition functions:

$$Z_{tot}(N) \approx Z_{tr} Z_{rot} Z_{vib} Z_{el} \mathrm{e}^{-\beta E_N}, \tag{8.14}$$

where E_N is the ground state energy. The choice of the zero of the energy as that of the completely separated monomers is particularly convenient here, as will be clear soon.

The translational partition function is easy to calculate because we already have the level density of a free particle (see (2.68)). The mass of the monomer is m, and then

$$\rho_{tr,N}(E) = \frac{1}{h^3} V 4\sqrt{2}\pi (mN)^{3/2} E^{1/2}, \tag{8.15}$$

with the canonical partition function:

$$Z_{tr}(N) = \int_0^\infty \rho_{tr,N}(E) \mathrm{e}^{-\beta E} \mathrm{d}E = V\sqrt{8\pi^3} \left(\frac{mNT}{h^2}\right)^{3/2}. \tag{8.16}$$

Alternatively, we could have gone back to basics and calculated it as

$$Z_{tr}(N) = \frac{1}{h^3}\int_V \mathrm{d}^3x \int_{\overline{p}} \mathrm{e}^{-\beta\frac{\overline{p}^2}{2m}}\mathrm{d}^3p. \tag{8.17}$$

The vibrational partition function depends on whether the particle is linear or not. We will assume that it is not. Then there are $3N-6$ vibrational degrees of freedom and, ignoring anharmonicities, we have the by now well-known vibrational partition function

$$Z_{vib}(N) = \prod_{i=1}^{3N-6}\left(1-\mathrm{e}^{-\beta\hbar\omega_i}\right)^{-1} \approx \left(\frac{T}{T_{vib}}\right)^{3N-6}, \tag{8.18}$$

where the high temperature limit was used in the last equality sign. The quantity T_{vib} is defined as the geometric average of all the vibrational frequencies of the particle, expressed in temperature units:

$$T_{vib} = \left(\prod_{i=1}^{3N-6}\hbar\omega_i\right)^{1/(3N-6)}. \tag{8.19}$$

It should be pointed out that this partition function is only a first approximation. Anharmonicities will increase the value and if melting occurs below T, it will contribute a significant factor to the right hand side of (8.18).

To calculate the rotational partition function of a particle we need to know the shape of it. For rotational motion, the shape is summarized in the moment of inertia tensor, defined as

$$I_{ab} \equiv \sum_i m_i\left((x_{i,1}^2+x_{i,1}^2+x_{i,1}^2)\delta_{a,b}-x_{i,a}x_{i,b}\right), \tag{8.20}$$

where the subscripts a, b can be any combination of x, y, z, the index i runs over all masses and $x_{i,a}$ is the a-coordinate of atom i when the center of mass is taken as origo of the coordinate system. This tensor determines the rotational motion of the particle. It can be brought on diagonal form by a coordinate transformation, equivalent to a diagonalization of a 3×3 matrix, and these diagonal elements are the ones we need. The particles we will deal with here will usually be spherical or close to spherical, for which shape the moments of inertia around the three principal axis are equal, $I_{xx}=I_{yy}=I_{zz}\equiv I$, making the particles what the spectroscopists call a spherical top. The moment of inertia for a sphere with constant density is already diagonal and the diagonal elements are equal to $I=2MR^2/5$. With the radius $R=r_1N^{1/3}$ and total mass $M=Nm$ this gives the N dependence

$$I = {}^2\!/_5\, mr_1^2N^{5/3}. \tag{8.21}$$

For a spherical top, the rotational energies are

$$E_{rot} = \frac{\hbar^2}{2I} L(L+1), \tag{8.22}$$

and the degeneracy of the levels are $(2L+1)^2$, where L is the orbital angular momentum of the motion of the nuclei. The partition function is then given by

$$Z_{rot}(N) = \sum_{L=0}^{\infty} (2L+1)^2 \mathrm{e}^{-\beta\hbar^2 L(L+1)/2I}. \tag{8.23}$$

The partition function also depends on the symmetry of the particle in a more subtle way than just the pure geometric symmetry reflected in the moment of inertia tensor. A spherical top is symmetric in the sense of moments of inertia, but it is a different question whether or not the wave function representing the particle also has spherical symmetry. If is has, the particle cannot rotate and the partition function of such a system reduces to unity. But a completely rotationally symmetric particle would have eliminated the granular structure of matter, which is not an easy thing to do.

Still, even after accounting for the presence of atoms at relatively well-defined position, there may be a residual, discrete symmetry left in the wave function. Such a symmetry of the wave function under rotations is usually summarized in the symmetry number, which is an integer that takes into account the symmetry of the entire wave function, both the electronic wave function, and the positions and vibrational motion of the nuclei. It represents the number of identical states of the entire system that can be reached by rotations alone. For high angular momentum states, where $L \gg \hbar$, one divides the unrestricted sum by the symmetry number to get the correct partition function. We will assume the symmetry number is unity, which is the classical value, because large particles at excited states rarely have a total wave function with enough symmetry to cause major quantum corrections. Putting it shortly, there is a correction factor and it is equal to one!

When $T \gg \hbar^2/2I$, the sum over L in (8.23) can be approximated with an integral, corresponding to the first term in the Euler-MacLaurin formula. If also the approximation $L(L+1) \approx (L+1/2)^2$ is used, we have

$$Z_{rot}(N) = \sqrt{\pi} \frac{(2IT)^{3/2}}{\hbar^3}. \tag{8.24}$$

In analogy to the vibrational partition function, the constant in this expression is often expressed as a temperature:

$$T_{rot} \equiv \frac{\hbar^2}{2I}. \tag{8.25}$$

We see from (8.25) that the larger the particle, or more precisely the larger the value of the moment of inertia I, the lower the rotational temperature. Spectroscopists use the rotational constant which is defined as

$$B \equiv T_{rot}\frac{k_B}{hc} = \frac{h}{8\pi^2 Ic}, \tag{8.26}$$

where c is the speed of light. The unit of this is inverse length, and usually cm^{-1} (wavenumber, see Appendix B for the value) is used. In this unit, energy is given by the number of wavelengths per cm for photons with that energy. Some values are given in Table 5.1 in this unit. If we stick with the characteristic rotational temperature we have

$$Z_{rot}(N) = \sqrt{\pi}\left(\frac{T}{T_{rot}}\right)^{3/2}. \tag{8.27}$$

When the moments of inertia along the principal axes are are not identical, the exact quantum mechanical rotational energy levels are more difficult to find. We will not attempt that here, but instead calculate the partition function semiclassically. From Chap. 2 the canonical partition function is semiclassically equal to

$$Z_{rot}(N) = \frac{1}{h^3}\int \mathrm{e}^{-\beta(L_x^2/2I_{xx}+L_y^2/2I_{yy}+L_z^2/2I_{zz})}\mathrm{d}^3\theta \int \mathrm{d}L_x\mathrm{d}L_y\mathrm{d}L_z, \tag{8.28}$$

where the three angles θ refer to the orientation of the particle. We have inserted the rotational energy expressed in terms of the classical angular momenta components $L_x,\ L_y,\ L_z$:

$$E_{rot} = \frac{L_x^2}{2I_{xx}} + \frac{L_y^2}{2I_{yy}} + \frac{L_z^2}{2I_{zz}}. \tag{8.29}$$

The angular integration takes some care. It goes as follows: Mark two antipodal positions on the particle. This defines an axis. The integral is then performed over all 4π of the orientation of this axis, and then independently over the 2π that defines the angular position around the axis. This gives $8\pi^2$. We then have

$$Z_{rot}(N) = \frac{8\pi^2}{h^3}\int \mathrm{e}^{-\beta(L_x^2/2I_{xx}+L_y^2/2I_{yy}+L_z^2/2I_{zz})}\mathrm{d}L_x\mathrm{d}L_y\mathrm{d}L_z. \tag{8.30}$$

The remaining integrals over the L's decouple and can be performed to give the resulting semiclassical rotational partition function

$$Z_{rot}(N) = \sqrt{\pi}\frac{T^{3/2}}{\sqrt{T_{rot,x}T_{rot,y}T_{rot,z}}}, \tag{8.31}$$

where the characteristic rotational temperatures are defined in analogy to the spherical case as

$$T_{rot,i} \equiv \frac{\hbar^2}{2I_{ii}}, \quad i = x, y, z. \tag{8.32}$$

For $I_{xx} = I_{yy} = I_{zz}$ (8.32) reduces to the expression for the spherical case, as it must.

Electronic excitations are much more specific to the chemical composition of the material than any of the other types of relevant excitations, and it is not possible to give a single, even very schematic, realistic description valid for all materials at the same time. We will therefore simply express that contribution in general terms as the canonical partition function in (8.14). It is in any case usually the contribution which is closest to unity (for the right choice of the zero of energy).

We can now put the partition functions for the separate degrees of freedom together:

$$\begin{aligned} Z_{tot}(N) &\approx \frac{1}{\pi}\left(\frac{2}{5}\right)^{3/2}\left(\frac{mT}{\hbar^2}\right)^3 r_1^3 V \left(\frac{T}{T_{vib}}\right)^{3N-6} N^4 Z_{el} \mathrm{e}^{\beta AN - \beta BN^{2/3}} \\ &\equiv c_N V \left(\frac{T}{T_{vib}}\right)^{3N-6} N^4 Z_{el} \mathrm{e}^{\beta AN - \beta BN^{2/3}} \quad (N \neq 1). \end{aligned} \tag{8.33}$$

For future use we also write down the partition function for the monomer. If it is not an atom, the following result needs to be multiplied by the partition function for the internal degrees of freedom. For an atom the partition function is simply that of translations,

$$Z_{tot}(1) = V\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2} \equiv c_1 V. \tag{8.34}$$

For the reference values $m = 1\,\mathrm{u}$ and $T = 293\,\mathrm{K}$, the constant has the value $c_1 = 9.5 \cdot 10^{29}\,\mathrm{m}^{-3}$. The corresponding factor for the multi-mers is defined as

$$c_N \equiv \frac{1}{\pi}\left(\frac{2}{5}\right)^{3/2}\left(\frac{mr_1 T}{\hbar^2}\right)^3 \tag{8.35}$$

and has, for $r_1 = 1\,\text{Å}$ and the same reference mass and temperature the value $1.8 \cdot 10^{31}\,\mathrm{m}^{-3}$.

Before we continue with calculations of the equilibrium distributions it is worthwhile to try to understand these two factors. It is easiest to begin with the monomer factor. The dimension of c_1 is that of a reciprocal volume. If one takes the cube root of this volume one gets the length

$$\lambda_t = \frac{h}{\sqrt{2\pi mT}}. \tag{8.36}$$

The denominator is a typical thermal momentum and λ_t is therefore on the order of the average wavelength of the monomer in its thermal motion, the so-called thermal de Broglie wavelength. The translational partition function can thus be expressed as

$Z_1 = V/\lambda_t^3$. If we consider the factor for the N-mer in this light, we see that it can be written as $c_N = c_1(r_1/\lambda_t)^3$. The reason the two factors c_1 and c_N differ by only one order of magnitude is that the thermal de Broglie wavelength of our reference atom under the reference conditions chosen is actually not very different from the size of an atom or a molecule. This is a result of the specific values of constants of nature, not of any thermal effect, and theories about thermal behavior assigning a different physical significance to this coincidence can safely be ignored.

8.3 Thermal and Chemical Equilibrium

To find out what the partition functions tell us about equilibrium abundances, consider a large collection of particles of all possible sizes in equilibrium with each other, including the gas of monomers which will also be present. We will examine the equilibrium between the monomer vapor and the particles containing N atoms. We can ignore the fact that other particles than those of sizes a and N are present and consider only atoms and N-mers. With the total number of monomers in these two types of particles denoted by n, we have $n = n_1 + Nn_N$, *c.m.n.*[1] The partition function for the collection of all N-mers is (see Chap. 2)

$$Z_N = \frac{Z_{tot}(N)^{n_N}}{n_N!}, \tag{8.37}$$

and similar for the monomer gas with N replaced by 1. In equilibrium the partitioning of atoms between the N-mers and the free atoms is given by the product distribution

$$P(n_N) = \frac{Z_{tot}(N)^{n_N}}{n_N!}\,\frac{Z_{tot}(1)^{n_1}}{n_1!} = \frac{Z_{tot}(N)^{n_N}}{n_N!}\,\frac{Z_{tot}(1)^{n-Nn_N}}{(n-Nn_N)!}. \tag{8.38}$$

We find the maximum of this distribution by setting the logarithmic derivative with respect to n_N to zero.

$$\frac{\partial}{\partial n_N}\ln\left(\frac{Z_{tot}(N)^{n_N}}{n_N!}\,\frac{Z_{tot}(1)^{n-Nn_N}}{(n-Nn_N)!}\right) = 0. \tag{8.39}$$

For the factorials we use Sterling's approximation (see Appendix C):

$$\ln(n!) = n\ln(n) - n + .. \tag{8.40}$$

and keep only the terms indicated. Using $n_1 = n - Nn_N$, a little algebra gives

[1] *cum manifestum notatio* = with an obvious notation.

$$n_N = Z_{tot}(N)\frac{n_1^N}{Z_{tot}(1)^N}. \tag{8.41}$$

Turning the equation around gives

$$n_1 = Z_{tot}(1)\left(\frac{n_N}{Z_{tot}(N)}\right)^{1/N}. \tag{8.42}$$

Introducing the partition functions from the previous section and considering a macroscopic amount of matter, $N \sim N_\mathrm{A} \ggg 1$, gives the abundance for monomers in equilibrium with a macroscopic amount of matter. With Stirling's approximation we get the number of monomers for the saturated vapor

$$n_1 = Z_{tot}(1)\left(\frac{T_{vib}}{T}\right)^3 \mathrm{e}^{-\beta A} = V\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\left(\frac{T_{vib}}{T}\right)^3 \mathrm{e}^{-\beta A}. \tag{8.43}$$

Let's digress to formulate the equilibrium condition in (8.39) in terms of the chemical potential. For the canonical partition function, the free energy is Helmholtz' free energy which is defined as (see Chap. 1)

$$Z_c = \frac{Z_{tot}(N)^{n_N}}{n_N!} = \mathrm{e}^{-\beta F}. \tag{8.44}$$

The chemical potential is the derivative of F with respect to the number of particles of the species;

$$\mu_M = \frac{\partial F}{\partial n_M} = -T\ln\left(Z_{tot}(M)/n_M\right), \tag{8.45}$$

where M can be 1 or N. If we use this with $M = N$ together with (8.42) we have

$$\mu_N = -T\ln\left(Z_{tot}(N)/n_N\right) = -T\ln\left(Z_{tot}(1)/n_1\right)^N = N\mu_1. \tag{8.46}$$

What we derived is therefore just the condition

$$\mu_N = N\mu_1. \tag{8.47}$$

Recapitulating, this relation was derived by considering the most probable partitioning of monomers between N-mer and monomer states.

We have now calculated the number of monomers in a given volume over a bulk piece of material. The corresponding pressure is known as the saturated vapor pressure. If the monomer is in equilibrium with the bulk and the N-mer is in equilibrium with both, we can calculate the N-mer concentration from the monomer abundances. For this purpose we use (8.41) with (8.43) with the requirement that N is much less than the macroscopic amount that qualifies as the bulk and which determines the vapor pressure. The result is

$$n_N = c_N V N^4 \left(\frac{T_{vib}}{T}\right)^6 \mathrm{e}^{-\beta B N^{2/3}}. \tag{8.48}$$

The factors that are left out in this approximation are

$$\left(\frac{T_{vib}}{T}\right)^{6N/N_0} N_0^{4N/N_0} (c_N V)^{-1/N_0} \mathrm{e}^{-\beta B N/N_0^{1/3}}, \tag{8.49}$$

and some tiny root of the electronic partition functions. All of these approach unity in the limit $N/N_0^{1/3} \to 0$. This condition is equivalent to considering only particles with radii much below 1 mm or so.

The distribution in (8.48) is the saturated vapor equilibrium abundances of particles of size N. The individual factors all have straightforward physical interpretations. The factor $c_N V N^4$ is the combined rotational and translational partition functions of the N-mer. The factor involving the sixth power of the vibrational temperature accounts for the decrease in the vibrational degrees of freedom that occurs when two particles are created out of one; $3N_0 - 6 \to 3(N_0 - N) - 6 + 3N - 6$. The specific power of six refers to harmonic oscillator level densities for the motion of the nuclei. The true values may be lower or higher and this will change the power from 6 to $2c_v$, where c_v is the heat capacity per atom, but this is usually not a major change. Finally, the exponential in (8.48) is the Boltzmann factor of the new surface area that appears when the N-mer is created (see Exercise 8.3). Note that the bulk cohesive energy per monomer, A, does not appear in (8.48).

An example of abundances calculated with (8.48) is shown in Fig. 8.4. The pre-exponential N^4 dependence is only a weakly growing function compared with the exponential decrease, and for $N = 100$, say, this term will only be a minor correction to a distribution which is already practically unobservable. Even when the temperature is so high that the value of βB is 2, which is usually a very high temperature, is the product of N^4 and the exponential less than 10^{-10} for this particle size.

The above calculation is based on schematic partition functions; as mentioned both anharmonicities in the vibrations and melting will increase the $N > 1$ abundances. Also finite size effects in the binding energies will be reflected in the abundances, and the results given here should only be considered as the trends. Nevertheless, the result derived clearly show that abundances of large cluster are vanishing in thermal and chemical equilibrium.

One may ask about the concentrations if the monomers are in chemical and thermal equilibrium with N-mers but not with a bulk phase. We can easily confirm our expectations that the N-mer concentrations are reduced in that case, compared to the bulk saturated vapor pressure. In the absence of a bulk phase, the monomer concentration must be less than the saturated vapor concentration. This means that the right hand side of (8.41), rewritten as

$$n_N = Z_{tot}(N) \left(\frac{n_1 V_q}{V}\right)^N, \tag{8.50}$$

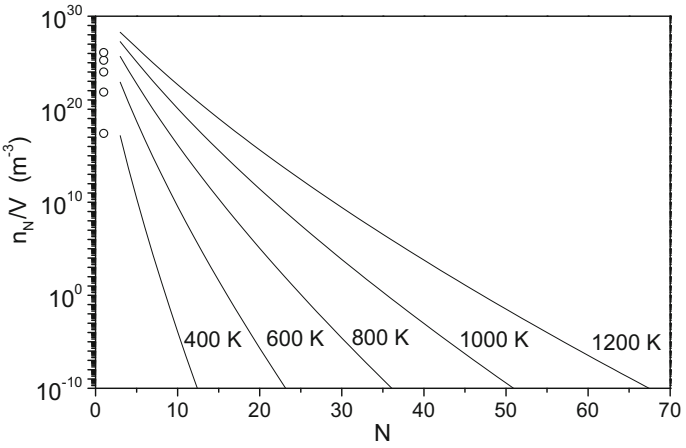


Fig. 8.4 Calculated abundances for clusters in equilibrium with the saturated monomer vapor pressure. The parameters used are $r_1 = 2.1$ Å, $m = 23$ u, $A = 1.12$ eV and $B = 0.718$ eV, corresponding to the values for sodium. The dimer is left out because the equation used to calculate the concentrations, (8.48), is not valid for this species. The monomer concentration is shown with open circles

will be reduced, as will therefore also n_N. If the lower pressure is obtained by reducing the monomer vapor pressure from the saturated pressure value, N-mers will dissociate and enrich the monomer concentration.

This is an example of what the chemists know as le Chatelier's Principle. In its general form the principle says that the system will tend to counteract an externally imposed change of its parameters. We will not calculate the concentrations as a function of the total concentration, but just conclude that in the absence of a bulk phase, the N-mer concentration is less than the already very small equilibrium value.

The most important lesson from this section is that particles of any reasonable size must be produced in some other way than by chemical equilibrium. The next sections describes the underlying statistical physics of some of those methods.

8.4 Polymerization

As a first attempt to explain quantitatively the size distributions and the rate of creation of particles from vapors (because they *do* exist, as we know), we will consider a model for polymerization of molecules. From our perspective the main virtue of this model is that it is exactly solvable and gives us some insight into size distributions in situations with thermal but not chemical equilibrium.

We will look at a collection of monomers and polymers up to an arbitrarily large size (polymers is a collective designation for X_N, where $N > 2$ and the bonds are of chemical nature). The collection is created by placing all monomers in a long string

and consider the distribution of sizes when every pair of neighboring monomers is bonded to each other with some probability p. A monomer has only two neighbors, and these never bond directly with each other. Effectively, this is therefore a one-dimensional theory. Keeping the model simple, the value of p is assumed the same over the whole length of the string and independent of the neighboring links.

The total number of monomers, bound to others or not, is a conserved number,

$$n_0 = \sum_{N=1}^{\infty} N n_N, \tag{8.51}$$

where n_N is the number of N-mers. We will take this number of monomers in the system to be so large that we can disregard end effects which are expected to influence the calculated probabilities with terms on the order of $1/n_0$.

The probability that a particular monomer belongs to a polymer of length N, which may be one, is the sum of the probabilities that the monomer is the jth monomer in the chain, summed over all j less than or equal to N. Each term is of the kind $(1-p)p^{j-1}p^{N-j}(1-p)$ because there is one broken link followed by $j-1$ unbroken ones before the jth monomer, and $N-j$ unbroken followed by one broken link after. Hence the total probability that the monomer is part of a polymer of length N is

$$\Pi_N = \sum_{j=1}^{N} (1-p)^2 p^{N-1} = N(1-p)^2 p^{N-1}. \tag{8.52}$$

The total number, n_t, of polymers of all sizes, including one, is

$$n_t = \frac{n_0}{\overline{N}}, \tag{8.53}$$

where $\overline{N}$ is the average polymer size. The average size is calculated as the probability that a chain is N monomers long, weighted with N.

The polymer size distribution is the probability that a given chain has a certain length. For any polymer, the first monomer has an broken link to one side, followed by $N-1$ bonds, which gives a factor of p^{N-1} on the probability. The terminating link, which is broken by definition, gives a factor of $1-p$. The probability that the chain is N monomers long is then

$$P_N = (1-p)p^{N-1}. \tag{8.54}$$

This is the polymer size distribution. It is different from the probability calculated in (8.52) which is the distribution of lengths of chains to which a randomly selected monomer belongs. Instead (8.54) gives the probability that a chain selected at random has length N. For a given p, the size distribution decreases with size as $\exp(N\ln(p))$, corresponding to a 'drop-off' scale of $-1/\ln(p)$.

The average size is

$$\overline{N} = \sum_{N=1}^{\infty} N P_N = \sum_{N=1}^{\infty} N(1-p)p^{N-1} = \frac{1}{1-p}. \tag{8.55}$$

For the number of polymers we then have

$$n_t = n_0(1-p). \tag{8.56}$$

Thus stationary state size distributions depend on p in a simple fashion.

Time does not appear explicitly in these distributions. In a more realistic situation where one starts with mainly or exclusively monomers, one expect p to increase with time from the value $p = 0$ to some finite value. If there is no reverse reaction, i.e. bonds don't break, the asymptotic value will be $p(t = \infty) = 1$. In the more general situation, with a bond formation rate of a and a bond breaking rate of b, we have the asymptotic value $p(\infty) = a/(a + b)$, which is easily found by equating the bond breaking and formation rates. All the way between the initial $p = 0$ and the asymptotic value, the distribution still only depends on p, but p will depend on time. The time derivative of p is the weighted sum of the rate constants for linking, a, and breakup, b:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = a(1-p) - bp, \tag{8.57}$$

which is easily understood as the sum of rate constants of formation and breakup, a and b, multiplied with the populations $1 - p$, p and applying the proper sign on the breakup term. We now just need to solve (8.57) to find $p(t)$, which then will give us the time dependence of all P_N. The standard technique is to compare (8.57) with the equation $\mathrm{d}(pg)/\mathrm{d}t = h$, which can be solved, and identify the functions g and h. The result is that $g = \exp((a + b)t)$, $h = ag$, and we then integrate h to get

$$p(t) = p(0)\mathrm{e}^{-(a+b)t} + \frac{a}{a+b}\left(1 - \mathrm{e}^{-(a+b)t}\right). \tag{8.58}$$

The problem of the time development can alternatively be formulated and solved by specifying all the terms that contribute to the growth and decay, and sum these up. We do it here because it is useful for comparison with more general descriptions of aggregation. It is a little easier to establish the equation that govern the rate of change of the number of polymers n_N instead of P_N, the relative population. To convert between the two, we have the following relation between n_N, the total number of polymers $n_0(1 - p)$, and the probability P_N:

$$n_N = P_N n_0(1-p), \tag{8.59}$$

The equation for the rate of change of n_N reads:

$$\frac{\mathrm{d}n_N}{\mathrm{d}t} = -2an_N + \sum_{i=1}^{N-1} an_{N-i}P_i - bn_N(N-1) + 2b\sum_{i=1}^{\infty} n_{N+i}. \tag{8.60}$$

The first term is the rate with which n_N decreases because either of the two ends fuse with a polymer of any length. The second term is the rate with which the corresponding fusion produces an N-mer from sizes $N-i$ and i. The third term represents the rate of breakup of an N-mer, which can occur at any of $N-1$ places, and the last term is the rate of increase due to the breakup processes that produce N-mers.

Instead of solving (8.60) directly, we will turn the argument upside down and use the known time dependence of p to show that the equation is correctly established. With the relation between n_N and P_N in (8.59), we can recast the equation into one for P_N:

$$\begin{aligned}\frac{\mathrm{d}P_N(1-p)}{\mathrm{d}t} =& -2aP_N(1-p) + \sum_{i=1}^{N-1} aP_{N-i}P_i(1-p) \\ & - bP_N(N-1)(1-p) + 2b(1-p)\sum_{i=1}^{\infty} P_{N+i}.\end{aligned} \tag{8.61}$$

With the help of the solutions for $P_N(p)$ in (8.54), the terms on the right hand side can be calculated. After a minor rearrangement of terms we get

$$\frac{\mathrm{d}P_N}{\mathrm{d}t} = p^{N-2}\left[(N-1)(1-p) - 2p\right](a(1-p) - bp) + p^{N-2}p\frac{\mathrm{d}p}{\mathrm{d}t}. \tag{8.62}$$

The last bracket in the first term is the time derivative $\mathrm{d}p/\mathrm{d}t$ and we therefore get

$$\frac{\mathrm{d}P_N}{\mathrm{d}t} = p^{N-2}\left[(N-1)(1-p) - 2\right](a(1-p) - bp)\,. \tag{8.63}$$

Calculating the left hand side directly from the solutions gives

$$\frac{\mathrm{d}P_N}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(1-p)p^{N-1} = \left[(N-1)p^{N-2} - Np^{N-1}\right]\frac{\mathrm{d}p}{\mathrm{d}t}. \tag{8.64}$$

Inserting (8.57) for the time derivative of p shows that this is identical to the time derivative calculated with (8.60), and consequently verifies this equation for this specific problem.

What we did here, in summary, was to write down and solve an equation that describes the temporal development of a simple model of polymer size distributions. The equation solved, (8.60), is a variation of what is known as the Smoluchowski

equation, or more correctly Smoluchowski equations, which describe the growth of particles by aggregation.

Although the polymer growth equation possesses the very appealing feature that it can be solved analytically, we must nevertheless establish a more realistic version for real life aggregation. The most severe shortcoming of the polymer equation is that it is one-dimensional whereas most systems are three-dimensional, or occasionally two-dimensional but rarely one-dimensional.

8.5 The Smoluchowski Equation

The Smoluchowski equations which will be established in more general form in this section is a set of ordinary, coupled and non-linear differential equations that describe the time development of the concentrations of particles as the result of aggregation, fusion and possibly break-up. Originally the equations were proposed without any break-up terms, corresponding to irreversible growth, but the expression is often used with such terms present, and this will also be the usage here.

We begin with the aggregation terms only. The rate of change of the concentration of species N, c_N is then

$$\left.\frac{\mathrm{d}c_N}{\mathrm{d}t}\right|_{growth} = \sum_{i=1}^{[N/2]} \frac{a_{i,N-i}}{2} c_i c_{N-i} - \sum_{i=1}^{\infty} (1+\delta_{i,N}) a_{i,N} c_N c_i . \tag{8.65}$$

The Kronecker delta in the negative sum appears because two particles of size N will fuse and disappear when i is equal to N, as opposed to the disappearance of a single N-mer for the processes where $i \neq N$. The upper limit of summation in the positive sum avoids double counting. Equation (8.65) can be rewritten as (see Exercise 8.17):

$$\left.\frac{\mathrm{d}c_N}{\mathrm{d}t}\right|_{growth} = \sum_{i=1}^{N-1} \frac{a_{i,N-i}}{2} c_i c_{N-i} - \sum_{i=1}^{\infty} a_{i,N} c_N c_i . \tag{8.66}$$

The coefficients $a_{i,j}$, known as the kernels, are symmetric in the indices,

$$a_{i,j} = a_{j,i}, \tag{8.67}$$

and have dimensions volume per time. The first sum in (8.66) is the production of size N particles from fusion of smaller particles, and the second is the depletion due to growth into larger particles. The factor $1/2$ compensates for double-counting in the first sum. The equations can be formulated as continuum integro-differential equation by treating the size as continuous (see Exercise 8.18), which is what Smoluchowski did.

The concentrations that enter the equations are mean values. That is a good procedure for sufficiently large systems, but for finite systems fluctuations will play a role. That can be taken into account by formulating the equations with probabilities as variables, and the equations are then known as the Markus-Lushinov equations. Introducing that complication is not making an already very difficult life any easier and we will tacitly ignore it and continue with the equations for the mean values with the expectation that they are sufficient for our purpose.

Even without the fragmentation terms, solution of (8.66) is a formidable task that has only been accomplished for a few different kernels. Smoluchowski provided the solution for constant kernels, $a_{i,j} = c$, for the $t = 0$ situation of monomers only;

$$c_N = \left(1 + \frac{t}{2}\right)^{-2} \left(\frac{t}{2+t}\right)^{N-1}, \quad a_{i,j} = c, \text{ all } i, j. \tag{8.68}$$

The constant c here and in the following two cases can be scaled away by a redefinition of the time as will be shown below with an explicit example, and we can assume $c = 1$. The solution of the equations in their continuous approximation has the corresponding form, with x denoting the size,

$$c_x = 4t^{-2}\mathrm{e}^{-2x/t}. \tag{8.69}$$

Already this solution, however schematic the physical situation it describes may be, displays an important property, namely that of scaling. The average size is

$$\langle x \rangle = \frac{t}{2}, \tag{8.70}$$

and with some loss of information about the specific solution but gain of generality, we can write the solution as

$$c_x = \langle x \rangle^{-2} \tilde{c}\left(\frac{x}{\langle x \rangle}\right). \tag{8.71}$$

In the constant kernel case, $\tilde{c}$ is an exponential function. In general these functions depend on the form of the kernels. An example of scaling will be calculated in detail later.

Another kernel for which the solution is known has the form

$$a_{i,j} = i + j. \tag{8.72}$$

This is must be expected to be more realistic for the description of real particles, with their fusion cross section which reasonably will grow with size. The solution of the continuous equations for this kernel is

$$c_x = (2\pi)^{-1/2}\, \mathrm{e}^{-t} x^{-3/2} \mathrm{e}^{-\mathrm{e}^{-2t}x/2}. \tag{8.73}$$

This solution has a diverging mean size. The problem arises at the low mass end of the spectrum. The problem is absent in the discrete solution, which has the less elegant form

$$c_N = \mathrm{e}^{-t} \frac{\left(N(1-\mathrm{e}^{-t})\right)^{N-1} \mathrm{e}^{-N(1-\mathrm{e}^{-t})}}{N!}. \tag{8.74}$$

A third exact solution has the interesting property that it gives gelation, i.e. causes the second moment of the size distribution size to be infinite after a finite time;

$$\langle N(t)^2 \rangle \to \infty, \quad t \to t_0 < \infty. \tag{8.75}$$

The kernel is

$$a_{i,j} = ij, \tag{8.76}$$

and the discrete solution is

$$c_N = N^{-1} \frac{(Nt)^{N-1} \mathrm{e}^{-Nt}}{N!}, \tag{8.77}$$

and the continuous

$$c_x = (2\pi)^{-1/2} x^{-5/2} \mathrm{e}^{-t^2 x/2}. \tag{8.78}$$

It can be shown that if the initial distribution is all monomers, the gelling transition occurs at the scaled time $t = 1$. In contrast, kernels that are homogeneous functions,

$$a_{si,sj} = s^{\gamma} a_{i,j}, \tag{8.79}$$

and have $\gamma \leq 1$ will not give rise to gelling.

After this brief overview of some mathematical results, we will specialize to a definite physical situation suitable to describe a number of situations one can encounter in practise. We will assume that concentrations of all species are uniform in space at all times, although they will in general of course depend on both size and time. The system is thus a gas composed of particles of different sizes, which is being homogenized continuously by the molecular motion so that rates can be expressed by position-independent concentrations. This assumption makes the probability for two particles to fuse proportional to the product of their concentrations.

The gas may or may not include an inert thermalizing component. The simplest equations are obtained if the thermalization with the gas is sufficiently fast compared to the rate of particle growth. With this condition, the heat of condensation which is added to the internal energy when two particles fuse will be dissipated and prevent that the large particles become hotter than the small ones. We will find quantitatively limits for this condition below for the aggregation-only case.

The a's are related to the collision frequencies between the two particles in a manner completely analogous to the fusion frequency we calculated in Chap. 5 on rate constants,

$$a_{i,j} = \sigma_{i,j} v_{i,j}, \tag{8.80}$$

where σ is the fusion cross section of the two particles and v is the average relative thermal speed. We will set the charge of all particles to zero. The cross section can be expected to be geometric to a reasonable approximation if the sticking coefficient is unity, i.e. they will fuse if they touch:

$$\sigma_{i,j} = \pi r_1^2 \left(i^{1/3} + j^{1/3}\right)^2. \tag{8.81}$$

The relative speed is identical to the average thermal speed of a particle with the reduced mass, $\mu_{i,j} = (1/im_1 + 1/jm_1)^{-1}$, because this is a two-body problem. It is calculated with the Maxwell-Boltzmann distribution to

$$v_{i,j} = \left(\frac{i+j}{ij}\frac{8T}{\pi m_1}\right)^{1/2}. \tag{8.82}$$

We therefore have for the kernels:

$$a_{i,j} = \pi r_1^2 \left(i^{1/3} + j^{1/3}\right)^2 \left(\frac{i+j}{ij}\frac{8T}{\pi m_1}\right)^{1/2} \equiv q \left(i^{1/3} + j^{1/3}\right)^2 \left(\frac{i+j}{ij}\right)^{1/2}. \tag{8.83}$$

The value of the parameter q, defined as

$$q \equiv \pi r_1^2 \left(\frac{8T}{\pi m_1}\right)^{1/2}, \tag{8.84}$$

depends only on the aggregating material and the temperature. It has dimension volume per time and will be used to scale the equations.

We will assume that this calculation is also valid for $i = j = 1$. The value of $a_{1,1}$ is important for the initial conditions where most of the gas is in monomeric form. The collision frequency should be given fairly well by (8.83), but energy conservation prevents dimers from being formed in collisions unless the monomers have internal degrees of freedom that can be excited in collisions with thermal energies and absorb the binding energy of the dimer. A third particle (monomer or cooling gas) may remove the excess energy in three-particle collisions and allow formation of a stable dimer, but this happens with a much lower rate than two-particle collisions. This mechanism is usually not very efficient and the dimer production therefore tends to act as a bottleneck for aggregation.

On the other hand, the bottleneck problem is often reduced by the presence of small clusters created prior to the production process. Also, evaporation of dimers or other small clusters from particles may occur and supply the gas with seeds of sizes beyond the bottleneck. In practice, particles are formed and we will use this as the justification to apply the simplest solution, viz. to keep the coefficients as described for all sizes. The assumptions are important to keep in mind when comparing theoretical predictions with experimentally measured spectra, though.

Even if the conditions for microscopically irreversible aggregation are fulfilled, the equations are still so complicated that they must usually be solved numerically because the coefficients are non-trivial functions of size. But although we cannot give the exact solutions in closed form, there are still things we can say about them. The first statement concerns time scales. If we use our reference values $r_1 = 1\,\text{Å}$, $T = 293\,\text{K}, m_1 = 1\,\text{u}$, we get the scaling parameter $q(r_1, T, m_1) = q_0 = 4.6 \cdot 10^{-18}\,\text{m}^3/\text{s}$. If we also use a standard initial monomer pressure of $P_0 = 1\,\text{mbar} = 100\,\text{Nm}^{-2}$, corresponding to an ideal gas concentration of $c_0 \equiv N/V = P/T = 2.47 \cdot 10^{22}\,\text{m}^{-3}$, we get the typical time for the aggregation-only situation of

$$t_0 \equiv 1/q_0 c_0 = 8.8\,\mu\text{s}. \tag{8.85}$$

The aggregation-only equations can be scaled by division of all concentrations by c which corresponds to setting the total monomer concentration equal to unity. If we also rescale the time to be the (dimensionless) time $\tau \equiv tqc$, the equations become

$$\begin{aligned}\frac{\mathrm{d}c_N}{\mathrm{d}\tau} = &\sum_{i=1}^{N-1} \frac{1}{2}\left(i^{1/3} + (N-i)^{1/3}\right)^2 \left(\frac{N}{i(N-i)}\right)^{1/2} c_i c_{N-i} \\ &- \sum_{i=1}^{\infty} \left(i^{1/3} + N^{1/3}\right)^2 \left(\frac{i+N}{iN}\right)^{1/2} c_N c_i,\end{aligned} \tag{8.86}$$

or in the more general form:

$$\frac{\mathrm{d}c_N}{\mathrm{d}\tau} = \sum_{i=1}^{N-1} \frac{1}{2} a_{i,N-i} c_i c_{N-i} - \sum_{i=1}^{\infty} a_{i,N} c_N c_i. \tag{8.87}$$

8.6 Conditions for Irreversible Aggregation

The condition that the aggregation is irreversible, also microscopically, is implicit in the treatment so far. This is not trivially fulfilled. Not only must the average size grow with time, but it also means that all monomers caught on some particle will stick for good and all particle-particle fusion processes is irreversible.

As an aside we note that it is not trivial either if the assumption of a spherical shape is consistent with this condition on the aggregation and equilibration speeds, because at low temperatures a low atomic mobility of the atoms on the surface of the particle may effectively prevent migration into a good spherical minimum energy candidate shape, and thereby result in shapes different from spherical, for example fractal structures. This is a Pandora's box question and we will keep the box closed.

We can, however, make an estimate of the temperature needed to make the aggregation irreversible, and this section will treat that question. First we note that without a cooling gas or some other cooling process, it is not possible to have irreversible

growth at long times because as the particles grow, matter will be found in larger and larger lumps, and these will contain the heat of condensation of the atoms released during formation. The temperature of these particles will converge toward a value on the order of $E_a/3$, which is very high and will be impossible to sustain for any reasonable period of time for individual particles. The large particles will simply re-evaporate if not cooled by an external agent.

The main cause of shrinking of particles is loss by monomer evaporation, and in some cases by loss of other small molecular fragments. Those situations arise if, for example, dimers are sufficiently stable. Then their binding energy will compensate the extra energy needed to eliminate two monomers at the same time from the shrinking particle. We will formulate the calculation in terms of monomer loss. A dimer loss, for example, is covered by the same equations if the proper energies for that process are used and will cause no change in the results.

We write down the requirement on irreversible aggregation as the condition that re-evaporation rates are small compared with aggregation rates,

$$\omega_{\overline{N},1} \mathrm{e}^{-\beta E_a} \ll v_{\overline{N}}\, c_{\overline{N}}\, \sigma_{\overline{N},\overline{N}}, \tag{8.88}$$

where $c_{\overline{N}}$ is the concentration of the average size $\overline{N}$, $v_{\overline{N}}$ is the average thermal speed of, σ is the attachment cross section, $\omega_{\overline{N},1}$ is the frequency factor for loss of a monomer from a particle of size $\overline{N}$, and $T = 1/\beta$ is the particle temperature which is elevated relative to the cooling gas temperature due to the heat of condensation deposited into the particles. The parameters scale with size approximately as

$$\omega_{\overline{N},1} = \omega \left(\overline{N}^{1/3} + \overline{N}^{1/3}\right)^2 = 4\omega \overline{N}^{2/3} \tag{8.89}$$

$$v_{\overline{N}} = v_1 \sqrt{2}\, \overline{N}^{\,-1/2} \tag{8.90}$$

$$c_{\overline{N}} = c \overline{N}^{\,-1} \tag{8.91}$$

$$\sigma_{\overline{N},\overline{N}} = \sigma \left(\overline{N}^{1/3} + \overline{N}^{1/3}\right)^2 = \sigma \overline{N}^{2/3}, \tag{8.92}$$

where ω, the frequency factor for evaporation of an atom, given in Chap. 5, can be taken constant with a value of 10^{15} to $10^{16}\,\mathrm{s}^{-1}$, v_1 is the thermal speed of the monomer and the factor $\sqrt{2}$ appears because the reduced mass of the two particles of size $\overline{N}$ should be used, σ is the cross section for monomer-monomer collision, and the quantities with other subscripts refer to the quantities pertinent for those particle sizes. The overall concentration of atoms, c, is defined as

$$c = \sum_{k=1}^{\infty} k c_k \tag{8.93}$$

i.e. counting all atoms whether free or bound in a particle.

Inserting the estimates in (8.89) into (8.88) gives the condition

$$\omega e^{-\beta E_a} < \frac{\sqrt{2}}{\overline{N}^{3/2}} \frac{v_1 c \sigma}{a}, \tag{8.94}$$

where a is the smallest ratio of aggregation rate to evaporation rate we are prepared to accept for the former to be large compared to the latter. A value of $a = 20 = \exp(3)$ is reasonable. Solving for T we get

$$T < \frac{E_a}{\ln(\omega \overline{N}^{3/2}/v_1 c \sigma) + \ln(a/\sqrt{2})} \tag{8.95}$$

We now need to find out how the temperature of the particles change with time as they aggregate particles and are cooled by the gas. We start with the cooling term. To find this requires knowledge of the cooling action of the gas. We assume that a single collision between a cooling gas particle and a particle changes the internal energy with an amount proportional to the difference between the particle temperature T and the gas temperature T_g, $\Delta E_{cool} = \alpha(T - T_g)$. Collisions occur with a frequency given by the product of the gas speed, v_g, the gas concentration, c_g, and the particle-gas molecule cross section, $\sigma_{\overline{N},g}$. This gives the cooling term

$$\left.\frac{dT}{dt}\right|_{cool} = -\alpha \frac{(T - T_g)}{C_v} v_g c_g \sigma_{\overline{N},g} = -\alpha \frac{(T - T_g)}{C_v} v_g c_g \sigma_{1,g} \overline{N}^{2/3}, \tag{8.96}$$

where C_v is the particle heat capacity. The cooling coefficient, α, summarizes the degree of thermalization of the outgoing gas atom to the particle temperature after the collision. The incoming cooling gas atom carries $2T_g$ on the average and the outgoing $2T$ if it is completely thermalized to the particle temperature (see Chap. 5), and α is therefore 2 for complete thermalization. A value of 1 may be a more realistic number but we will keep this parameter free. The value of $\sigma_{\overline{N},g}$ has been set to $\sigma \overline{N}^{2/3}$, similar to the expression for the attachment cross section of atoms to particles of size $\overline{N}$, and we have ignored correction due to effective masses.

The heating rate is parametrized as the heat of fusion for two particles of size $\overline{N}$, ΔE, times their collision rate:

$$\left.\frac{dT}{dt}\right|_{heat} = \sqrt{2} v_1 c \sigma \frac{\Delta E}{C_v} \overline{N}^{-5/6}, \tag{8.97}$$

where the collision frequency on the right hand side of (8.94) was used. With liquid drop binding energies and the parameters E_a and B, we have the heat of fusion $\Delta E = B\overline{N}^{2/3}(2 - 2^{2/3}) \approx 0.28 E_a \overline{N}^{2/3}$, where we have used the approximate empirical relation $B = 2E_a/3$, ignoring the difference between E_a and the bulk value A.

Adding the cooling and heating equations with these choices and calculating numerical constants gives

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\frac{(T - T_g)}{C_v} v_g c_g \sigma_{1,g} \overline{N}^{2/3} + 0.39 v_1 c \sigma \frac{E_a}{C_v} \overline{N}^{-1/6}. \tag{8.98}$$

For the sake of argument we will now assume that the cooling gas temperature and density is constant, neglecting the effects of heating of the gas by the heat of condensation and cooling by e.g. adiabatic expansion in gas expansion sources. Then $\mathrm{d}T/\mathrm{d}t = \mathrm{d}(T - T_g)/\mathrm{d}t$ which simplifies the subsequent reasoning. To find the gas parameters that provide sufficient cooling to exclude re-evaporation, we require that the particle temperature remains constant and equal to the highest possible temperature consistent with our aggregation-only condition. The condition $\mathrm{d}T/\mathrm{d}t = 0$ gives

$$T - T_g = \frac{v_1 c \sigma}{v_g c_g \sigma_{1,g}} \frac{0.39 E_a}{\alpha \overline{N}^{5/6}}, \tag{8.99}$$

which must hold for all values of $\overline{N}$.

The required cooling gas concentration is found by combining (8.95, 8.99) and solving for the ratio of concentrations,

$$\frac{c_g}{c} \gtrsim \frac{0.39 E_a}{\alpha \overline{N}^{5/6}} \frac{v_1 \sigma}{v_g \sigma_{1,g}} \left[\frac{E_a}{\ln\left(\omega \overline{N}^{3/2} / v_1 c \sigma\right) + \ln(a/\sqrt{2})} - T_g \right]^{-1}. \tag{8.100}$$

This expression imposes the hardest conditions on the small particles, formally for $\overline{N} = 1$, and we can reformulate the condition, using this size, as

$$\frac{c_g}{c} \gtrsim \frac{0.39 E_a}{\alpha} \frac{v_1 \sigma}{v_g \sigma_{1,g}} \left[\frac{E_a}{\ln\left(\omega / v_1 c_0 \sigma\right) + \ln(a/\sqrt{2})} - T_g \right]^{-1}. \tag{8.101}$$

One of the important features in this result is the divergence on the required ratio of the gas to monomer concentration when T_g approaches a temperature which is some fraction of the binding energy. A more convenient estimate of this point is obtained if we dispense with our safety factor $\ln(a/\sqrt{2})$, set the prefactor except E_a equal to unity and rewrite (8.101) as

$$\frac{c_g}{c} \gtrsim E_a \left[\frac{E_a}{\ln\left(\omega / v_1 c \sigma\right)} - T_g \right]^{-1}, \tag{8.102}$$

or

$$\frac{c}{c_g} \lesssim \frac{1}{\ln(\omega / v_1 c \sigma)} - \frac{T_g}{E_a}, \tag{8.103}$$

which will often suffice in practical applications.

8.7 The Break-up Terms

If the cooling of the particles discussed in the previous section is insufficient, particles break up. This gives rise to two more terms in the equation:

$$\left.\frac{\mathrm{d}c_N}{\mathrm{d}t}\right|_{break-up} = -\sum_{i=1}^{N-1}\frac{1}{2}b_{N,i}c_N + \sum_{i=1}^{\infty} b_{N+i,i}c_{N+i}. \tag{8.104}$$

The first term describes the loss due to breakup, summing over all the possible channels, and the second the gain due to breakup of larger sizes. Adding these break-up terms to the growth terms gives us the set of equations

$$\begin{aligned}\frac{\mathrm{d}c_N}{\mathrm{d}t} &= \sum_{i=1}^{N-1} a_{N-i,i}c_i c_{N-i} - \sum_{i=1}^{\infty} a_{N,i}c_N c_i - \sum_{i=1}^{N-1}\frac{1}{2}b_{N,i}c_N + \sum_{i=1}^{\infty} b_{N+i,i}c_{N+i} \qquad (8.105)\\ &= \sum_{i=1}^{N-1}\frac{q}{2}\left(i^{1/3}+(N-i)^{1/3}\right)^2\left(\frac{N}{i(N-i)}\right)^{1/2} c_i c_{N-i}\\ &\quad - \sum_{i=1}^{\infty} q\left(i^{1/3}+N^{1/3}\right)^2\left(\frac{i+N}{iN}\right)^{1/2} c_N c_i - \sum_{i=1}^{N-1}\frac{1}{2}b_{N,i}c_N + \sum_{i=1}^{\infty} b_{N+i,i}c_{N+i},\end{aligned}$$

where the special choice of a-kernels were inserted in the last equality. Equation (8.105) is the generalized Smoluchowski equations for nucleation.

We can simplify the equations by restricting the decay channels to be open only for loss of a monomer, i.e. setting $b_{N,i} = b_{N,i}\delta_{i,1}$. This can be expected to be a good approximation for many materials and was already used in the previous section. If a liquid drop breaks into two particles, $N \to (N-n)+n$, the cost in energy is the increase in surface energy,

$$B(N-n)^{2/3} + Bn^{2/3} - BN^{2/3} \simeq Bn^{2/3}. \tag{8.106}$$

This energy is usually much bigger than the temperature and the decay channels open will therefore tend to be monomer evaporation, or loss of some other small fragments. This can happen for particularly stable dimers, as mentioned but also when the $N-n$ particle is particularly stable due to a strong shell structure (see Chap. 10).

For monomer loss only the complete set of equations can be written in the simplified form, using the notation $b_{N,1} \equiv b_N$

$$\frac{\mathrm{d}c_N}{\mathrm{d}t} = \sum_{i=1}^{N-1} a_{N-i,i}c_i c_{N-i} - \sum_{i=1}^{\infty} a_{N,i}c_N c_i - b_N c_N + b_{N+1}c_{N+1}. \tag{8.107}$$

8.8 Solution of the Aggregation Equation

With the physical conditions for an aggregation-only situation established in Sect. 8.6, an attempt on the solution of this version of the Smoluchowski equations can be made for the relevant kernels.

We start with the mean size, $\overline{N}$. It is given by

$$\overline{N} = \frac{\sum_{N=1}^{\infty} N c_N}{\sum_{N=1}^{\infty} c_N} = \frac{1}{\sum_{N=1}^{\infty} c_N}, \tag{8.108}$$

because of particle number conservation. The time derivative is

$$\frac{\mathrm{d}\overline{N}}{\mathrm{d}\tau} = -\frac{1}{\left(\sum_{N=1}^{\infty} c_N\right)^2} \sum_{N=1}^{\infty} \frac{\mathrm{d}c_N}{\mathrm{d}\tau} = -\overline{N}^2 \sum_{N=1}^{\infty} \frac{\mathrm{d}c_N}{\mathrm{d}\tau}. \tag{8.109}$$

We can replace the time derivatives on the right hand side by the aggregation Smoluchowski equation, (8.87) to get

$$\frac{\mathrm{d}\overline{N}}{\mathrm{d}\tau} = -\overline{N}^2 \sum_{N=1}^{\infty} \left(\sum_{i=1}^{N-1} \frac{1}{2} a_{i,N-i} c_i c_{N-i} - \sum_{i=1}^{\infty} a_{i,N} c_N c_i \right). \tag{8.110}$$

This looks rather complicated. We must do two double sums over a function we don't know and which is multiplied by some coefficients that are known but non-trivial. We first note that the two sums are actually identical because

$$\sum_{N=1}^{\infty} \sum_{i=1}^{N-1} a_{i,N-i} c_i c_{N-i} = \sum_{N=1}^{\infty} \sum_{i=1}^{\infty} a_{i,N} c_i c_N. \tag{8.111}$$

(See Exercise 8.17). The derivative therefore reduces to

$$\frac{\mathrm{d}\overline{N}}{\mathrm{d}\tau} = \overline{N}^2 \sum_{N=1}^{\infty} \sum_{i=1}^{\infty} \frac{1}{2} a_{i,N} c_i c_N. \tag{8.112}$$

The key observation that will help us out is that the a's are reasonably slowly varying functions, at least when neither of the indices are very small. We can therefore approximate the $a_{i,N}$ with $a_{\overline{N},\overline{N}}$, which is equal to $4\sqrt{2}\,\overline{N}^{1/6}$ and take this factor outside the summation signs,

$$\frac{\mathrm{d}\overline{N}}{\mathrm{d}\tau} = \frac{1}{2} 4\sqrt{2}\,\overline{N}^{1/6} \overline{N}^2 \sum_{N=1}^{\infty} \sum_{i=1}^{\infty} c_i c_N. \tag{8.113}$$

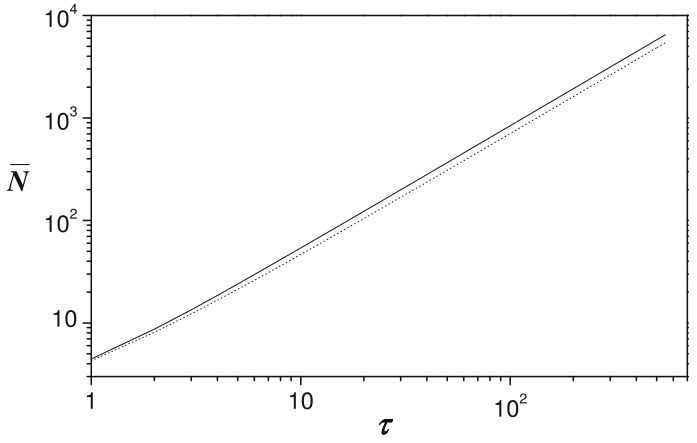


Fig. 8.5 Numerically simulated (full line) and approximate (dotted line) mean particle sizes for the aggregation-only Smoluchowski equations

The remaining two sums decouple and are identical and equal to $1/\overline{N}$ because of (8.108). We therefore have

$$\frac{\mathrm{d}\overline{N}}{\mathrm{d}\tau} = 2\sqrt{2}\,\overline{N}^{1/6}. \tag{8.114}$$

Note that this result is obtained without knowing the distribution. We solve the equation to

$$\overline{N} = \left(\frac{5\sqrt{2}}{3}\tau + \tau_0\right)^{6/5}. \tag{8.115}$$

To convert τ to a physical time, use $t = 4\tau/c_0 v_1 \sigma$.

Average sizes from numerically calculated solutions of (8.87) are shown in Fig. 8.5 together with the curve given by (8.115). The simulations were started with monomers at $\tau = 0$, which fixes τ_0 to be 1. The double-logarithmic plot makes it clear that the simulated mean size grows with a power slightly above unity. The main error in (8.115) is the deviation of a multiplicative factor from unity, whereas the power of 6/5 is reproduced reasonably well. The theoretical curve is 16% below the numerical at large τ. This is the result of the error made in the estimate of the constant factor on the right hand side of (8.109). An error of only 16% is surprisingly good considering the level of approximation in the calculation.

We now show explicitly that the kernels used in these equations will accommodate scaled solutions. When solutions scale, the concentrations can be written on the form

$$c_N = \frac{1}{\overline{N}^2}\tilde{c}\left(\frac{N}{\overline{N}}\right), \tag{8.116}$$

where $\tilde{c}$ is some, so far unknown, function and the scaling size can be taken to be the mean size, $\overline{N}$, without loss of generality. The prefactor $\overline{N}^{-2}$ takes into account the fact that summation of abundances weighted with the size should be independent of $\overline{N}$ because of particle number conservation:

$$\sum_{N=1}^{\infty} N c_N \approx \int_0^{\infty} \frac{N}{\overline{N}^2} \tilde{c}\left(\frac{N}{\overline{N}}\right) \mathrm{d}N = \int_0^{\infty} x\tilde{c}(x)\mathrm{d}x = 1. \tag{8.117}$$

The time derivative of the scaled concentrations are

$$\frac{\mathrm{d}c_N}{\mathrm{d}\tau} = -2\frac{\dot{\overline{N}}}{\overline{N}^3}\tilde{c}_N - \frac{N\dot{\overline{N}}}{\overline{N}^4}\tilde{c'}_N = \frac{\dot{\overline{N}}}{\overline{N}^3}\left(-2\tilde{c}_N - \frac{N}{\overline{N}}\tilde{c'}_N\right), \tag{8.118}$$

where the prime on $\tilde{c}$ indicates the derivative with respect to the argument. The term in the last bracket is only a function of $N/\overline{N}$. By the Smoluchowski equation this derivative is also equal to

$$\begin{aligned}\frac{\mathrm{d}c_N}{\mathrm{d}\tau} &= \sum_{i=1}^{N-1} \frac{1}{2} a_{i,N-i} c_i c_{N-i} - \sum_{i=1}^{\infty} a_{i,N} c_N c_i \\ &\approx {}^1\!/_2 \int_0^N a_{i,N-i} c_i c_{N-i} \mathrm{d}i - c_N \int_0^{\infty} a_{i,N} c_i \mathrm{d}i,\end{aligned} \tag{8.119}$$

where sums were approximated by integrals, which is permissible if the summation runs over a large number of terms of similar magnitude. The kernels are homogeneous functions with exponent 1/6:

$$a_{\alpha i,\alpha j} = \alpha^{1/6} a_{i,j}, \tag{8.120}$$

which, incidentally, shows that the kernels don't cause a gelling transition. If we use (8.120) together with the scaling expression in (8.116) for the concentrations, (8.119) can be written as

$$\frac{\mathrm{d}c_N}{\mathrm{d}\tau} \approx \overline{N}^{-17/6} \frac{1}{2} \int_0^N a_{x,y-x} \tilde{c}_x \tilde{c}_{y-x} \mathrm{d}x - \overline{N}^{-17/6} \tilde{c}_y \int_0^{\infty} a_{x,y} \tilde{c}_x \mathrm{d}x, \tag{8.121}$$

where $y \equiv N/\overline{N}$. The right hand side is a product of the factor $\overline{N}^{-17/6}$ and a function of $N/\overline{N}$ which we for notational simplicity will call g. Equating (8.118, 8.121) gives

$$\dot{\overline{N}}\,\overline{N}^{-1/6} = \frac{g\left(\frac{N}{\overline{N}}\right)}{f\left(\frac{N}{\overline{N}}\right)}, \tag{8.122}$$

where the shorthand $f\left(\frac{N}{\overline{N}}\right) \equiv -2\tilde{c}\left(\frac{N}{\overline{N}}\right) - \left(\frac{N}{\overline{N}}\right)\tilde{c}'\left(\frac{N}{\overline{N}}\right)$ has been used. The left hand side of this equation does not depend on N and the right hand side does not depend on time. Hence a separation of variables has been achieved, making both sides equal to the same constant. We have calculated this constant to be approximately $2\sqrt{2}$. We therefore have

$$\frac{g\left(\frac{N}{\overline{N}}\right)}{f\left(\frac{N}{\overline{N}}\right)} = 2\sqrt{2}. \tag{8.123}$$

What we have seen is that there are solutions to the aggregation-alone Smoluchowski equation that scale and that the mean value varies as in (8.115), when the coefficients in the Smoluchowski equation vary as in (8.120). Note that to derive this result, the precise functional form of the coefficients are not important, only that they have the scaling that result in the factor $\alpha^{1/6}$. For the same reason we cannot expect to get the constant of proportionality for the scaled solution (if all coefficients were a factor 100 smaller, the scaling would be exactly the same but $\overline{N}$ would definitely be much smaller for a given time).

We can, however, use the scaling properties to find a partial differential equation for the solution. The partial derivative with respect to time is

$$\begin{aligned}\frac{\partial c_N}{\partial \tau} &= -2\frac{\dot{\overline{N}}}{\overline{N}^3}\tilde{c}\left(\frac{N}{\overline{N}}\right) - \frac{N\dot{\overline{N}}}{\overline{N}^4}\tilde{c}'\left(\frac{N}{\overline{N}}\right) \\ &= -2\frac{\dot{\overline{N}}}{\overline{N}}c_N - \frac{N\dot{\overline{N}}}{\overline{N}^4}\tilde{c}'\left(\frac{N}{\overline{N}}\right)\end{aligned} \tag{8.124}$$

where the dot and the prime indicate differentiation with respect to time and the argument of the function, respectively. The derivative with respect to size is

$$\frac{\partial c_N}{\partial N} = \frac{1}{\overline{N}^3}\tilde{c}'\left(\frac{N}{\overline{N}}\right). \tag{8.125}$$

Substituting the last equation into the first and using that

$$\frac{\mathrm{d}\overline{N}}{\mathrm{d}\tau} = \frac{6}{5}\frac{\overline{N}}{\tau} \tag{8.126}$$

gives

$$\frac{\partial c_N}{\partial \tau} = -\frac{6}{5\tau}\left(2c_N + N\frac{\partial c_N}{\partial N}\right) \Rightarrow \frac{\partial \ln(c_N)}{\partial \ln \tau} = -\frac{6}{5}\left(2 + \frac{\partial \ln(c_N)}{\partial \ln N}\right). \tag{8.127}$$

As seen by inspection, this equation is solved by a log-normal function which is a function of the form

$$c_N = c'\overline{N}^{-2} \exp\left(-\frac{1}{2s^2}\left(\ln(N) - \ln(N_0)\right)^2\right), \tag{8.128}$$

provided the mean size has the dependence on τ we found previously (8.115).

The constants of integration c', s and N_0 are unknown at this point, but the known time dependence of the mean size and mass conservation help to fix relations between these quantities. Replacing summation with integration, the mass conservation reads

$$1 = c' \int_0^\infty N\overline{N}^{-2} \exp\left(-\frac{1}{2s^2}\left(\ln(N) - \ln(N_0)\right)^2\right) \mathrm{d}N. \tag{8.129}$$

With the substitution $x = \ln(N/N_0)/\sqrt{2}s$ the integral is calculated by completing the square which gives the normalization constant as

$$c' = \frac{\overline{N}^2}{N_0^2} \frac{1}{s\sqrt{2\pi}} \mathrm{e}^{-2s^2}. \tag{8.130}$$

The reciprocal of the mean size is calculated similarly;

$$\overline{N}^{-1} = \sum_{N=1}^{\infty} c_N = c' \int_0^\infty \overline{N}^{-2} \exp\left(-\frac{1}{2s^2}\left(\ln(N) - \ln(N_0)\right)^2\right) \mathrm{d}N = N_0^{-1} \mathrm{e}^{-3s^2/2}. \tag{8.131}$$

This relation says that the average size is larger than the peak value of the size distribution by the factor $\exp(3s^2/2)$ and that therefore

$$c' = \frac{1}{s\sqrt{2\pi}} \mathrm{e}^{s^2}. \tag{8.132}$$

It also implies that the peak value size, N_0, varies with time as $\overline{N}$, i.e. as $\tau^{6/5}$. This is of course exactly as required for a function that scales as prescribed. In particular we have that

$$\frac{\dot{\overline{N}}}{\overline{N}} = \frac{\dot{N}_0}{N_0}, \tag{8.133}$$

which will be used below.

To find s, we calculate the time derivative of some specific concentration. The best choice of this concentration is one not too far from the peak value because this is representative of the bulk part of the distribution, which is where our approximate solution is expected to be best. We will therefore use the peak size, N_0. From the scaled expression in (8.128) we have

$$\dot{c}_{N_0} = -2c' \frac{\dot{\overline{N}}}{\overline{N}} \overline{N}^{-2} = -4\sqrt{2}c'\overline{N}^{-2-5/6}, \tag{8.134}$$

where we have made use of (8.133) and the known time dependence of $\overline{N}$.

This derivative can also be calculated with the (still aggregation-only) Smoluchowski equation. For c_{N_0} we have:

$$\dot{c}_{N_0} = {}^1\!/_2 \sum_{i=1}^{N_0-1} a_{N_0-i,i} c_{N_0-i} c_i - c_{N_0} \sum_{i=1}^{\infty} a_{N_0,i} c_i . \tag{8.135}$$

In the loss term we approximate the a by the replacement $i \to \overline{N}$. With a constant a the sum gives $1/\overline{N}$, making this term approximately equal to $-c_{N_0} a_{N_0,\overline{N}}/\overline{N}$. After inserting the expression for c_{N_0} from the scaled solution, (8.128), the term becomes:

$$-\, c' \overline{N}^{\,-3} a_{N_0,\overline{N}}. \tag{8.136}$$

We can express a in terms of $\overline{N}$ and s alone because N_0 and $\overline{N}$ are related. This gives the loss term

$$-\, c' \, \overline{N}^{\,-2-5/6} (1 + \mathrm{e}^{-s^2/2})^2 (1 + \mathrm{e}^{3s^2/2})^{1/2}. \tag{8.137}$$

What remains to be calculated in (8.135) is the gain term which is a self-convolution of the abundances. It is calculated with a saddle point expansion:

$$\begin{aligned}
&{}^1\!/_2 \sum_{i=1}^{N_0-1} a_{N_0-i,i} c_{N_0-i} c_i \qquad (8.138)\\
&\approx {}^1\!/_2 \; c'^2 \overline{N}^{\,-4} \int_0^{N_0} a_{N_0-i,i} \mathrm{e}^{\left(-\frac{1}{2s^2}(\ln N - \ln N_0)^2 - \frac{1}{2s^2}(\ln(N_0-N) - \ln N_0)^2\right)} \mathrm{d}N.\\
&\approx {}^1\!/_2 \; c'^2 \, \overline{N}^{\,-4} a_{N_0/2,N_0/2} \, \mathrm{e}^{-(\ln 2)^2/2s^2} \int_0^{N_0} a_{N_0-i,i} \mathrm{e}^{-\frac{4}{s^2 N_0^2}(N - N_0/2)^2} \mathrm{d}N\\
&\approx c' \;\; \overline{N}^{\,-2-5/6} 2^{-1/6} \mathrm{e}^{-(\ln 2)^2/2s^2 - 3/4\, s^2}
\end{aligned}$$

where the value of c' was used in the last equality.

Setting the time derivative of (8.134) and the sum of (8.137, 8.138) equal gives the solutions $s = 0.98$ and $s = 0$, of which only the former is physically relevant.

An example of a numerically calculated distribution is shown in Fig. 8.6. The main qualitative difference to the distributions generated by the polymerization equation, where the coefficients on the production and loss terms are size independent, is the strong reduction of abundances at the smallest sizes. The same distribution is shown in a log-lin plot in Fig. 8.7, from where it is clear that it is very close to a log-normal distribution. The numerically determined value of s at large $\overline{N}$ is 1.27, which is a little larger than the value estimated above. The value is reached already around $\tau = 30$. This gives a ratio of the two masses where the concentrations are half the maximum values of $\exp(2s\sqrt{2 \ln 2}) = 20$, which is a very wide distribution, by any measure. The full-width-half maximum is equal to $4.24 N_0 = 0.34 \overline{N}$. The standard deviation

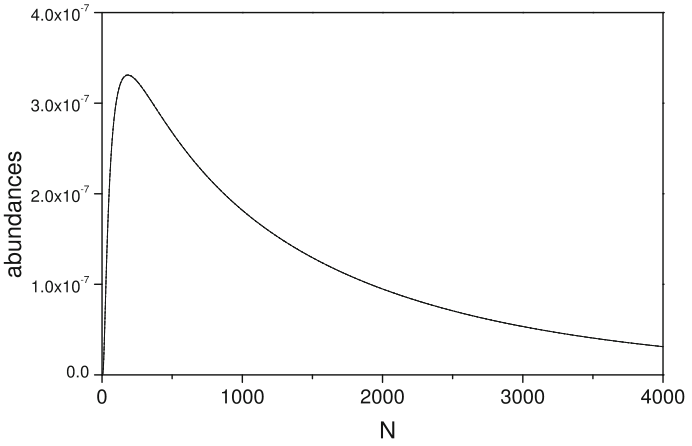


Fig. 8.6 The numerically calculated aggregation-only abundance spectrum at $\tau = 188$. The abundance spectra at $\tau = 376, 549$ are also plotted after the scaling described in text. They scale so well that they coincide with the $\tau = 188$ curve and are not visible in the figure

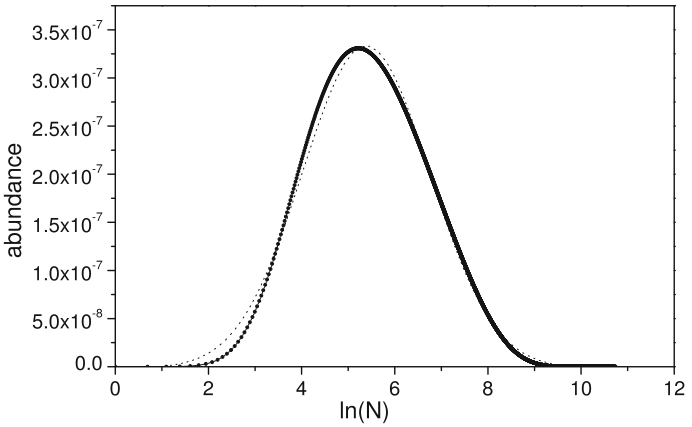


Fig. 8.7 The aggregation-only abundance spectrum at $\tau = 188$, but here plotted with log-lin scales. The dotted line is a fit with a log-normal distribution

of the size distribution, which is probably the most relevant of these parameters of the distribution from an experimental point of view is $\sigma = 1.5\overline{N}$.

A test of the scaling of the numerically calculated distributions is made by comparing spectra for different times. From the scaling equation we have for the spectra sampled at τ_1 and τ_2 that

$$c_{N'}(\tau_2)\frac{\overline{N}(\tau_2)^2}{\overline{N}(\tau_1)^2} = c_N(\tau_1), \tag{8.139}$$

where

$$N' \equiv N \frac{\overline{N}(\tau_2)}{\overline{N}(\tau_1)} \tag{8.140}$$

The numerical simulations indicate that scaling holds very well. Figure 8.6 shows three curves that are scaled as prescribed above and are indistinguishable on the scale of the figure.

8.9 Supersaturated Gases and the Critical Size

In the previous sections of this chapter we have been concerned with two extreme cases of particle abundances, the distribution of sizes in complete thermal and chemical equilibrium at one end of the scale, and irreversible aggregation at the other, as well as the less extreme cases of microscopically reversible but macroscopically irreversible growth described with the generalized Smoluchowski equations.

The cases that described growth all displayed a growth for all particles, akin to the picture shown in Fig. 8.7. In this and the following section we will look at near-equilibrium situations where growth occurs, albeit slow enough to make equilibrium considerations the starting point of the description. This leads to a growth that is concentrated on a very small number of outliers, in what can be considered fluctuation-based growth.

The situation is realized physically by increasing the vapor pressure of the monomer gas above the saturated vapor pressure and is therefore a non-equilibrium situation with respect to chemical composition, whereas thermal equilibrium is maintained to a good approximation. Such a vapor is called supersaturated. The supersaturation, s, is the relative vapor pressure above the saturated value, p_0, at the given temperature:

$$s = \frac{p}{p_0} - 1. \tag{8.141}$$

When s is positive but not too big, the whole system is in a quasi-equilibrium state where growth and decay almost balance but where there is a small net growth with time, concentrated on a very small number of particles. The growth process involves a barrier in the free energy for these drops. The presence of the barrier means that below a certain critical size, a drop will tend to evaporate back into monomers and that above this size it will experience a net accumulation of monomers and keep growing, rolling down the free energy hill.

This situation is an idealized description of atmospheric nucleation and the concept of a critical size is of overwhelming importance in the sectors of meteorology and climate science that deals with nucleation of water vapor and other species into small particles that ultimately form clouds or raindrops. An understanding of the origin and consequences of the critical size is therefore essential to begin to understand nucleation.

We will parameterize the free energy of a drop, irrespective of size, in analogy to the liquid drop expression for ground state energies as

$$F_d = Nf + 4\pi r_1^2 N^{2/3}\gamma = Nf + BN^{2/3}, \tag{8.142}$$

where N is the number of monomers in the drop, f is the bulk free energy per molecule in the condensed phase, $4\pi r_1^2 N^{2/3}$ is the surface area of the drop, and γ the surface tension. The parametrization is equivalent to assuming that the surface tension accounts for the whole difference between the bulk and the drop free energy. Expression (8.142) gives a chemical potential of

$$\frac{\partial F_d}{\partial N} = \mu_d = f + \frac{8\pi}{3} r_1^2 \gamma N^{-1/3}. \tag{8.143}$$

The molecules in the gas phase have a chemical potential which can be calculated with the partition function

$$Z_g = \frac{1}{n!}\left(\frac{V z_{int}}{V_q}\right)^n, \tag{8.144}$$

where c_1 is the number of gas phase molecules, z_{int} the partition function of the internal degrees of freedom for a single free monomer, and V_q is the thermal quantum volume. Both z_{int} and V_q depend only on the temperature, and z_{int} includes all degrees of freedom of the molecule apart from the translational. The gas chemical potential is then

$$\mu_g = T \ln\left(\frac{c_1 V_q}{V z_{int}}\right) = \mu_0 + T \ln\left(\rho/\rho_0\right), \tag{8.145}$$

where ρ_0 is a reference value of $\rho \equiv n/V$. If we chose ρ_0 to be the value for a saturated vapor, the value of μ_0 will be f. Then

$$\mu_d - \mu_g = \frac{8\pi}{3} r_1^2 \gamma N^{-1/3} - T \ln(\rho/\rho_0). \tag{8.146}$$

For $\rho = \rho_0$ the drop's chemical potential is always above that of the gas, although the difference goes to zero for asymptotically large drops, consistent with our choice of reference density. For $\rho < \rho_0$, the gas is undersaturated and then $\mu_d > \mu_g$ for all sizes, i.e. drop formation is uphill all the way and does not occur, except as a fluctuation excursion. For $\mu_d = \mu_g$, the drop is in equilibrium with the gas. Using the ideal gas law we have $c_1 = P/T$, and the equilibrium pressure over a drop of radius r is therefore

$$P = P_0 \exp\left(\frac{2\gamma}{nrT}\right), \tag{8.147}$$

where $n \equiv (4\pi r_1^3/3)^{-1}$ is the monomer number density in the drop. This is known as the Kelvin equation (after the person, not the unit).

When the gas is supersaturated, $\rho > \rho_0$ by definition. The chemical potential difference is still positive for small enough N, but crosses to a negative value at the critical size, N_c, determined by $\mu_g = \mu_d$. From (8.146) we get

$$0 = \mu_d - \mu_g = \frac{8\pi}{3} r_1^2 \gamma N_c^{-1/3} - T \ln(\rho/\rho_0), \tag{8.148}$$

which is easily solved to give

$$N_c^{1/3} = \frac{8\pi}{3} \frac{r_1^2 \gamma}{T \ln(\rho/\rho_0)}. \tag{8.149}$$

Particles with sizes below this critical size will tend to evaporate spontaneously because the evaporation rate is slightly higher than the aggregation rate, whereas those above will grow because the opposite inequality holds. For small values of the supersaturation

$$s = \frac{\rho}{\rho_0} - 1 \ll 1, \tag{8.150}$$

one can expand the logarithm to get the critical size

$$N_c^{1/3} = \frac{8\pi}{3} \frac{r_1^2 \gamma}{Ts} = \frac{2B}{3sT}. \tag{8.151}$$

Another way to find the critical size is to consider the difference in free energies of the gas and the drop, instead of the difference in chemical potentials. From (8.142) and the integrated version of (8.145) we get

$$\Delta F = 4\pi r_N^2 \gamma - NTs = BN^{2/3} - NTs. \tag{8.152}$$

The free energy difference peaks at N_c where the barrier height is

$$\Delta F_c = BN_c^{2/3} - N_c Ts = \frac{4}{27} B \left(\frac{B}{Ts} \right)^2 = \frac{1}{3} BN_c^{2/3}. \tag{8.153}$$

This value enters the quasi-equilibrium distributions as the argument of an exponential:

$$\exp\left(-\Delta F_c / T\right) = \exp\left(-\frac{4B^3}{27 s T^3} \right). \tag{8.154}$$

The free energy difference in (8.152) is illustrated in Fig. 8.8 for water which has the value $B = 0.22\,\text{eV}$, for $T = 300\,\text{K}$ and a relative supersaturation of 10%. Also shown is the free energy difference given as a function of the cube root of N. The critical size has a radius of 11 nm for this supersaturation. Even at this high supersaturation is the height of the free energy barrier still much higher than the temperature, which is measured in a couple of tens of meV. When describing

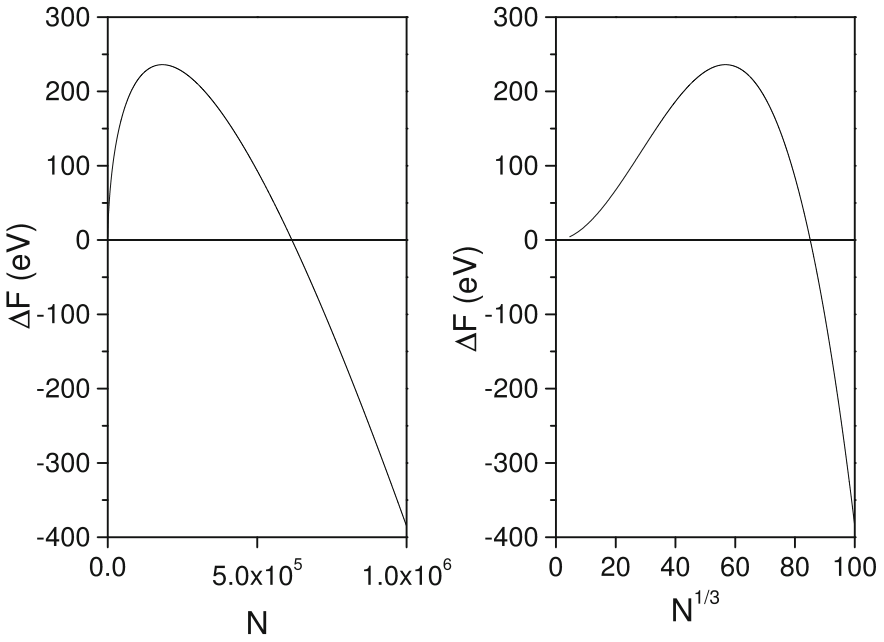


Fig. 8.8 The difference in free energy of an N water molecule particle and the state where all molecules are in the gas phase. The parameters are given in the main text

atmospheric nucleation rates, the barrier height enters as the factor $\exp(-\Delta F/T)$. Even a relatively small error in ΔF, caused by an incorrect surface tension, will cause a major error after exponentiation. Calculated nucleation rates can be off by many orders of magnitude, and it remains a formidable challenge to describe the free energies with the required precision.

The presence of other components than water, or whichever species is providing the feedstock in a nucleation situation, may change the estimate of the critical size dramatically. The mechanism behind a reduced critical size is that the saturated vapor pressure is reduced because the condensing nucleus has a different composition. What is a low water supersaturation for water drops may therefore be a very large supersaturation for other species. A reduction in the critical size dramatically changes the free energy barrier for nucleation and can potentially compensate for even a minuscule concentration of the species.

A current development in nucleation theory focuses the attention on heterogeneous nucleation involving very small clusters of sulfates and ammonia and, in particular, amines which have very low water vapor pressures and reduced critical sizes, reaching down to clusters composed of a few molecules, even at realistic supersaturation levels. An electric charge on the droplet will have the same effect, and this has observable consequences for atmospheric nucleation rates, although how much the effect matters in the big picture is still a disputed question in the highly politicized field of global warming, or climate change.

8.10 Nucleation

Once the critical size has been introduced and the conceptual foundation for the process of quasi-equilibrium nucleation has been laid, we can attack the problem of the growth of particles beyond the critical size, with due attention to the caveats about the real life applications made in the previous section. The core of what follows is known as classical nucleation theory (CNT). It gives the rate of formation of supercritical condensation nuclei, a topic which is immensely important for the description of the formation of the condensed state, be it liquid or solid, from a supersaturated vapor. It also has bearings on the melting and freezing of bulk matter, although these processes can be somewhat more complicated to describe.

Classical nucleation theory considers the addition to and loss of monomers from particles, i.e. the process

$$X_N + X \leftrightarrow X_{N+1}. \tag{8.155}$$

As a first approximation, one may calculate the nucleation rate as the product of the quasi-equilibrium concentrations of the critical size and the monomer addition rate. The quasi-equilibrium concentrations are given by the free energy barrier in (8.153):

$$c_{N_c} = c_1 \exp\left(-\Delta F/T\right) = c_1 \exp\left(-\frac{4}{27}\frac{B^3}{T^3 s^2}\right). \tag{8.156}$$

The rate of monomer addition, the a's, are calculated assuming that the sticking coefficient (also known as the accommodation coefficient in these types of studies) is energy independent. Then aggregation coefficients are calculated as the thermal equilibrium values

$$a_{N-1,1} = \sigma_{N-1,1}\overline{v}_{N-1,1}, \tag{8.157}$$

where σ is the capture cross section of the $N-1$-mer and the monomer, assumed geometric,

$$\sigma_{N-1,1} = \pi r_1^2 \left((N-1)^{1/3} + 1\right)^2 \approx \pi r_1^2 N^{2/3} \tag{8.158}$$

and $\overline{v}_{N-1,1}$ is the thermal average relative speed of the two particles. As always for a two-body problem, this can be calculated as a one-body problem, and the result is given in (8.82). The mass in that relation is the reduced mass, $\mu_{N-1,1}$, but this can be set equal to the monomer mass here without loss of much precision;

$$\overline{v}_{N-1,1} \approx \left(\frac{8T}{\pi m_1}\right)^{1/2}. \tag{8.159}$$

The first approximation to the nucleation rate, J, is then

$$J = c_1 \pi r_1^2 N_c^{2/3} \left(\frac{8T}{\pi m_1}\right)^{1/2} c_1 \exp\left(-\Delta F/T\right) \tag{8.160}$$
$$= \pi r_1^2 \left(\frac{2B}{3sT}\right)^2 \left(\frac{8T}{\pi m_1}\right)^{1/2} c_1^2 \exp\left(-\frac{4}{27}\frac{B^3}{T^3 s^2}\right).$$

One may divide this by two because the process can go either way and loss of a monomer is almost equally probable in this quasi-equilibrium scenario, but this will turn out to the least of the corrections needed.

We immediately spot the need for one correction, which arise because the result is calculated with an inconsistent procedure. The concentration of the critical size is supposed to be depleted by a growth reaction for which there is no reverse reaction. Once the particles have grown to a certain sized they no longer contribute to the equilibrium. This is inconsistent with chemical equilibrium. We must therefore expect that the concentration of the critical size is not given simply by the free energy difference as in (8.156), and that it in fact must be smaller.

It turns out that this reduction can be calculated in the stationary state, i.e. when the concentrations are constant, except those of the large sizes that beyond the critical size. In this situation of lack of chemical equilibrium, all concentrations need to be reconsidered. The more rigorous theory, known as Becker-Döring theory, also has the addition and elimination of monomers as the starting point. It is, in fact, a variation of the Smoluchowski equations with only monomers present in any appreciable amount. Since both addition and break-up are predominantly by monomers, we can begin by un-cluttering the notation by defining

$$a_N \equiv a_{N,1}, \quad b_N \equiv b_{N,1}, \tag{8.161}$$

The rate of the net change of concentration of particles that change size from $N-1$ to N per time unit, J_N, is

$$J_N \equiv \frac{\mathrm{d}c_N}{\mathrm{d}t} = a_{N-1} c_1 c_{N-1} - b_N c_N. \tag{8.162}$$

The rate of change of the concentration c_N is the continuity equation for the flow described by the J_N's:

$$\frac{\mathrm{d}c_N}{\mathrm{d}t} = J_N - J_{N+1}. \tag{8.163}$$

This sets the stage for the calculation of the nucleation rate. This is done with a trick, based on the fact that the right hand side of (8.162) contains two terms of opposite sign, involving the concentrations of two consecutive cluster sizes. The strategy is to multiply the equation by a (size-dependent) constant which is chosen so that when the terms in the sequence are added, the sum becomes a telescoping series, which leaves us with the first and last term.

The first step is to eliminate the b's in favor of the a's. This is achieved with the relations between rates in complete chemical and thermal equilibrium. Then

all J_N's in (8.162) are zero. If we express the decay rate in terms of equilibrium concentrations, denoted by the superscript e, we have

$$b_N = a_{N-1}\frac{c^e_{N-1}c^e_1}{c^e_N}. \tag{8.164}$$

Inserting this b_N, (8.162) becomes

$$J_N = a_{N-1}c_1c_{N-1} - a_{N-1}\frac{c^e_{N-1}c^e_1}{c^e_N}c_N. \tag{8.165}$$

If we divide this equation by $a_{N-1}c^e_{N-1}c^e_1$ and use that $c_1/c^e_1 = 1 + s \equiv S$ is the vapor pressure relative to the equilibrium value, the rate can be expressed as

$$\frac{J_N}{a_{N-1}c^e_{N-1}c^e_1} = S\frac{c_{N-1}}{c^e_{N-1}} - \frac{c_N}{c^e_N}. \tag{8.166}$$

Division by S^N then gives

$$\frac{J_N}{S^N a_{N-1}c^e_{N-1}c^e_1} = S^{-(N-1)}\frac{c_{N-1}}{c^e_{N-1}} - S^{-N}\frac{c_N}{c^e_N}. \tag{8.167}$$

This is the desired form for the terms in the telescoping series. In a stationary state where $J_N = J$ for all N, this also becomes the equation that determines the nucleation rate. Summing (8.167) from 2 to a large size, M say, gives

$$J = \left[\sum_{N=2}^{M-1}\left(S^N a_{N-1}c^e_{N-1}c^e_1\right)^{-1}\right]^{-1}\left(1 - \frac{c_M}{S^M c^e_M}\right) = \left[\sum_{N=2}^{M-1}\left(S^N a_{N-1}c^e_{N-1}c^e_1\right)^{-1}\right]^{-1} \tag{8.168}$$

where the term $c_M/S^M c^e_M$ vanishes for sufficiently large values of M because $S > 1$ and $c_M < C^e_M$ (see Exercise 8.23). In the evaluation of the sum we first approximate a_{N-1} with a_{N_c} and take this factor and the monomer equilibrium concentration times the supersaturation out of the sum. Setting the summation limit to infinity gives

$$J = a_{N_c}Sc^e_1\left[\sum_{N=1}^{\infty}\left(S^N c^e_N\right)^{-1}\right]^{-1}. \tag{8.169}$$

The reciprocal of the terms in the sum are equal to

$$S^N c^e_N = S^N c^e_1 \mathrm{e}^{-\beta\Delta F^e_N} = c^e_1 \mathrm{e}^{-\beta\Delta F^e_N + N\ln(S)} = c^e_1 \mathrm{e}^{-\beta\Delta F_N}, \tag{8.170}$$

where ΔF^e_N is the equilibrium free energy difference and ΔF_N the free energy difference for the supersaturated situation.

These terms have a minimum at the critical size (this is how it is defined). The smallest reciprocal term gives the largest contribution. If only this terms is included in the sum, the nucleation rate is, with $c_1 = c_1^e S$, equal to

$$J \approx (a_{N_c} c_1) c_1^e \mathrm{e}^{-\beta \Delta F_N}. \tag{8.171}$$

This agrees with the previous estimate. We now know how to improve this, viz. by a better approximation of the sum in (8.169). For this purpose the sum is converted to an integral,

$$\sum_{N=1}^{\infty} \left(S^N c_N^e\right)^{-1} \approx {c_1}^{e-1} \int_{-\infty}^{\infty} \mathrm{e}^{\beta \Delta F_N} \mathrm{d}N, \tag{8.172}$$

which is calculated by a saddle point expansion of the integrand using (8.152) to give

$$\int_{-\infty}^{\infty} \mathrm{e}^{\beta \Delta F_N} \mathrm{d}N \approx \frac{4\sqrt{\pi}}{3} (\ln S)^{-2} \left(\frac{B}{T}\right)^{3/2} \exp\left(\beta \Delta F_{N_c}\right). \tag{8.173}$$

Inserting this into (8.169) we get the final result which is the so-called classical nucleation rate per volume

$$J = a_{N_c} S (c_1^e)^2 \frac{3}{4\sqrt{\pi}} (\ln S)^2 \left(\frac{T}{B}\right)^{3/2} \mathrm{e}^{-\beta \Delta F_{N_c}}. \tag{8.174}$$

The extra factor due to the summation is known as the Zeldovitch factor.

$$Z = \frac{3}{4\sqrt{\pi}} (\ln S)^2 \left(\frac{T}{B}\right)^{3/2}. \tag{8.175}$$

It is less than unity, as expected. It is essentially the population of the critical size relative to the value one would have in the absence of nucleation, and is calculated self-consistently in the presence of the irreversible drain on the population of this size by the nucleation loss.

The rate can be written in a form which is useful because it allows an experimental determination of the critical size. Inserting the free energy difference from (8.152), evaluated at $N = N_c$, into the equation gives

$$J \propto S \mathrm{e}^{N_c \ln S/2}. \tag{8.176}$$

The factors left out are independent of S or varies only logarithmically. Taking the double logarithmic derivatives, taking into account the dependence of N_c on S, gives

$$\frac{\mathrm{d} \ln J}{\mathrm{d} \ln S} = N_c + 1. \tag{8.177}$$

This is known as the first nucleation theorem.

The theory presented here, classical nucleation theory, has been tested in great detail due to its enormous importance in atmospheric science. Defects have been found for realistic applications, but the theory remains a good starting point for understanding the process of formation of condensed phase from a supersaturated vapor.

Exercises

8.1 Show that spherical particles with a size independent density have radii that are proportional $N^{1/3}$ and surface areas proportional to $N^{2/3}$. Show also, under the same conditions, that the addition of a layer of atoms (a shell) to the surface increases the radius by $4r_1/3$.

8.2 Calculate the surface tension for a 2d system with the type of calculation that was used to arrive at the 3d result in (8.6).

8.3 Find the total ground state energy of two particles with $N/2$ monomers each and compare it to a single particle with N monomers with identical bulk and surface energy parameters. Calculate the difference and compare it with the bulk contribution, using $B = 2A/3$ and $N = 10^{24}, 10^{10}$ and 10^3. Confirm that negative values of B would lead to spontaneous breakup of particles.

8.4 Discuss the consequences of a surface tension B which is bigger than A, with both A and B size-independent.

8.5 The packing of atoms into dense structures produces a shell structure where one layer is wrapped around the previous as atoms are added. The 'magic numbers', i.e. shell closings for which the structures are particularly stable are, for the special case of the icosahedral shell structure, equal to $N_K = 1, 13, 55, 147, \ldots$ This series can be expressed in closed form, as polynomium in the shell number K, which is an integer that simply numbers the shells. Argue that the highest power of K is 3, and find the coefficients for all the powers of K in the expression, using the series given.

8.6 Show that in the general expression for the number of monomers in a closed, geometric shell; $N_K = aK^3 + bK^2 + cK + d$, with K a positive integer, the coefficients a and b must be integer multiples of $1/3$ and $1/2$, respectively.

8.7 The translational partition function for an ideal gas is sometimes expressed in terms of the thermal de Broglie wavelength, which is defined as the wavelength of a gas molecule when it has the average thermal momentum. This differs from the wavelength used in (8.36).
(a) Find the ratio of the two.
(b) Find the average de Broglie wavelength for an ideal gas and compare with the other two lengths.

8.8 Show that the extremum found for (8.38) has a curvature of

$$\frac{\mathrm{d}^2 P}{\mathrm{d}c_N^2} = -\frac{1}{c_N} + \frac{N^2}{c_1}. \tag{8.178}$$

Use the calculated values of c_N and c_1 in equilibrium with bulk to show that this curvature is negative for most sizes. What are the parameter values leading to a positive curvature? What are the consequences of a positive curvature?

8.9 Why aren't there any rotational degrees of freedom for a single atom? Or are there?

8.10 Give a reason why thermal excitation energies can't grow faster than linear with N when the size increases. If you cannot find a reason, try instead to find an example where it does happen.

8.11 Prove (8.53), i.e. show that the average particle size in an arbitrary collection of particles is equal to the total number of monomers in the system, divided by the number of particles, including particles of size one.

8.12 Show that Π_N in (8.52) is correctly normalized and the probabilities therefore sum to unity;

$$\sum_{N=1}^{\infty} \Pi_N = 1. \tag{8.179}$$

Then show that (8.54) is normalized.

8.13 Show that the equilibrium value of p for a collection of polymers with a bond formation rate of a and a bond breaking rate of b is the one given in the text. Estimate the fluctuations in the values of p as a function of the number of polymers.

8.14 In (8.55), why didn't we use (8.52), to calculate the average size?

8.15 Show, using (8.60), that the solution in (8.54) is stable in the sense that small deviations from the solution will tend to disappear with time.

8.16 What is the mean free path, i.e. the average distance a particle flies between two collisions, for our standard monomers and conditions in Sect. 8.5? Compare that to the average distance between the monomers in the gas, to the thermal de Broglie wavelength and to the monomer radius r_1.

8.17 Show that (8.111) is right. The easiest is probably to tick off the terms in an (N, i) diagram and show that both double sums cover all points exactly once. Next, use the same technique to show that (8.65) and (8.66) are equivalent. Finally, show that the aggregation-only Smoluchowski equation in (8.66) conserves the number of monomers, $\sum_N N c_N$.

8.18 Approximate (8.66) as a set of equations in a continuous particle size variable.

8.19 Show that (8.127) is also solved by an exponential in N, $c_N = \overline{N}^{-2} \exp(-N/\overline{N})$. Use the Smoluchowski equation to calculate the concentration of the monomer as a function of time and show that it differs from $c_1 = \overline{N}^{-2} \exp(-1/\overline{N})$ when $\overline{N} \propto \tau^{6/5}$. For this calculation assume that all $c_{N>1}$ are given by the exponential expression and make the necessary approximations of the a's.

8.20 Find the ratio between the mean size, $\overline{N}$, and the peak value, N_0, of the log-normal distributions that are solutions to the aggregation-only Smoluchowski equation. Use both the value $s = 0.98$ and 1.27. Find also the ratio of the two half values, i.e. the sizes for which the concentrations are half the peak concentration.

8.21 Show that (8.139) follows from (8.116).

8.22 Calculate the free energy of a water drop as a function of radius at $s = 0.5\%$ and $T = 300$ K. Find the critical size and the height of the free energy barrier, ΔF_c. Calculate $\exp(-\Delta F_c/T)$.

8.23 In the calculation of the nucleation rate, the term $c_M/S^M c_M^e$ in (8.168) is postulated to vanish for large enough M. This can be shown by application of (8.166). Use that equation recursively to find the ratios of concentrations c_M/c_M^e in terms of J and the monomer concentrations c_1, c_1^e. Divide by S^M and conclude.

Chapter 9
Molecular Dynamics and Monte Carlo Simulations

This chapter presents the basics of the two most used simulation tools in numerical work on both small and large systems, Molecular Dynamics (MD) and Monte Carlo (MC) simulations.

Both MD and MC simulations require that one has a realistic Hamiltonian available. The easiest situations to handle numerically are those that only involve two-body interactions. Occasionally, one may simplify these interactions more and truncate the range of the potential, thus ignoring the interaction energy for particles that are far apart. In other situations, a problem may require that e.g. excited electronic states are included into the dynamics, or other distinctly quantal effects are considered. Occasionally, such effects can be parametrized in terms of many-body potentials. In any case, for both types of simulations, the results of a simulation will in general not be better than the Hamiltonian used.

Another common feature for MC and MD is the need to establish the system on the computer. In a purely classical simulation this amounts to assign coordinates and possibly momenta to the atoms before a run starts. This corresponds to placing the system in a point, s, in phase space;

$$s(\overline{x}, \overline{p}) \equiv (x_1, x_2, x_3, \ldots, x_{3N}, p_1, p_2, \ldots, p_{3N}), \tag{9.1}$$

for a system of N atoms at the starting time. From this initial configuration the next one is generated by methods that represent a good deal of the art of MD and MC simulations. For both types of simulations, the new configuration depends on the previous one. In MD simulations a new configuration at time $t + \mathrm{d}t$ is generated from the one at time t by propagating the system with Newton's (or Hamilton's) equations of motion. In MC there is no time and a new configuration is generated from the previous with stochastic methods in a process called an iteration. Whereas all coordinates and momenta in MD are calculated in a timestep in MD, there is more freedom to chose the scope of an iteration in MC, and one may prefer to update a single, several or all degrees of freedom at the time, as reasons of efficiency dictate.

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_9

In general MD simulations do not include quantum effects whereas MC calculations may do that. We will only be concerned with classical potentials here, possibly augmented with tractable quantum corrections to the motion of the atoms, such as the Feynman-Hibbs potential in Chap. 2.[1]

A simulation program typically consists of a set of nested loops where a new configuration is generated from the old one and the relevant quantities recorded. As in all other programming activities, it is strongly recommended to keep the code as simple as possible. One may easily save a significant factor on computer time by implementing a clever algorithm, but this saving can easily be eradicated several times over by the extra time it takes to write the code and make it work error-free. Another advantage of a simple program is that it is easy to change when new questions arise.

One cannot expect to be able to start a system in a representative configuration. It is therefore necessary to let the code run a number of steps before one begins to sample quantities for averages. How long this thermalization time should be depends on the system, but it is advised to use a good fraction of the simulation time to thermalize. Alternatively one may store results from different parts of the simulation series separately so one can analyze the thermalization off-line.

In order to get some tools to estimate how well your system is thermalized, let's recapitulate a few statistical facts. The first is the Central Limit Theorem: If you average sufficiently many numbers sampled from the same distribution, no matter what that distribution is, you will end up with a Normal (or Gaussian) distribution for the average. The mean value of the average will be the same as the mean value of the original distribution and the standard deviation of this mean value will be the one for the original distribution divided by $\sqrt{n-1}$, where n is the number of sampling points. The n points need to be stochastically independent, which is achieved by sampling them at phase space points sufficiently far apart. Too closely sampled points cause the reduction to be less than $\sqrt{n-1}$. When sampled too close in phase space, points are said to be correlated. A time correlation function for a quantity A in an MD simulation is defined as (with $\langle A\rangle \equiv \overline{A}$)

$$c(\Delta t) \equiv \left\langle \left(A(t)-\overline{A}\right)\left(A(t+\Delta t)-\overline{A}\right)\right\rangle = \langle A(t)A(t+\Delta t)\rangle - \left\langle A^2\right\rangle, \qquad (9.2)$$

where the averaging is over time. One may of course also define correlations over lengths or other quantities and the averaging may be over a continuous or a discrete variable. One will often observe a decreasing correlation function with Δt, approaching zero monotonically, possibly modified by superposed oscillations with one or several different frequencies. In MC similar correlation functions can be defined where time is replaced by the iteration number, although they will have a different meaning and be more of a diagnostic tool for the programmer.

If the averages you sample drift as you continue to sample, beyond the natural statistical fluctuations, you can be pretty certain that you have not yet sampled a

[1] All interatomic potentials are ultimately quantal by nature, apart from the naked Coulomb interaction. The question is how quantal one treats the motion, given this potential.

sufficiently large part of phase space for the simulation to be representative. There is only one solution to this problem: Keep simulating. Even if the averages seem to behave, however, there is no guarantee that you actually have sampled enough. The simulation can be trapped in a metastable minimum in the potential energy surface from which it will escape only after some time. Metastable states appear when there is a sufficiently high energy or free energy barrier between two local minima in the high dimensional potential energy surface describing the system. The crossover from one of these minima to another can take exponentially long time, and in MD simulations where the total energy is conserved, some crossover events may not happen at all. One partial remedy of this problem is to simulate several replicas of the system with different initial conditions and compare the relevant physical quantities calculated in the different runs. There is no guarantee that you solve the trapping problem this way, but chances are that you can at least spot its presence.

9.1 Basics of Molecular Dynamics Simulations

MD simultaneous are usually used to solve the equations of motion for isolated classical systems, i.e. to find the time development of microcanonical systems, or to trace the motion of a system on a constant energy surface in phase space. Each trajectory in phase space is uniquely determined by the initial conditions once the equations of motion have been fixed, with the exception of chaotic systems for which different trajectories will appear even from practically identical initial conditions. During a run it is essential to check the conservation of energy and other quantities that should be conserved, such as the total momentum and angular momentum of a free particle. An energy which is not conserved indicates a problem with the simulation; Somewhere in the algorithm you calculate something incorrectly.

The core of MD algorithms is the Newtonian equations of motion;

$$\frac{\mathrm{d}\overline{v}_i}{\mathrm{d}t} = -\frac{1}{m_i}\overline{\nabla}_i V(\overline{x}), \tag{9.3}$$

$$\frac{\mathrm{d}\overline{r}_i}{\mathrm{d}t} = \overline{v}_i, \tag{9.4}$$

where $\overline{v}_i$ is the velocity of particle i, and the gradient is taken with respect to the coordinates of this particle. These equations need to be discretized and used with a finite timestep.

As a first attempt one may try to use the approximations:

$$\overline{x}_i(t+\delta t) = \overline{x}_i(t) + \delta t\ \overline{v}_i(t), \tag{9.5}$$

and

$$\overline{v}_i(t+\delta t) = \overline{v}_i(t) - \delta t \frac{1}{m} \overline{\nabla}_i V(\overline{x}(t)), \tag{9.6}$$

that are valid to first order in the time step δt and use them to propagate the system from time t to $t + \delta t$. In principle this algorithm, together with a prescription for V, the initial conditions of $\overline{x}$, $\overline{v}$, and a computer with a compiler is all you need to do MD simulations (see Appendix A for access to free compilers).

The specific time propagation algorithm, however, should be chosen with care, and this is one place where sophistication in the coding pays. The reason is that exponentially growing energies can appear in relatively long simulations if the algorithm is chosen unwisely.

The phenomenon can be understood by considering a one-dimensional harmonic oscillator. With frequency ω and unit mass the equations of motion become, after an expansion to first order in the time increment, δt,

$$\begin{pmatrix} x(t+\delta t) \\ v(t+\delta t) \end{pmatrix} = \begin{pmatrix} 1 & \delta t \\ -\omega^2 \delta t & 1 \end{pmatrix} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \tag{9.7}$$

After N timesteps the phase space coordinates are therefore

$$\begin{pmatrix} x(t+N\delta t) \\ v(t+N\delta t) \end{pmatrix} = \begin{pmatrix} 1 & \delta t \\ -\omega^2 \delta t & 1 \end{pmatrix}^N \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = U \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} U^{-1} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}, \tag{9.8}$$

where U^{-1} is the matrix that diagonalises the matrix in (9.7) that transfers the system from t to $t + \delta t$. The λ's are the eigenvalues of this matrix, and if the modulo of just one of them exceeds unity, the coordinates and velocities will increase unlimited with N. This is the case here; The eigenvalues are $\lambda = 1 \pm \mathrm{i}\omega\delta t$, with moduli squared $|\lambda_i|^2 = 1 + (\omega\delta t)^2$. For most purposes this divergent behavior is not acceptable even if it is second order in δt, partly because the number of timesteps scales with $1/\delta t$. One must find alternatives.

The root of the problem with the simplest approach is that the time derivative was approximated asymmetrically. This is avoided with the so-called Verlet algorithm. It uses the configurations at both time t and the time $t - \delta t$ to determine the configuration at $t + \delta t$. If you expand the position of a particle to third order in $\pm\delta t$ around t and add the two equations you get

$$\overline{x}_i(t+\delta t) + \overline{x}_i(t-\delta t) = 2\overline{x}_i(t) - \delta t^2 \frac{1}{m} \overline{\nabla}_i V(\overline{x}(t)), \tag{9.9}$$

or, expressing the future, $\overline{x}_i(t+\delta t)$, in terms of the past and the present, $\overline{x}_i(t-\delta t)$ and $\overline{x}_i(t)$:

$$\overline{x}_i(t+\delta t) = 2\overline{x}_i(t) - \overline{x}_i(t-\delta t) - \delta t^2 \frac{1}{m} \overline{\nabla}_i V(\overline{x}(t)), \tag{9.10}$$

The appearance of δt^2 instead of δt is already a promising sign. The momentum is not needed in this algorithm, but it may be useful to know for other reasons. It is given by

$$\overline{p}(t) = m\frac{\overline{x}(t+\delta t) - \overline{x}(t-\delta t)}{2\delta t}, \tag{9.11}$$

and can only be calculated after the next step has been taken. The algorithm can be modified to replace coordinates by momenta, and is then known under the name 'velocity-Verlet' algorithm. The terms left out in the Verlet algorithm are on the order of δt^4 instead of δt^2 for (9.5, 9.6), allowing larger timesteps which speeds up simulations considerably.

Figure 9.1 shows a comparison of simulations using both (9.5, 9.6) and (9.10) for a unit mass particle moving in a $V = x^4$ potential, started at $x = 1$ with zero velocity. The data plotted are the calculated distances after 1000 time units. It is glaringly obvious that the Verlet algorithm is so much better than the first order prescription that some effort in implementing it pays of. What is not clear from the figure because of the logarithmic abscissa, is that at the error in the simple algorithm scales linearly with δt (for small values) and it scales with the square for the Verlet algorithm from the smallest all the way to the maximum δt value of 0.2.

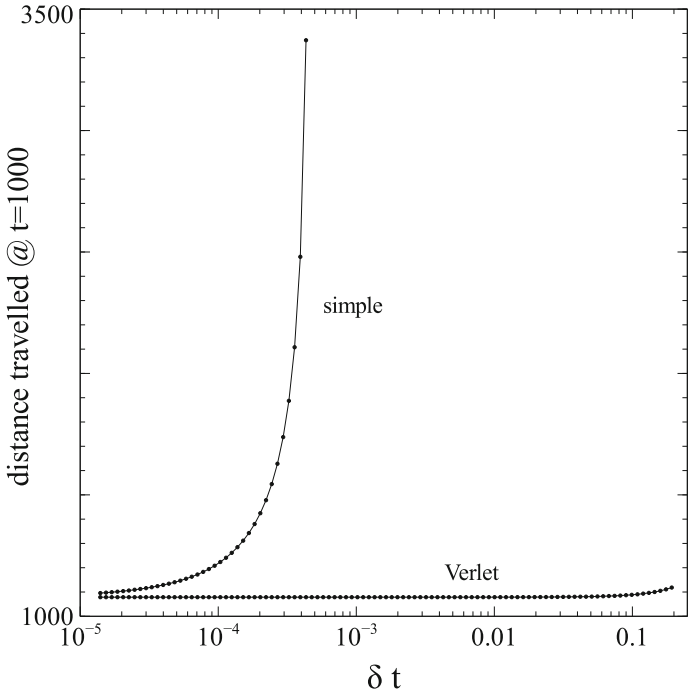


Fig. 9.1 The travel distance vs. the time steps for the simple algorithm in (9.5, 9.6) and the Verlet prescription (9.10)

The values of δt that should be used can be estimated in different ways. The simplest is to simulate with one value of δt, sample what you need, and then repeat the procedure with a smaller δt to see if you get the same result. This is repeated until your averages converge. Although this may sound like brute force, it still only doubles the computation time relative to the optimal timestep if you have a factor of two between consecutive trial values of δt. The idea also demonstrates a sound principle of numerical simulations: Start crude and refine until you see no change in the result. With an ingenious use of this procedure, one can even use such finite resolution simulations to extrapolate values to infinite time resolution, even if they are simulated at fairly coarse resolution. The discipline is called 'Finite Size Scaling' and is used in very time-demanding simulations.

Another trick to speed up simulations is to use the fact that timesteps need not have the same length at all times. The magnitude of the terms left out of (9.10) will in general change with time, and a simple trick is to adjust the timesteps to give the same precision for all steps. This requires an estimate of the neglected terms and it may or may not pay to do this calculation.

One significant drawback of MD simulations is that it is difficult to carry them to long times. When timesteps are of sub-femtosecond magnitude, it is difficult to reach time scales of microseconds, say, for large systems and/or realistic potential energy surfaces that often require some amount of quantum mechanical computation for every point reached in phase space.

9.2 Thermostats in MD Simulations

There is no explicit, externally defined temperature in MD simulations and no time in MC simulations. Occasionally, one wants both in the same simulation, for example when one wants to monitor how chemical reactions proceed in time and at a known temperature. Procedures have been designed to perform the inherently microcanonical MD simulations such that they mimic a canonical ensemble. Such numerical devices are called thermostats.

One popular choice, called the Berendsen thermostat, modifies the velocities at every timestep with a scaling factor f determined as

$$f = \left(1 + \frac{\delta t}{\tau}\left(\frac{T_0}{T} - 1\right)\right)^{1/2} \tag{9.12}$$

where the instantaneous 'temperature' T is defined as $T \equiv 2E_{kin}/s$, where s is the effective number of degrees of freedom, i.e. the number of degrees of freedom with a kinetic energy component, T_0 is the prescribed temperature, δt is the timestep, and τ is a relaxation time which is preset by the programmer. Application of this thermostat will not give the canonical properties rigorously, and the deviations are particularly severe for small systems.

Another thermostat is the so-called Langevin thermostat which is implemented with the addition of two extra terms to the force derived from the potential energy:

$$m\frac{\mathrm{d}^2\overline{x}_i}{\mathrm{d}t^2} = -\overline{\nabla}_i V(\overline{x}) - m\Gamma\frac{\mathrm{d}\overline{x}_i}{\mathrm{d}t} + \overline{W}_i. \tag{9.13}$$

The second term on the right hand side is a friction force with the magnitude determined by the parameter Γ which has dimension of inverse time. The last term compensates for the effects of the friction, which alone would cause all motion to come to a complete standstill, by adding stochastic excitations to the kinetic energy of each particle. The excitations are time dependent, spherically symmetric and uncorrelated both from one particle to the next and for different times. The variance is adjusted to give the right temperature by the prescription

$$\langle\overline{W}_i(t)\overline{W}_j(t')\rangle = \delta_{i,j}\delta(t-t')6m\Gamma T. \tag{9.14}$$

The factor 6 in the last term is the value for three dimensions (each dimension contributes a factor 2). The generation of the random numbers needed for the implementation of the method is the subject of Sect. 9.6.

Finally, the Andersen thermostat should be mentioned. This is the conceptually simplest of them all. It amounts to updating the momenta to the distributions they have in a canonical distribution of the prescribed temperature, i.e. for each momentum component p to pick a value at random from the distribution

$$P(p)\mathrm{d}p \propto \mathrm{e}^{-p^2/2mT}\mathrm{d}p. \tag{9.15}$$

The old momentum is simply discarded and replaced with one drawn from this distribution. You can update all three momenta of an atom at once or separately. The frequency of update is a free parameter in your simulation. Andersen suggests to use an exponential distribution

$$f(t_u) = \nu\mathrm{e}^{-\nu t_u}, \tag{9.16}$$

where t_u is the time since the previous update. You can implement this distribution by updating to a new momentum with the probability $\delta t\nu$ and retain the current value with the complementary values, $1-\delta t\nu$. You need to decide the value of ν, however.

In an equilibrium situation, MD simulations with a good thermostat will converge to canonical properties as simulations grow longer and averages taken over larger parts of phase space. These averages can also be calculated with the Monte Carlo technique described below and the two techniques must give identical results in that limit.

Probing static thermal properties in a MD simulation at constant temperature may seem a little redundant, in particular since the implementation of MD requires more analytical work than the much simpler MC. One advantage is that MD in some cases give much faster convergence because the system is moving dynamically in phase

space. In contrast, the MC technique tend to diffuse a system in phase space, a process which is potentially much slower.

Another, and much more important reason to use thermostated MD is that the method can simulate situations close to thermal equilibrium but where chemical reactions nevertheless occur, i.e. where the components in a system split up, form new compounds or whatever dynamics that changes the macrostate of the system. Such reactions have at least one intrinsic timescale. This means that the choice of the equilibration time to the heat bath becomes important. Too weak coupling between the two systems and the simulated process will effectively be microcanonical; too short time and dynamic effects can be suppressed. The latter may mean, for example, that a localized release of energy will be absorbed immediately into the heat bath, whereas in the real system it can be used to overcome energy barriers before it is dissipated. For a realistic implementation of MD thermostats it therefore important to consider the specific system simulated and its coupling to whatever heat bath is present in the experimental settings.

9.3 Measuring Temperature in MD Simulations

Instead of pre-setting the temperature with a thermostat, one may measure it. A measured MD temperature in simulations without a thermostat will by necessity be the microcanonical temperature. There are several different suggestions in the literature on how to do this. One is to use the average kinetic energy of atoms in classical simulations, viz.

$$\langle E_k \rangle = \frac{s}{2} T, \tag{9.17}$$

as also used above in the prescription of thermostats. We may consider an N-atom, free, non-rotating particle in the center of momentum frame, in which case $s = 3N - 6$. Equation (9.17) is usually justified by reference to equipartition (see Chap. 2) which is a very general statement about classical statistical systems. However, equipartition is derived for canonical ensembles and does not hold microcanonically, so we can't really use (9.17).

Another definition is obtained with the microcanonical temperature given in Chap. 3, where it is associated with the logarithmic derivative of the level density. The level density in the classical limit is calculated as

$$\rho(E) = \frac{1}{h^s} \int \delta(H - E) \prod_i \mathrm{d}x_i \mathrm{d}p_i \tag{9.18}$$

where H is the Hamiltonian, x_i, and p_i are the s coordinates and momenta. We will consider the big class of classical Hamiltonians where the coordinates and momenta enter in a separable way and the kinetic energy can be written as a sum over momenta squared;

$$H = \sum_i \frac{p_i^2}{2m_i} + V(\{x_i\}), \tag{9.19}$$

where $\{x_i\}$ denotes the set of all coordinates. The p_i's and x_i's may be generalized momenta and coordinates in which case also m_i generalizes. This will have no effect on the results derived below.

The separation of the Hamiltonian into kinetic and potential energy means that we can calculate level densities as convolutions of two independent contributions, one from the kinetic energy, ρ_p, and one from the potential energy, ρ_x:

$$\rho(E) = \int_0^E \rho_p(E_k)\rho_x(E - E_k)\mathrm{d}E_k, \tag{9.20}$$

where E is the total excitation energy and E_k is the kinetic energy. The definitions of the level densities ρ_p and ρ_x in (9.20) are unambiguous, apart from a multiplicative factor, because we can redefine the ρ's as

$$\begin{aligned} \rho_x(E - E_k) &\to \frac{1}{c}\rho_x(E - E_k), \\ \rho_p(E_k) &\to c\rho_p(E_k), \end{aligned} \tag{9.21}$$

which leaves the integrand in (9.20) unchanged and therefore gives the same total level density. If we insert the convolution (9.20) into the definition of the temperature we get

$$\begin{aligned} T^{-1} &= \frac{\mathrm{d}\ln(\rho(E))}{\mathrm{d}E} \\ &= \frac{\rho_p(E)\rho_x(0)}{\rho(E)} - \frac{1}{\rho(E)}\int_0^E \rho_p(E_k)\frac{\mathrm{d}\rho_x(E - E_k)}{\mathrm{d}E_k}\mathrm{d}E_k. \end{aligned} \tag{9.22}$$

Partial integration of the last term gives

$$T^{-1} = \frac{1}{\rho(E)}\int_0^E \rho_x(E - E_k)\frac{\mathrm{d}\rho_p(E_k)}{\mathrm{d}E_k}\mathrm{d}E_k, \tag{9.23}$$

provided $\rho_p(0) = 0$, as will be the case when there are more than two independent momenta in the system (see below). This condition will be fulfilled if we simulate one or more atoms in an external potential in 3d space.

The advantage of recasting the expression as in (9.23) is that the level density of the kinetic energy is known and simple. Given the canonical equipartition value in (9.17) it is (see Chap. 2), for s independent momenta, equal to

$$\rho_p(E_k) = aE_k^{s/2-1}. \tag{9.24}$$

From (9.24) it follows that the derivative in the integrand in (9.23) can be expressed as

$$\frac{\mathrm{d}\rho_p(E_k)}{\mathrm{d}E_k} = \frac{s-2}{2E_k}\rho_p(E_k). \tag{9.25}$$

If we insert (9.25) into (9.23) we see that the right hand side in (9.23) is nothing but the microcanonical ensemble average of the reciprocal of the kinetic energy, multiplied by $s/2 - 1$,

$$T^{-1} = \frac{1}{\rho(E)}\int_0^E \frac{s-2}{2E_k}\rho_x(E-E_k)\rho_p(E_k)\mathrm{d}E_k = \frac{s-2}{2}\left\langle\frac{1}{E_k}\right\rangle. \tag{9.26}$$

The microcanonical temperature in a MD simulation can therefore be determined by

$$T = \frac{2}{(s-2)\langle\frac{1}{E_k}\rangle} \qquad \text{(microcanonical)}. \tag{9.27}$$

For free particles, the translational and angular momenta are conserved and s is therefore reduced to $3N - 6$. This number holds for simulations of a particle with zero total angular momentum. Non-zero angular momenta require some care, because the rotational energy, as opposed to angular momentum, is not conserved and can be exchanged with non-rotational energy. The development of this theory is for elsewhere.

The heat capacity is also of interest in microcanonical physics. It is defined as

$$C^{-1} = \frac{\partial T}{\partial E}. \tag{9.28}$$

An expression in terms of average values can be found. A calculation along the same lines as that for T gives

$$C = \left(1 - \frac{(s/2-2)\langle\frac{1}{E_k^2}\rangle}{(s/2-1)\langle\frac{1}{E_k}\rangle^2}\right)^{-1}. \tag{9.29}$$

The condition of applicability here is that there are more than four independent momenta in the system.

We will demonstrate some of these concepts with a MD simulation of a simple but non-trivial one-dimensional system. With positions denoted by $x_i, i = 1, \ldots, N$, it has the energy

$$E = \sum_{i=2}^{N}\frac{p_i^2}{2m} + \sum_{i=1}^{N-1}\left(\left((x_{i+1}-x_i)^2 - x'^2\right)^2 - b(x_{i+1}-x_i)\right). \tag{9.30}$$

This Hamiltonian represents a chain of particles with identical mass m that interact with the potential given by the last sum. The first particle is glued to an infinitely heavy wall and doesn't contribute to the dynamics, and we therefore have $N-1$ particles with non-trivial dynamics. The Nth particle couples only to the $N-1$'th particle. The model is a variation of the toy model for a rubber molecule, which is usually given in discrete form as a two-state model. The present form permits calculations of dynamics as opposed to purely thermodynamic questions.

We can rescale the coordinates and the time to remove three of the four parameters m, a, x' and b. Three is what we can reasonably hope for because we can only scale with the three dimensionful quantities time, mass and length. If we scale the coordinates with x', $x_i \to x_i/x' \equiv x_i'$, and the time as $\tau = t/(ax'^2/m)^{-1/2}$, the scaled energy becomes

$$\frac{E}{ax'4} = \sum_{i=2}^{N} 1/2 \left(\frac{\mathrm{d}x_i'}{\mathrm{d}\tau}\right)^2 + \sum_{i=1}^{N-1} \left(\left((x_{i+1}' - x_i')^2 - 1\right)^2 - \alpha(x_{i+1}' - x_i')\right), \qquad (9.31)$$

where $\alpha \equiv b/ax'^3$. After the scaling is accomplished we do not need the primes on the coordinates and we will drop them in the following.

In order to understand the effect of the remaining (dimensionless) parameter α in this expression, we first set it to zero. Then there are two minima for each term, both with zero potential energies, viz. $x_{i+1} - x_i = \pm 1$, separated by a barrier of height 1. The distances between neighboring atoms, $x_{i+1} - x_i$, are independent and can be optimized separately, and the potential energy minimum therefore has a degeneracy of 2^{N-1}.

Addition of the term with α lifts this degeneracy. A positive/negative value of α causes the minimum energy difference to be located at the positive/negative values of the differences. Because a non-zero α does not introduce any new coupling terms between sets of atoms that are not already coupled, the sets $x_{i+1} - x_i$ remain independent and we can still find this minimum for all distances separately. The minima are two of the roots of

$$4\left((x_{i+1} - x_i)^2 - 1\right)(x_{i+1} - x_i) - \alpha, \qquad (9.32)$$

and are located close to ± 1 when α is not too big. The remaining root gives the top of the barrier which is close to zero. For $\alpha = 0.3$, which is the example used in the following, the absolute minimum is attained when $x_{i+1} - x_i \equiv x_0 \approx 1.04$, and the other is at -0.96 (see Fig. 9.2). The lowest energy state of the entire chain is therefore the stretched configuration with a total length of $1.04(N-1)$. At non-zero excitation energies the chain contracts into a shorter configuration.

An MD simulation of the length vs. microcanonical temperature is shown in Fig. 9.3 for $N = 100$. The timesteps of the calculation were 10^{-3} in scaled units and the total number of timesteps per point was $2 \cdot 10^7$. Also shown is the temperature

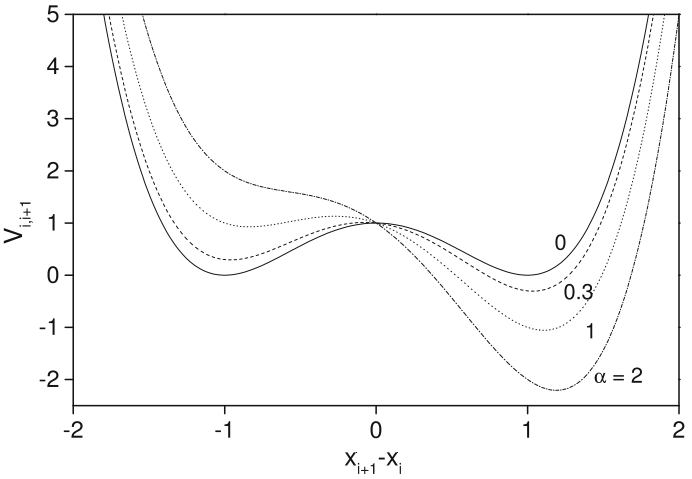


Fig. 9.2 The contribution of each term in (9.31) to the potential energy for different values of the asymmetry parameter α. Both abscissa and ordinate are in units of the scaled variables

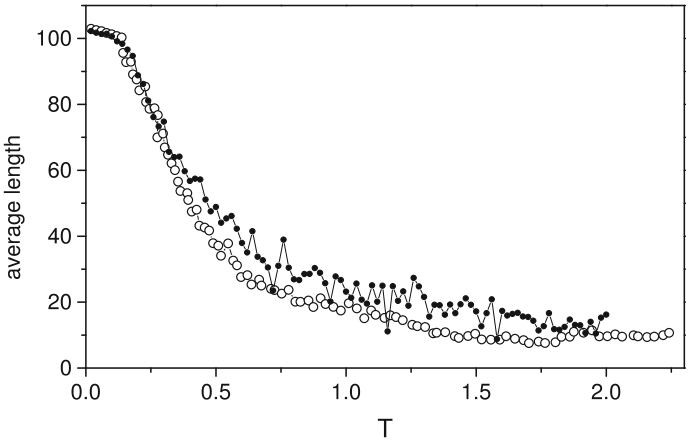


Fig. 9.3 The length of the $N = 100$ atom rubber molecule vs. temperature. The open circles are the results of a MD simulation, the filled circles those of a MC simulation. Note that the fluctuations in the MC points are significantly higher than those of the MD points, although the statistics of the MC simulations is high (10^8 iterations vs. $2 \cdot 10^7$ timesteps, or 20,000 time units)

dependent length calculated in an 10^8 iteration long MC simulation, calculated with methods described in the section on MC calculations below.

The canonical MC and the microcanonical MD caloric curves for the model with 10 atoms is given in Fig. 9.4. The heat capacity of the canonical system is seen to be larger than for the microcanonical system. This behavior persists for even higher temperatures than shown here. It is a finite size effect. For high energies the potential varies approximately as x^4 and the canonical partition function is calculated to $Z_{can} \propto$

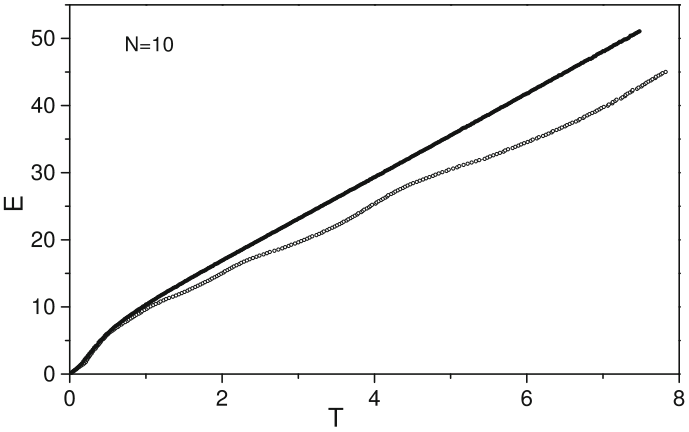


Fig. 9.4 The caloric curves for the rubber molecule described by (9.31) with $\alpha = 0.3$ and composed of 10 atoms. The open circles are the MD values, the filled circles the MC. The energy is the total excitation energy

$\beta^{-3(N-1)/4}$ ($N-1$ is the number of degrees of freedom for an N atom chain). In this expression, $(N-1)/2$ is contributed by the kinetic energy and $(N-1)/4$ from the potential energy part. The canonical heat capacity is therefore $3(N-1)/4$ (see Chap. 2), equal to 6.75 for our $N = 10$ system. The microcanonical caloric curve can also be calculated approximately in the high energy limit. The level density for the potential energy is $\rho_x \propto E^{(N-1)/4-1}$ and for the kinetic energy it is $\rho_p \propto E^{(N-1)/2-1}$. The partitioning of the total energy between the kinetic and the potential part is found to leading order from the maximum of the integrand in the convolution integral, $\rho_x(E-\varepsilon)\rho_p(\varepsilon)$. This gives an average kinetic energy of

$$E_k \approx E\frac{(N-1)/2-1}{3(N-1)/4-2}. \tag{9.33}$$

Approximating the reciprocal of this peak value with the average of the reciprocal kinetic energy, one gets

$$T \approx \frac{2}{N-2}E\frac{(N-1)/2-1}{3(N-1)/4-2}. \tag{9.34}$$

For our $N = 10$ atom system this gives a heat capacity of 5.43, 20% lower than the canonical value. The difference decreases with the size of the system. For $N = 100$ the difference is already very small and the caloric curves for the two different ensembles almost coincide.

Another example of information that can be extracted from a MD simulation is the two correlation functions shown in Fig. 9.5. The main frame is the time correlation

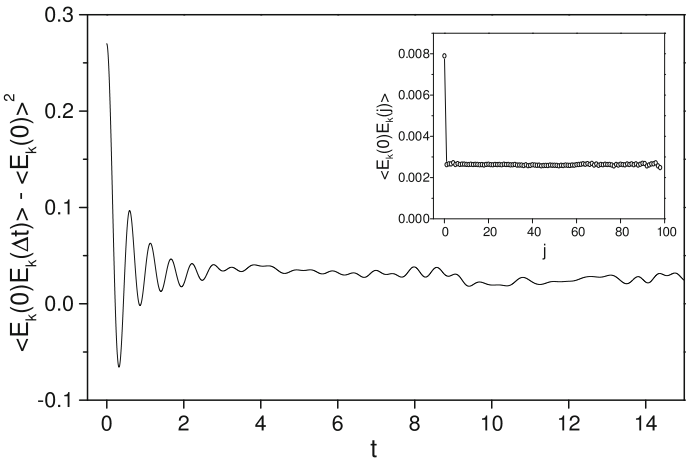


Fig. 9.5 The time correlation function of the kinetic energy for a chain of 100 particles with the rubber molecule Hamiltonian given by (9.31) and the total energy of 9.9 scaled units above the ground state. The small deviation of the long time average from zero is due to an estimated value of $\langle E_k \rangle$ that is slightly off. The inset shows the correlation function for the particles' kinetic energy vs. their separation, for the same total energy. All are equal to the square of the average, showing the absence of correlations, except the distance zero, where the value is the average of the square

function of the total kinetic energy and the inset is the correlation between the kinetic energies of pairs of atoms, plotted as a function of the number of particles between them. The amplitudes of these correlation functions at zero argument are the subject of Exercise 9.8.

Whereas the correlation function in the inset is constant for all non-zero separations, there clearly is a finite correlation in time for this system. This correlation approaches zero in an oscillatory manner with a period equal to 0.54 or a frequency of $2\pi/0.54 = 11.6$, as found from inspection of the figure.

To understand this period we need to examine the dynamics of the system. For small deviations from the equilibrium position of our chain we get for the potential energy to second order in $x_{i+1} - x_i$

$$V \approx V_0 + \sum_{i=1}^{N-1} 4.44(x_{i+1} - x_i - x_0)^2. \tag{9.35}$$

The equations of motion are

$$\begin{aligned} -\ddot{x}_j &= 2 \cdot 4.44(2x_j - x_{j+1} - x_{j-1}), \qquad 1 < j < N, \\ -\ddot{x}_N &= 2 \cdot 4.44(x_N - x_{N-1} - x_0), \qquad j = N. \end{aligned} \tag{9.36}$$

The ground state is $x_j = (j-1)x_0$. An excited state can be written as

$$x_j = b_j(t) + (j-1)x_0, \tag{9.37}$$

with $b_j \neq 0$ for non-zero excitation energy. For an eigenmode (normal mode), which is harmonic in time, the value of b can be written as $b_j = a_j \sin(\omega t + \phi)$, where a_j are coefficients to be determined. Inserting this into (9.36) and cancelling the time dependent factor gives the relations

$$\omega^2 a_j = 2 \cdot 4.44(2a_j - a_{j+1} - a_{j-1}), \qquad 1 < j < N, \tag{9.38}$$

$$\omega^2 a_N = 2 \cdot 4.44(a_N - a_{N-1}), \qquad j = N. \tag{9.39}$$

These are linear difference equations and we can attempt a solution of the form

$$a_j = a \sin(k(j-1)), \tag{9.40}$$

where a is an amplitude and k is a wave vector. Inserting this Ansatz into (9.38, 9.39) and performing some manipulations of trigonometric functions give the two equations

$$\omega^2 = 4 \cdot 4.44 \cdot (1 - \cos k), \qquad 1 < j < N, \tag{9.41}$$

$$\omega^2 = 2 \cdot 4.44 \cdot (\sin(k(N-1)) - \sin(k(N-2))), \qquad j = N. \tag{9.42}$$

Elimination of ω^2 and some more trigonometric exercise provides the set of valid k-values by the relation

$$\sin(kN) = \sin(k(N-1)). \tag{9.43}$$

This has an infinity of solutions for k that differ by multiples of 2π. The $N-1$ smallest positive solutions are

$$k = (2n-1)\frac{\pi}{2N-1}, \qquad n = 1, 2, \ldots, N-1. \tag{9.44}$$

The frequencies of the normal modes are therefore

$$\omega_n^2 = 4 \cdot 4.44 \cdot \left(1 - \cos\left((2n-1)\frac{\pi}{2N-1}\right)\right). \qquad n = 1, 2, \ldots, N-1. \tag{9.45}$$

The highest value, corresponding to $n = N-1$, is $\omega_{N-1} \approx 5.96$, and there are a number of frequencies close to this value because the cosine in (9.45) has zero derivative at that value of the argument.

Naively one would therefore expect that the oscillations in Fig. (9.5) would have the period $2\pi/5.96 = 1.05$. But the period observed in Fig. 9.5 is half of that, 0.54. The reason that the correlation has a period which is a factor of two shorter than the period of vibration that is both the shortest and the dominant in the vibrational spectrum, can be understood by consideration of the analogous quantity for the simpler problem of a single harmonic oscillator. Inspect the phase space of this

system in Fig. 2.4. In the trajectories the kinetic energy (and the potential energy, for that matter) is periodic with a period which is half the period of oscillation because $-p$ gives the same kinetic energy as p. This gives a higher kinetic energy correlation at time differences equal to an integer number of half oscillation periods than for any other time differences, equivalent to a frequency in the correlation function which is twice the oscillation frequency. Basically, the kinetic energy is the square of the momentum, which is a periodic, sine-like function, and the square of a sine function has twice the frequency of the function itself. The more quantitative formulation of this argument is the subject of Exercise 9.7.

9.4 Monte Carlo Simulations

The term 'Monte Carlo' suggests what makes this method work; It is named after the location of the famous casino in the small principality Monaco on the French Riviera. Roulette is gambling on probabilities, and in spite of decades wasted by losing gamblers on pseudo-analysis, there is no way one can predict the result of the next roll of the ball on the basis of the previous results (on an honest device, obviously). That is for all practical purposes completely random.[2]

Even if a process can be considered random in practise, this does not mean that one cannot derive anything meaningful from it. A very early application of pseudo-random processes, predating silicon-based computers with a couple of centuries, was the use of match-throwing on a pattern of parallel lines to get an experimental determination of π. A modern version of the idea, very easily implemented on a computer, is to generate two stochastically independent random numbers x, y in the interval between -1 and 1. Then all points in the square $-1 \leqslant x \leqslant 1$, $-1 \leqslant y \leqslant 1$ are generated with equal probability. The probability that a point is generated inside the circle with unit radius is the area of the circle divided by the total area, or $\pi/4$. Repeating the random number generating process sufficiently many times, one gets an estimate of the value of $\pi/4$ as the fraction of points inside the unit circle. Figure 9.6 shows an example of such a calculation.

Although the calculation of π is not the most convincing reason to learn how to run MC simulations, it is clear from the example that even random events, if disciplined, have the ability to provide non-trivial information.

In MC simulations of physical systems, the transfer from one configuration to the next is stochastic. If one labels the microstates of a system, i.e. effectively the points in phase space, with a single integer, an iteration will take the system into another microstate, selected from a range of other states with a prescribed probability which sum up to unity for all the possible new states. The next iteration will then take each

[2] This does not imply that the dynamics is completely random. A group of physics students in the US decided to try to predict the winning numbers of roulette, based on measured trajectories of the ball and on-line computation of the dynamics. Their story is told in Thomas A. Bass 'The Newtonian Casino' (Penguin Press Science).

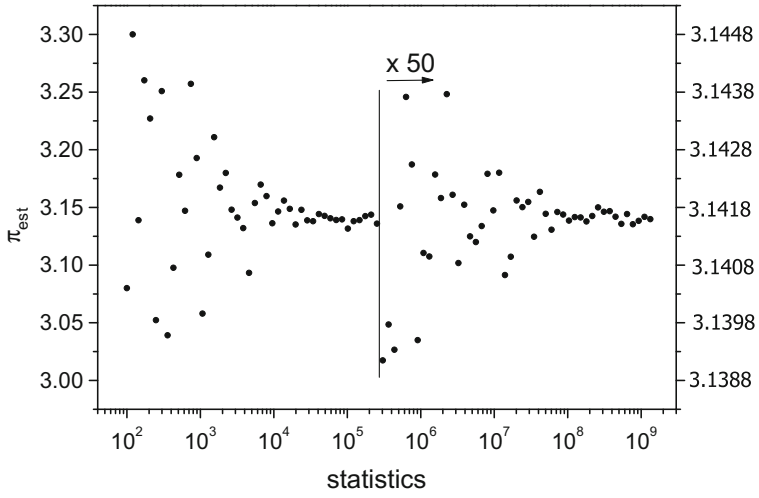


Fig. 9.6 The estimated value of π based on the fraction of randomly generated points that fall inside a circle of unit radius in a 2×2 square. For statistics above 250,000 hits the difference to π is multiplied with 50, as indicated in the figure

of these microstates to one of a number of other microstates with their own prescribed set of probabilities etc.

For a collection of simulations, the state of a system can therefore be considered a vector, $\overline{s}$, of populations of the microstates, with real, non-negative elements that sum to unity, and the transition probabilities as forming a matrix, $\overline{\overline{M}}$, with rows and columns indexed by the initial and the final microstates. These are the characteristics of what is called a Markov chain or Markov process. Basically, the move into the next state is decided from the present state of the system only, with no memory of previous history.

After a sufficient number of iterations, the state of the system is that of equilibrium, $\overline{s}_e$, for which

$$\overline{s}_e = \overline{\overline{M}}\overline{s}_e. \tag{9.46}$$

This result is independent of the choice of the matrix elements as long as they conform to some general requirements, of which the detailed balance is the most important. This ensures that $\overline{s}_e$ in (9.46) exists. It only represents the true equilibrium distributions if the dynamics of the system, as represented by $\overline{\overline{M}}$, is ergodic, i.e. every state in the system is ultimately visited. Another way of stating the ergodicity criterion is that precisely one eigenvalue of $\overline{\overline{M}}$ is unity and all other other are less than unity. The unit eigenvalue corresponds to the state $\overline{s}_e$. Two or more unit eigenvalues means that there is more than one state that fulfil the criterion in (9.46). That must mean that there are two or more regions of phase space that do not mix, which is another way of saying that the system is non-ergodic.

In practice, simulations are not performed by generating the matrix $\overline{\overline{M}}$; the space is far too big to make this practical, and Markov processes are mainly a conceptual help for us. In practice, individual trajectories through phase space are generated on a case-by-case basis.

One way to organize this move through phase achieve which is guaranteed to give ergodicity is to have a non-zero probability to reach any state from any other state. This clearly does the job, but is equally clearly an unpractical protocol to implement. Less can do it, though, and for the Markov chains we implicitly construct with the prescriptions given below, every microstate can be reached from any other initial state by a finite number of steps through intermediate microstates.

Note though, that although non-ergodic dynamics is absent asymptotically, it may well be present for real simulations with their finite number of iterations. Free energy barriers will be crossed with a probability that varies with the free energy, F, as $\exp(-\beta F)$. F may be so high that the probabilities to cross them will be vanishingly small.

The absence of history, physically manifested in the absence of a momentum which is the pointer to the near future in the simulations, severely restricts the amount of detail on the dynamics one can extract from these simulations. MC simulations can, on the other hand, be used much more generally than MD simulations, and they are routinely applied in a variety of connections that have nothing to do with the inner structure of particles, from the calculation of integrals over assessments of particle detector efficiencies to calculations of the fall-out from polluting smokestacks.

The basis of a MC simulation of the thermal properties of a particle or molecule is, apart from the Hamiltonian, a random number generator. The aim is to generate a canonical probability distribution stochastically:

$$P(\{x_i\},\{p_i\})\prod_i \mathrm{d}x_i\mathrm{d}p_i \propto \rho(E(\{x_i\},\{p_i\}))\mathrm{e}^{-\beta E(\{x_i\},\{p_i\})}\prod_i \mathrm{d}x_i\mathrm{d}p_i, \tag{9.47}$$

where ρ is the level density of the system and $\beta \equiv 1/T$. If the momenta and the coordinates do not mix in the Hamiltonian, the distribution factorizes:

$$P(\{x_i\},\{p_i\})\prod_i \mathrm{d}x_i\mathrm{d}p_i \propto \rho_x(V(\{x_i\}))\times \mathrm{e}^{-\beta V(\{x_i\})}\prod_i \mathrm{d}x_i \times \mathrm{e}^{-\Sigma_i \frac{\overline{p}_i^2}{2m_i T}}\prod_i \mathrm{d}p_i, \tag{9.48}$$

where the last factor is the product of the Gaussian distributions of all the momenta. They can be integrated out and yield a factor that depends on T and m_i but not on the coordinates. Hence it can simply be ignored in the simulation. If we need the momentum or kinetic energy distributions, we simply write them down afterwards.

This takes care of half the dimensions of phase space. The other half is the time consuming and interesting part of the potential energy. It is explored by prescribing how one moves from one point on the potential energy surface to the next. A simple and generally applicable choice, obviously not unique, is to update a single coordinate or the coordinates of a single atom at the time. At a given configuration, you pick out

the coordinate that should potentially be changed, x_i say. First you chose whether you go left or right, $x_i \to x_{i,new} = x_i - \delta x_i$ or $x_i \to x_{i,new} = x_i + \delta x_i$. The value of δx_i is arbitrary but must be the same for both directions.[3] It will be shown below how to generate random numbers of this kind.

If all these suggested updates were accepted, one would end up with a distribution that would be completely flat over the whole phase space and that had nothing to do with the potential energy. To get a distribution weighted with the Boltzmann factor $\mathrm{e}^{-\beta V}$, it is necessary to first calculate the energy of the trial configuration, $V_n \equiv V(x_1, x_2, \ldots, x_j^{(n)}, \ldots)$ (n for 'new') and compare it with the energy of the initial state, $V_i \equiv V(x_1, x_2, \ldots, x_j^{(i)}, \ldots)$ (i for 'initial'). In equilibrium we want the two configurations to be populated with a relative weight given by

$$P(V_i) \propto \frac{\mathrm{e}^{-\beta V_i}}{\mathrm{e}^{-\beta V_i} + \mathrm{e}^{-\beta V_n}}, \tag{9.49}$$

for the initial configuration, and

$$P(V_n) \propto \frac{\mathrm{e}^{-\beta V_n}}{\mathrm{e}^{-\beta V_i} + \mathrm{e}^{-\beta V_n}} \tag{9.50}$$

for the candidate configuration. Detailed balance requires that we accept the candidate with a probability $p_{i\to n}$ which is related to the probability for the inverse choice, $p_{n\to i}$, according to the rule

$$p_{i\to n} \frac{\mathrm{e}^{-\beta V_i}}{\mathrm{e}^{-\beta V_i} + \mathrm{e}^{-\beta V_n}} = p_{n\to i} \frac{\mathrm{e}^{-\beta V_n}}{\mathrm{e}^{-\beta V_i} + \mathrm{e}^{-\beta V_n}}. \tag{9.51}$$

If we shave off the common factor we have

$$p_{i\to n} \mathrm{e}^{-\beta V_i} = p_{n\to i} \mathrm{e}^{-\beta V_n}. \tag{9.52}$$

This determines only the ratio of the two probabilities. We will now make the choice that all moves to lower energy are accepted with unit probability, i.e. $p_{i\to n} = 1$. If $V_n - V_i \leqslant 0$, the new configuration is thus accepted unconditionally. This optimizes the efficiency of this part of the procedure: Unity is the maximum possible value. Less would correspond to holding the hand brakes while you kept pushing the pedals of your bicycle.

To find the probability when the candidate energy is higher than the starting configuration, $V_n - V_i \geqslant 0$, you just solve (9.52) for $p_{n\to i}$ with $p_{i\to n} = 1$. Interchanging the subscripts i, n this gives:

$$p_{i\to n} = p_{n\to i} \frac{\mathrm{e}^{-\beta V_i}}{\mathrm{e}^{-\beta V_n}} = \mathrm{e}^{-\beta(V_i - V_n)}. \tag{9.53}$$

[3] Strictly speaking it does not have to be equal for the two directions, but to avoid unnecessary complications you better adopt this policy of symmetric choice from the start.

The decision on whether or not to move to the new configuration with the probability given by (9.53) is made by generating a evenly distributed random number. If it is less than $e^{-\beta(V_n - V_i)}$ the move is accepted. Otherwise it is rejected and one proceeds to the next coordinate or iteration.

For averaging statistics, which this is all about, it is important to count the state in the averages after an attempted move, irrespective of whether the attempted change of the configuration was successful or not.

This choice of transition probabilities is called the Metropolis algorithm after the name of the first author of the article where the method was introduced. It should be clear from the above that the Metropolis algorithm is by no means the only one possible, but it has done quite well since the introduction. A very good reason for this is the generality and the simplicity of the algorithm which can be implemented in very few lines of code.

By the nature of the procedure, a fraction of the attempted moves will be rejected. The precise number depends on the values of the δx_is one uses. For MC simulations of molecular systems these parameters are the single most important place to pay attention to computer time economy of code execution. First of all, it does not pay to make δx_i too large because then most attempts will be rejected and you tend to be frozen in whatever configuration you happen to be in. Too small a value is also a problem, because then most attempts will be accepted but each step will be too small to move the system much.

A suitable compromise is an acceptance rate of 50% for a move. This success rate can be obtained by adjusting δx_i with a small factor that depends on the fraction of attempted updates that were accepted. If a successful update results in an increase of δx_i by a factor $(1 + \alpha)$, and a failure in an decrease $1/(1 + \beta)$, after M_a successful updates out of a total number of M attempts, one has changed δx_i as

$$\delta x_i(M) = \delta x_i(0)\,(1+\alpha)^{M_a}\,(1+\beta)^{-(M-M_a)}\,. \tag{9.54}$$

If the value of δx_i is equilibrated, $\delta x_i(M) \approx \delta x_i(0)$ holds and for small α, β the success-to-failure rate is then α/β, corresponding to a total success rate of $\alpha/(\alpha + \beta)$.

In principle, adjustment of the step size can't be done on-line when you do the simulation because it will violate the detailed balance condition, but often one does it anyway, and if $\alpha, \beta \ll 1$ it is only a slightly dirty procedure. If the modification of δx_i is not done after each step but instead based on an average of a large number of attempted moves, the problem is also reduced.

Not all simulations are of a nature where the step size can be adjusted continuously. When dealing with discrete systems it is often a good policy to make the attempted moves as small as possible. This is not only an advantage when writing the code, but it will often be essential for moving the degrees of freedom around at all. If one has, in one jump, to overcome an activation energy which is five times the temperature, one needs an average of $e^5 = 148$ attempts before success. This is a lot but the simulation is still feasible. Trying to move two at the same time will then require typically e^{10} attempts before success, and a success rate of $e^{-10} \approx 5 \cdot 10^{-5}$ is equivalent to failure.

9.5 Microcanonical MC

MC simulations are usually canonical, because points in phase space are accepted with the Boltzmann factor as the weight factor. MD simulations are born microcanonical, although they can be made (quasi-)canonical with the thermostats mentioned. If one is interested in microcanonical static properties, it is however possible to use the efficiency of MC sampling. This is accomplished by replacing updates of the potential energy involving energy exchanged with an infinitely large heat bath of temperature T, with exchange of energy between the potential and the kinetic energies of the particle itself. The addition of the kinetic energy as a heat bath is quite natural in the light of the systems we are simulating, and we can even apply the equations derived for T and C in MD microcanonical simulations directly to find the microcanonical caloric curves.

The only difference between orthodox MC and microcanonical MC is the coordinate update. In the Metropolis algorithm the moves are accepted unconditionally if the potential energy decreases, as in canonical MC. A trial configuration with higher potential energy is accepted with probability

$$p_{i\to n} = \frac{\rho_p(E - V_n)}{\rho_p(E - V_i)} = \left(\frac{E - V_n}{E - V_i}\right)^{3N/2-1} . \tag{9.55}$$

which replaces (9.53). Numbers with these probabilities can be calculated with the methods given in the next section.

9.6 Random Number Generation

Generating random numbers is an essential part of MC. Most compilers come equipped with a random number generator which will give you a prototype random number, distributed evenly between zero and one, called a univariate random number. The generator may need to be initialized by a seed number. This may be done semi-automatically by calling a computer clock, or by manually entering a number. Once the generator is initialized, it can be called a huge number of times to give new random numbers. Ultimately, however, it will return to the starting point and repeat the sequence. That is, if it is of the conventional, algorithmic type. Random number generators based on quantum randomness in the hardware will not suffer from this problem.

The period of a generator depends on the compiler and the random number seed, if relevant, and can be long. And sometimes it needs to be long. A period of, for example 10^{10} sounds like a lot, but it is actually possible to run codes that will return to square one and repeat. Happily, in MC simulations of particles with their many degrees of freedom, it is rarely a problem because the system is in a completely different configuration after running through one period of the random number, or

you are updating another variable than the first time around, but it is still worth keeping in mind.

The generator is supposed to provide numbers without any correlation. Before you use one of those black boxes, be advised to do a simple check of the mean, standard deviation and correlation function. The correlation function for the sequence of univariate random numbers $x_1, x_2, x_3, \ldots$,

$$C_j = \langle x_{k+j} x_k \rangle - \langle x \rangle^2 = \langle x_{k+j} x_k \rangle - 1/4. \tag{9.56}$$

should be zero for $j \neq 0$. Another simple test is to add the values up as

$$\text{sum} \equiv \sum_{i=1}^{n} (x_i - 1/2). \tag{9.57}$$

Figure 9.7 shows such a sum. By plotting this series, a period will appear clearly and indicate that the generator has reached the end of its length and has restarted, a behavior that can otherwise be a little tricky to identify for long series. The example here shows no obvious periodicity, indicating a period exceeding 10^{12}. The curve is calculated with a laptop and illustrates the tremendous computational power available with even such small devices if your programming language is not too high level and bogged down by too many features.

When it has been established that the standard, evenly distributed random numbers can be trusted, one can perform a long range of simulations, based on that alone. A first application of a univariate random number may be the generation of symmetrically distributed values for use with the Metropolis algorithm. The question of whether one

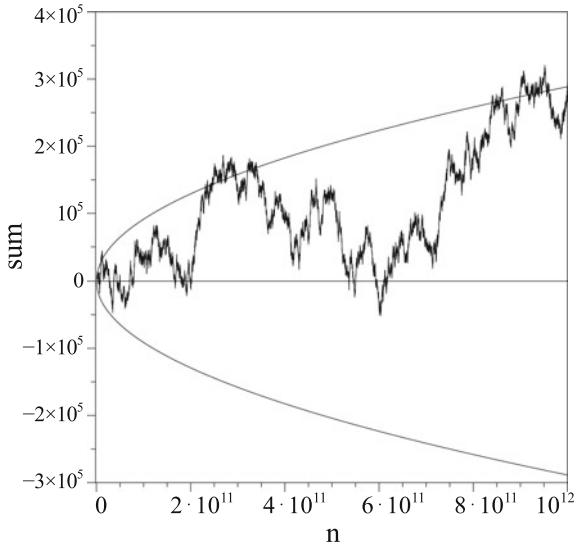


Fig. 9.7 A sum of consecutive random numbers offset with −1/2. The square roots are the calculated standard deviations on the sum, $\sigma = \sqrt{n/12}$

goes left or right in this algorithm is most easily decided by generating a univariate random number, u, and decide for an attempted move to the right if the number exceeds 0.5 and to the left otherwise. Alternatively, one can translate u to the interval $[-1;1]$ and multiply by δx_i, $x_{j,n} = x_{j,i} + \delta x_j(2u - 1)$, and use this as an attempted move.

There are several different methods to generate random numbers drawn from more complicated distributions and one needs to know some of these to make use of the full power of the MC technique. The following is a short overview of some useful techniques.

The first we will mention is applicable to a distribution defined on a finite interval and without divergences. Then the maximum or at least an upper limit of the distribution can be found. With this knowledge, the prescription is to generate an evenly distributed random number, u_1, and shift it with the linear transformation that maps [0;1] on the relevant interval. Generate a second number, u_2, and multiply it with the maximum of the probability distribution. If this product is less than the value for the probability distribution at u_1, $P(u_1)$, u_1 is accepted as coming from the distribution, otherwise it is rejected. The idea behind the method is not very different from the procedure used to determine the value of π, once you think about it. It is essential for this method to be efficient that the product of the interval length and the maximum of the (normalized) probability distribution is not extremely large compared to one. Otherwise one gets too many rejects. In an obvious extension of the method, it can also be used to generate multi-dimensional variables.

This method, 'acceptance-rejection' as it is sometimes called, is the simplest possible and is often a good alternative when the direct inversion described below is not possible. One application of the method, for $\sin^2 x$ on the interval $x \in [0; 2\pi]$, is shown in Fig. 9.9, together with examples of other methods.

The second method is much more elegant, more economical with computer power, but unfortunately not very general. One integrates the probability distribution and inverts it. If a random number, evenly distributed between 0 and 1, is used as the argument of this inverted function, the functional value will be a number sampled randomly from the original distribution. Formally, for the integral $F(x)$ of the probability distribution $P(x)$,

$$F(x) = \int_{-\infty}^{x} P(x')\mathrm{d}x', \tag{9.58}$$

the variable x generated with the univariate random number u as

$$x = F^{-1}(u) \tag{9.59}$$

will be distributed with the probability distribution $P(x)$. Figure 9.8 illustrates (9.58, 9.59).

The exponential decay provides a simple case for this method. With the time constant τ, the normalized function is $\tau^{-1}\exp(-t/\tau)$. It integrates to

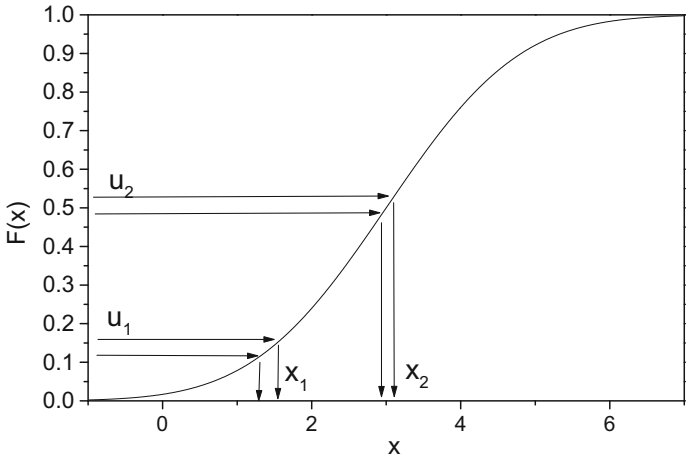


Fig. 9.8 The logic behind the inversion of the integral of the probability density. The two univariate random variables u_1 and u_2 are used to find the two different random values x_1 and x_2. The two small intervals indicated have identical widths around the u's but clearly not on their projections on the x-axis, indicating a higher density at the places with a high slope of F, as expected

$$\int_0^t \tau^{-1} \exp(-t'/\tau) \mathrm{d}t' = 1 - \mathrm{e}^{-t/\tau}. \tag{9.60}$$

Equating this to the univariate random number u and solving for t gives

$$t = -\tau \ln(1-u) = -\tau \ln(u). \tag{9.61}$$

The last equality does not imply that $u = 1/2$ but is a manipulation which is allowed because $1 - u$ has the same distribution as u. There is no loss with this procedure; all generated numbers are accepted.

Another important case where number can be generated without loss is the Gaussian. For this, one initially generates numbers from a slightly different distribution, $P \propto r\mathrm{e}^{-r^2}$, which is the distribution of the distances of points from Origo in a two-dimensional Gaussian distribution. It can easily be integrated and inverted. This gives a value of r sampled from the two-dimensional distribution. It is then projected on the x-axis, using a second random number, u_2, to generate the random phase that goes into the projection cosine. Numbers generated this way will be normally distributed with mean zero and variance one. To get a non-zero mean, $\langle x \rangle$, and a standard deviation, σ, numbers are generated and scaled as

$$\langle x \rangle + \sigma \sqrt{-2 \ln(u_1)} \cos(2\pi u_2). \tag{9.62}$$

The scaling and shift obviously works with any distribution, not just a Gaussian.

The inversion method can also be used if a distribution can be written in a scaled form used for all parameter values, temperatures for example, and therefore only needs to be generated once. Then the integral of the distribution can be calculated and inverted numerically to be stored in an array.

The methods discussed above can be combined. It may be convenient to extract one part of a distribution if it can be generated, for example, by the inversion method and the rest of the distribution ends up in a form that can be generated more efficiently. Then the random number generated by the first method is fed into the second method applied to the modified distribution.

Often one encounters the need to convert a distribution of one variable into another, because the converted distribution is easier to generate than the original. In one dimension, the general expression for the distribution of the variable $y(x)$ is, when the distribution for x is given as $P(x)$, equal to:

$$P(y) = P(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|, \tag{9.63}$$

a formula that holds whenever you transform distributions, not just for the purpose of random number generation. In higher dimensions the derivative becomes a matrix known as the Jacobian and the absolute value of the derivative is then replaced by the absolute value of the determinant of this matrix.

Occasionally it is possible to update a degree of freedom by generating the whole distribution for this degree of freedom directly, even if it is coupled to other degrees of freedom. As an illustration, consider a simulation of a system with a potential energy that can be written in the form

$$V = f(\overline{x})y + g(\overline{x})y^2, \tag{9.64}$$

where f and g are some functions of all coordinates except the y we will update. When y alone is updated, f and g can be considered constants and we can write the distribution of y as

$$P(y)\mathrm{d}y \propto \mathrm{e}^{-\beta(fy+gy^2)} \propto \mathrm{e}^{-\beta g(y+\frac{f}{2g})^2}, \tag{9.65}$$

where the constants of proportionality do not depend on y. The problem is thus reduced to generating normally distributed numbers, which can be done very rapidly with the procedure described in (9.62).

Finally one should mention a brute force method that is easy to apply, viz. the random walk method. From a chosen start-value, a new value is attempted with a symmetric distribution, for example as $y_n = y_i + \delta y(2u - 1)$, where u is a univariate number and δy is optimally similar in magnitude to the width of the distribution, if known. The trial value is then tested and accepted if $f(y_n) \geqslant f(y_i)u_2$ where u_2 is another and uncorrelated univariate number. The method is thus of the same trial-and-error type as the Metropolis method.

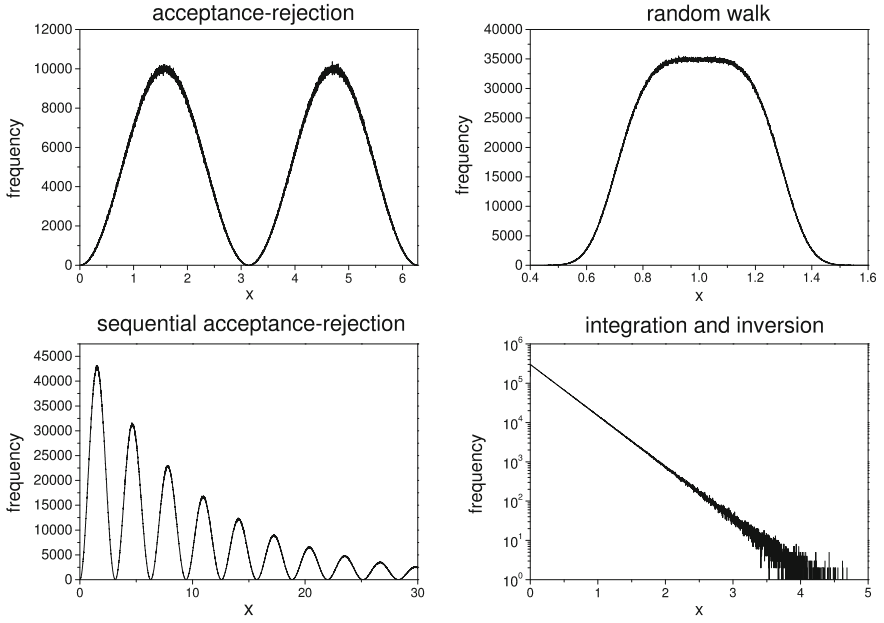


Fig. 9.9 Random numbers generated with the methods described in the text. The attempt statistics is in all cases 10^8. The target functions are, clockwise from top left: $\sin^2(u)$, $\exp(-u^4/u_0^4)$, a simple exponential, and finally $\sin^2(u)\exp(-u/u_0)$ generated with the sequential direct inversion and acceptance-rejection method. For this, first numbers u from the exponential distribution are generated. These are then accepted if a new randomly generated univariate number is less than $\sin^2(u)$

9.7 Optimization: Simulated Annealing

Simulated annealing is a method to locate local minima on complicated potential surfaces. These potential surfaces are not necessarily models of particles but can be anything which has a potential or more generally a cost associated with it. The method was in fact conceived as a tool for optimizing the placement of electronic components on a circuit board. That is an example of what is called an NP-complete problem, which means that the number of possible solutions for the optimization problem grows faster than any polynomial of the number of constituents. This rules out a brute force exploration of configuration space, consisting in calculating all possible configurations to test for cost.

Instead one uses the following recipe: Associate an energy with every configuration and shift the components around with the Metropolis algorithm. Initially the temperature in the Metropolis algorithm is set very high with the consequence that all possible configurations are generated, mainly ones of quite poor quality, i.e. high energy, because that's where the electronics industry equivalent to entropy is found. As the temperature is slowly lowered, the system tends to move around closer to

configurations with well-defined minima, and finally, when the temperature is so close to zero motion is effectively frozen out, the configuration can be taken seriously as a candidate for the best placement of the components. The method works surprisingly well, although one can rarely, if ever, be sure to have discovered the best solution of a realistic problem, and for large enough numbers of components, one most likely has not.

In case of a collection of atoms, the cost function will be the potential energy surface, and minimizing this corresponds to finding the ground state of the lowest energy isomer. Isomer is the name for a configuration which is stable against small deviations from the configuration, equivalent to an atomic configuration for which the Hessian matrix has only positive eigenvalues, and hence for which all vibrational frequencies are real.

The procedure is illustrated in Fig. 9.10, where the ground state of the rubber molecule with the energy given by (9.31) is attempted located. The ground state for this system is the state where $x(i) = x(i-1) + 1.036$ (for $\alpha = 0.3$), and the simulation provides a demonstration of the efficiency of the algorithm. For this particular simulation, 95 out of 99 atomic distances end up in the correct low energy position. The initial temperature is 2, the final temperature is 10^{-6}, and the decrease was exponential, with a factor of $1 - 10^{-3}$ per 50,000 iterations. Snapshots of the value of $x(2)$ is shown in Fig. 9.11. Because $x(1) = 0$, this also represents a sampling at varying temperatures of the thermal distribution of positions in the potential shown in Fig. 9.2.

Another illustration is a two-dimensional model of neon absorption on a graphene sheet. Graphene is a hexagonal, two-dimensional lattice of carbon atoms, and a neon

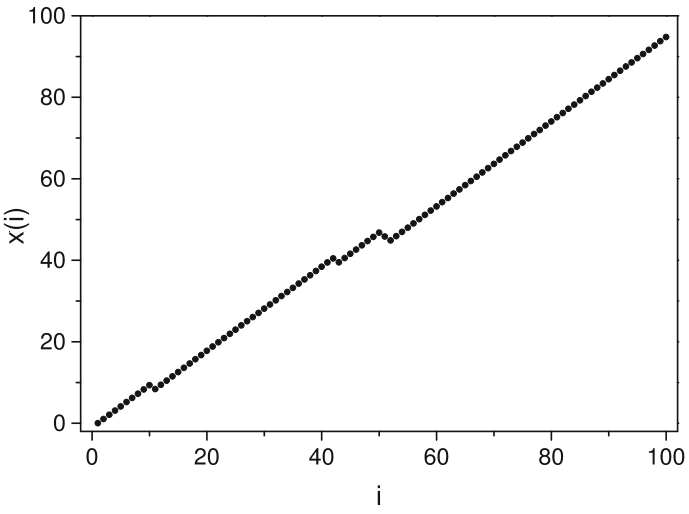


Fig. 9.10 The final configuration at the end of a simulated annealing run of a rubber molecule with 99 movable atoms

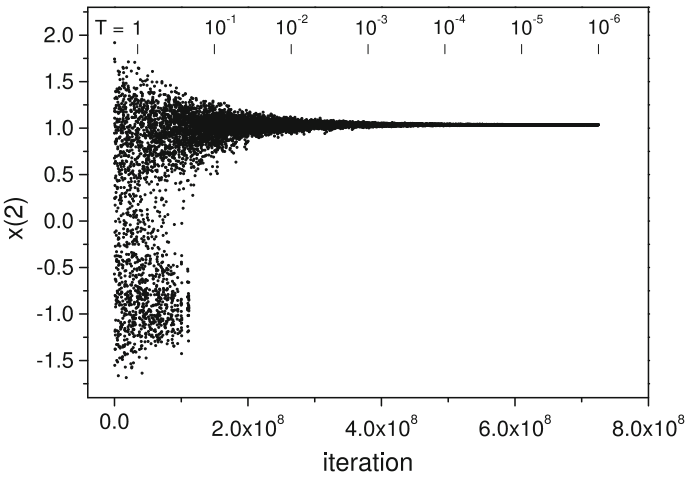


Fig. 9.11 The distance between atom 1 (fixed at $x(1) = 0$) and atom 2 in the rubber molecule as a function of the MC iteration number. The data are from the same simulation as the data shown in Fig. 9.10. The position was sampled every 50,000 iterations. The temperatures are shown on top

atom can be accommodated in a potential depression in each hexagon. The difference between the top and bottom of this potential well is 30 K, with the high energy attained with the neon atom on top of a carbon atom, lower values when the neon is balancing on the bonds connecting the carbon atoms and the minimum value when the atom is in the position equally far from all the C atoms in a hexagon. In addition to interactions with the carbon atoms, the neon atoms interact with each other via the Lennard-Jones potential.

The total potential energy is

$$V = c\left(\sum_{j=1}^{3}\cos(\overline{x}\cdot\overline{k}_j) - \sum_{j=4}^{6}\frac{2}{3}\cos(\overline{x}\cdot\overline{k}_j) + E_0\right) + \sum_{i>j}4\varepsilon\left(\left(\frac{\sigma}{r_{i,j}}\right)^{12} - \left(\frac{\sigma}{r_{i,j}}\right)^{6}\right). \tag{9.66}$$

The terms contained in the first bracket is the interaction with the substrate. The parameters are $c = 1.875$ K, the (dimensionless) energy offset is $E_0 = 9$ which makes the minimum energy for a single neon atom zero and the maximum 30 K. The wavenumbers for the modulation of the surface potential are $k_j = 2\pi\cos(2\pi(j-1)/3)/s$, for $j \leqslant 3$, and half these values for $4 \leqslant j \leqslant 6$. The length parameter s that appear in the graphene part has the value 1.42/1.298 Å. This all ensures an approximately hexagonal shape with the correct graphene lattice constant of the substrate and the correct difference between maximum and minimum in the potential energy landscape. The graphene-neon potential has its maxima at the midpoints of

the C-C bonds and not at the vertices, but this is a minor error in this connection, if it even is an error. A feature, which is not immediately obvious from the functional form of the potential, is that the potential energy for a single atom in one of the graphene minima is not quadratic in the displacement from the minimum, as is usually the case, but has a quartic dependence.

The second term in the potential energy is the Lennard-Jones (LJ) interaction, where $\sigma = 2.9$ Å is the LJ length scale for neon, and $\varepsilon = 36.7$ K is the neon-neon interaction energy.

The potential minima on the lattice are separated by $1.42\sqrt{3}$ Å $= 2.46$ Å, whereas the optimal distance between two neon atoms in the absence of a substrate is $2^{1/6}\sigma = 3.3$ Å. This difference will cause a mismatch in the optimal lattice structure for the Lennard-Jones neon-neon interaction and that of the neon-graphene interaction. We expect the LJ structure to dominate because the number of bonds make that binding stronger, but it is not obvious how the resulting lattices will be oriented relative to each other.

The problem was simulated with 50 neon atoms on a lattice 35×35 Å^2, corresponding to 238 lattice sites of which 35 were located at the boundary. The simulation started at a temperature $T = 100$ K, and was lowered according to the algorithm $T_{n+1} = T_n \exp(-1/(C + 10))$, where C is the heat capacity measured as $(\langle E^2\rangle - \langle E\rangle^2)/T^2$. This non-exponential cooling is designed to make the cooling rate slower when the heat capacity is high and the system explores new configurations with a higher frequency than at configurations with low heat capacities. The kinetic energy was not simulated and a safety term of 10 is added to the heat capacity to prevent incidences of too fast cooling rates. The statistics was $5 \cdot 10^4$ iterations for each temperature, of which 10^4 were used for thermalization.

The initial ($T = 100$ K) coordinates of the particles were chosen randomly within the confining volume. One iteration consisted in choosing new trial coordinates for each atom separately with the Metropolis algorithm, with the suggested new coordinates given by $x_n = x_i + (2u - 1)\Delta x$, with u a univariate variable and Δx the maximum single step displacement. The amplitude of the trial steps, Δx was adjusted after each attempt according to whether or not it was successful, by the factor 1.002 or 1/1.002. As discussed above, this adjustment gives an average acceptance rate of 0.5. The value of Δx was atom-specific, with the motivation that movement of atoms residing on the border of a cluster are much less constrained than movements of atoms located in the interior.

Figure 9.12 shows the energy measured during the annealing run. Little happens with the potential energy before the temperature reaches values around the barrier height of 30 K. From 30 to 10 K, the atoms condense into a lattice structure which is mainly determined by the LJ interaction between the atoms. A condensation from a 2d gas into a 2d cluster is indicated by the increased heat capacity in this interval, from values close to zero to almost 300 at the maximum. After the lattice structure is established, the heat capacity reduces to 50, which is the value for the potential energy contribution for 50 harmonic oscillators in two dimensions. Isolated atoms in this particular lattice would have only half that heat capacity because of the x^4 potential in the single atom minima.

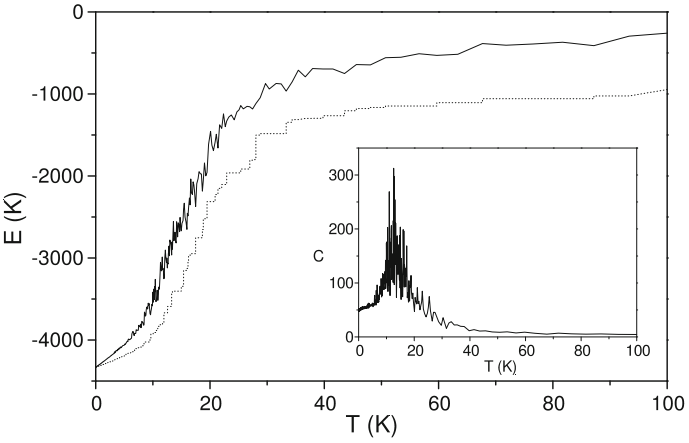


Fig. 9.12 The average energy as a function of temperature for the neon-graphene model system with 100 neon atoms (full line). The dotted line shows the 'best-so-far' energy, which is a piecewise constant function. The inset shows the heat capacity. The kinetic energy contribution is not included

Figure 9.13 shows the positions of the atoms after the annealing. The mismatch of the graphene lattice constant and the range of the interatomic LJ potential is clearly manifested in the different lattice constants of the neon crystal and the graphene substrate. The neon crystal is almost circular, providing a visual demonstration of the presence of a surface tension (the fact that low energy states are single clusters provides another, see Chap. 8). The figure also shows the annealed structures of two other hypothetical types of atoms with Hamiltonians identical to that of graphene and neon, except that the length scale in the LJ potential has been adjusted to give a smaller mismatch between the two lattices, $\pm 10\%$. These particles are also almost circular but are structured with a single atom per hexagon. At larger coverages one must expect that domains form and a lot of other interesting effects. We will leave the subject to future work.

9.8 Optimization: Genetic Algorithms

Genetic algorithms is another optimization technique which has seen widespread use. Like simulated annealing it is stochastic in nature. It also requires the definition of a cost parameter which, as a cadeau to the name of the procedure, is called a fitness parameter. This parameter is used to decide which of a swarm of replicas of the system has the best probability of survival and procreation from one generation to the next. Small changes are then induced, likewise randomly, from one generation to the next in the structure of the system. In a typical application of the technique to nanoscience, one calculates the geometric structure of a many-atom particle for a number of particles and selects preferentially those with lowest energies.

Like for simulated annealing, optimization of particle structures is far from the only application of the technique. It is used to generate protocols to quantum engineer

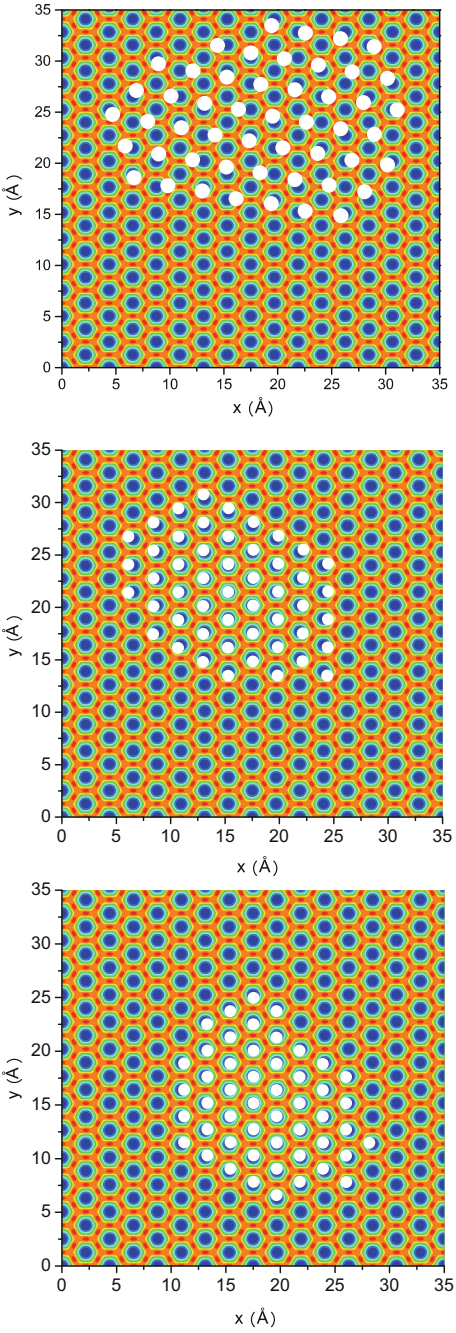


Fig. 9.13 The positions of the 50 neon atoms shown as white dots at the end of an annealing run with the potential in (2.44) and the procedure described in the text (top frame). Two sets of 50 make-believe atoms with interatomic potentials that give a smaller mismatch to the graphene surface are also shown (+10% mismatch middle frame and −10% mismatch bottom frame)

chemical reactions with light, for example, and to generate desirable optical properties of multilayer, 2d materials, to mention just two examples with very practical motivations.

Variations of the basics principle exist, not surprisingly.[4] A first obvious modification is to mate two systems, for example by cutting each into two and gluing two halves together, making it all look very biological. One may speculate that if the system is big enough, it may pay to splice three pieces etc. The risk with this and similar procedures is that they can rapidly develop into an activity that breaks the first rule of numerical simulations: Keep it simple.

Exercises

9.1 Estimate the terms left out in (9.5) and (9.6) to leading order in δt. Use this to establish a criterion for the maximum size of the time step you can use in a MD simulation with this propagation prescription.

9.2 Show that the procedure for updating momenta in the Andersen thermostat in Sect. 9.2 give a distribution of update time which is close to exponential for small timesteps.

9.3 Verify that the distance traveled in the 1000 time units in Fig. 9.1 at small δt's is correct.

9.4 It is of general interest to consider the limit of large systems in numerical calculations, because in this limit the differences between the ensembles used to calculate thermal properties presumably becomes irrelevant, and for example canonical and microcanonical values should become equal.

To show that this holds for the canonical and microcanonical temperatures for a system comprised of harmonic oscillators, first show that the averages of the kinetic energy to the powers 1 and -1 for a system of n h.o. are

$$\begin{aligned} \langle E_k \rangle &= \frac{1}{2} E \\ \langle E_k^{-1} \rangle &= \frac{2n-2}{n-2} E^{-1}. \end{aligned} \tag{9.67}$$

Use this to express T in terms of $\langle E_k \rangle$ and show that it is

$$T = \frac{2}{n-1} \langle E_k \rangle \qquad \text{(microcanonical, harmonic oscillators)}. \tag{9.68}$$

Compare this result with the canonical result.

9.5 Show that the microcanonical heat capacity in (9.29) is correct.

[4] Diversity of implementations of an idea based on diversity seems like a must.

9.6 Show that the values given in (9.44) are the wavenumbers for the problem described.

9.7 The time correlation function of the kinetic energy of a single harmonic oscillator takes the form, disregarding offsets and normalization:

$$c(\Delta t) \propto \int_0^{\frac{2\pi}{\omega}} \sin^2(\omega t) \sin^2(\omega(t+\Delta t)) \mathrm{d}t. \tag{9.69}$$

Show that the oscillating part of $c(\Delta t)$ has a period of π/ω, or half the period of the motion.

9.8 The correlation functions in Fig. 9.5 both have values for zero arguments that can be calculated. For the particle-particle correlation, the value is $\langle E_k^2 \rangle$, where E_k is the kinetic energy of a single particle. The difference, for $j > 0$, $\langle E_k(0)^2 \rangle - \langle E_k(0) E_k(j) \rangle \approx 0.0049$ is approximately equal to $c_v T^2$, where $c_v = 1/2$ is the kinetic energy contribution to the heat capacity from a single particle, and T is the microcanonical temperature of the system which is close to 0.1 at these energies. Assume that the energy in the rubber molecule is low enough to describe the system as harmonic oscillators. Write down the combined level density for the kinetic and the potential energy with this assumption. Use the expression to show that the average total kinetic energy is $E/2$ in this limit, where E is the total energy. Next show that the variance of the total kinetic energy is approximately half the value found in the canonical ensemble and compare that with the figure. Use the fact that there are 99 independent oscillators in the system in Fig. 9.5 to facilitate the approximate calculation of integrals. Show also that the difference of a factor of two persists to infinitely large systems.

9.9 What happens if the Markov matrix in (9.46) would have an eigenvalue exceeding unity? A negative eigenvalue?

9.10 Establish the Markov matrix for a two-state system where moves occur according to the Metropolis algorithm. Show that it has precisely one eigenvalue equal to unity and one numerically less than unity, provided the two states are not degenerate. Note that states are normalized by summing their elements, not the square of their elements, as in quantum mechanics.

9.11 Explain how one can generate random points evenly distributed on a three-dimensional spherical shell of zero thickness.

9.12 Find C_0 in (9.56).

9.13 Design an algorithm to generate kinetic energies, randomly distributed according to the Maxwell-Boltzmann distribution.

9.14 Fill in the details in the derivation of (9.62).

9.15 Write and execute a MC code to calculate the classical heat capacity of a particle in a x^4 potential. Give the variance of $\langle |x| \rangle$ and $\langle x \rangle$ as a function of the number of iterations included in the averaging.

Chapter 10
Thermal Excitation of Valence Electrons

The dynamics of valence electrons is an essential part of the description of small particles, from chemical reactions, over the photo-physics of molecules to the electron transport in nanoscale devices. And as photo-physics in particular has shown, transitions from one Born-Oppenheimer surface to another also happen in the absence of incoming or outgoing photons. Processes called Internal Vibrational Relaxation (IVR) convert electronic excitations to vibrational in a dissipative process, as also discussed in some detail in Chap. 11. Time reversal requires that the inverse process can also happen, and if it can happen, it will happen, at least for a system which is described by statistical tools. Electronically excited states are therefore not just gateways to dissipate energy; they are also players in equilibrium. Thermally excited electronic states are a fact of life.

In most small molecules the energy required to excite electronically is usually a significant fraction of and may even exceed the total excitation energy contained in a molecule, also for particles that would otherwise be considered hot and therefore susceptible to thermal electronic excitation. For those species the thermal population of electronically excited states is usually of little relevance. There are examples, however, of even very small molecules containing sufficient energy to thermally excite states and consuming more than half the total excitation energy in the process. The excited states will be populated for only brief periods of time, but long enough to give observable effects, like strongly enhanced ionization cross sections by photons below the ground state threshold of the IE, or the recurrent fluorescence mentioned in Chap. 6. Although the excited states will be fairly thinly populated, enhanced optical properties or reactivities of these states can make even small occupation numbers important, as the observation of these two examples demonstrate.

For large particles the relevance of thermal electronic excitations becomes even more obvious, in particular for metallic substances. It is worth pointing out that most of the elements in the periodic table are metallic.

Metals are characterized by a good electrical conductivity, which is the result of the delocalization of the valence electrons. In a simple independent particle picture, electrons move in a mean field potential where the size of the wave function is

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_10

determined by the volume of the positive ionic cores. When the number of atoms in the particle and hence the size of the mean field potential increases, quantum mechanical energy levels, both occupied and unoccupied, will decrease in energy. One can estimate the size dependence of levels in a metallic cluster by the similarities to a particle-in-a-box situation and consider the behavior of the levels. They vary with size as

$$E(\{n_q\}, r) - V_0 \approx (E(\{n_q\}, r_0) - V_0) \left(\frac{r_0}{r}\right)^2 \tag{10.1}$$

where r is the particle radius, r_0 is an arbitrary reference size, $\{n_q\}$ is the set of quantum numbers that characterize the state, and V_0 is energy at the bottom of the potential. This size dependence also holds reasonably well for more realistic mean field potentials with a flat bottom and smooth edges. The scaling of the electronic levels implies that an increase in size reduces the gap between the highest occupied orbital (among chemists known as the Highest Occupied Molecular Orbital, or HOMO) and the lowest unoccupied level in the ground state (Lowest Unoccupied Molecular Orbital, LUMO). This gap is the most important parameter that determines the amount of thermal electronic excitation in a particle at a given temperature.

The treatment of the electronic excitations in terms of single particle excitations is obviously not exact. Life is never that easy. Electronic states, including excitations, involve all valence electrons and in principle even the core electrons. In practice, however, one must usually base calculations of thermal properties on excitation spectra derived from single particle levels with the expectation that the results, although not numerically exact, will nevertheless give enough quantitative insight into the phenomenon studied to justify the significant simplification. The good explanatory power of the free electron gas model for simple metals such as the alkali metals supports this expectation.

For metals one uses the term Fermi energy, E_F, to denote the highest electronic kinetic energy in the mean field potential. An equivalent Fermi temperature is defined as $T_F \equiv E_F/k_B$. The Fermi energy for such a free electron gas, unsurprisingly called a Fermi gas, is calculated with the semiclassical expression for a free particle from Chap. 2. With a sloppy use of differentials we have

$$g(E)\mathrm{d}E = \frac{2}{h^3}\mathrm{d}^3x\mathrm{d}^3p = cE^{1/2}\mathrm{d}E, \tag{10.2}$$

where a factor 2 for the spin has been included. g denotes the single particle density of states, not to be confused with the level density of the entire valence electron system. The total number of electrons is calculated as

$$N = \int_0^{E_F} g(E)\mathrm{d}E = \int_0^{E_F} cE^{1/2}\mathrm{d}E, \tag{10.3}$$

from which we find c and the density of states at the Fermi level

$$g(E_F) = \frac{3}{2}\frac{N}{E_F}. \tag{10.4}$$

The average level spacing at the Fermi level is then

$$\Delta \simeq \frac{1}{g(E_F)} = \frac{2}{3}\frac{E_F}{N} \quad \text{(no spin degeneracy)}. \tag{10.5}$$

If the spin degeneracy is taken into account, levels become doubly degenerate and an extra factor of 2 in the spacing appears to give

$$\Delta \simeq \frac{4}{3}\frac{E_F}{N} \quad \text{(with spin degeneracy)}. \tag{10.6}$$

The Fermi energy is related to the electron density in a Fermi gas. The relation is found by noting that (10.3) can also be written as (including the factor two for spin)

$$N = \frac{2}{h^3}\int_V \int_0^{p_F} \mathrm{d}^3x\mathrm{d}^3p = 8\pi\frac{V}{h^3}\frac{p_F^3}{3}, \tag{10.7}$$

with the Fermi momentum defined as

$$E_F \equiv \frac{p_F^2}{2m}, \tag{10.8}$$

where m is the mass of the electron. The set of occupied states form what is called the Fermi sphere in a free electron gas. The electron density, $n = N/V$, and the Fermi energy are related as

$$n = \frac{2\sqrt{2}}{3\pi^2}\frac{m^{3/2}}{\hbar^3}E_F^{3/2} \Rightarrow E_F = \left(3\pi^2 n\right)^{2/3}\frac{\hbar^2}{2m}. \tag{10.9}$$

The number of electrons, N, is the number of atoms in the volume V times the element's valency, i.e. the number of electrons contributed to the Fermi gas per atom. This density is often expressed in terms on the Wigner-Seitz radius, which is the radius of a hypothetical sphere with the density N/V:

$$\frac{4\pi}{3}r_s^3 = n^{-1} = (z\rho)^{-1}, \tag{10.10}$$

where z is the valency, a positive and usually small integer, and ρ is the density of atoms.[1]

[1] Occasionally, r_s is used to denote the corresponding effective atomic radius, which corresponds to setting z to one in (10.10). This is at odds with standard solid state notation and is not recommended.

Table 10.1 Fermi energies in eV as calculated in the Fermi gas approximation from the density and valency

Group 1	Group 2	Group 11	Group 13
Li 4.7	Be 14.4		
Na 3.2	Mg 7.1		Al 11.7
K 2.0	Ca 4.7	Cu 7.0	Ga 10.4
Rb 1.9	Sr 3.9	Ag 5.5	In 8.6
Cs 1.5	Ba 3.6	Au 5.5	Tl 8.1

The Fermi energies cover an order of magnitude in energy, from 1.5 eV for Cs to 14 eV for Be (see Table 10.1). For the somewhat arbitrary choice of 3 eV for the Fermi energy and 800 K for the temperature, the ratio of Δ and the temperature is

$$\frac{\Delta}{T} \sim \frac{2 \cdot 35000K}{3NT} \approx \frac{30}{N}, \tag{10.11}$$

or twice that if the spin degeneracy is included. For metal clusters that are describable with an equidistant level spacing, one must therefore expect that thermal electronic effects become important at sizes around 30 and larger, give or take some factors of two.

With the onset of electron thermodynamics for such relatively small sizes, it becomes relevant with a comparison of equilibrium thermal electronic excitations and those of the vibrational motion of the nuclei. The latter have a typical total thermal energy of $\overline{E}_v \sim 3NT$ for temperatures above the quantum energy of the vibrations, which can be collectively represented by the Debye temperature. For a strongly degenerate Fermi gas, i.e. one where the temperature is much less than the Fermi energy, only electrons in a narrow energy band with a width on the order of T are excited. They each have a typical thermal energy of $\sim T$. The fraction of valence electrons in this region is $\sim T/E_F$ and the thermal energy is then $\overline{E}_e \sim NT^2/E_F$. Relative to the phonon thermal energy, we therefore have an electronic excitation energy of

$$\frac{\overline{E}_e}{\overline{E}_v} \sim \frac{\frac{NT^2}{E_F}}{3NT} = \frac{T}{3E_F}, \tag{10.12}$$

which is small. Any particle at a temperature equal to the Fermi energy will disintegrate in a spray of atoms in less time than a typical vibrational period. The electrons will therefore carry only a small fraction of the total excitation energy under most circumstances.

At low energies the picture changes. In bulk, the heat capacities of solids vary as T^3 below the Debye temperature. Bulk electronic heat capacities are approximately proportional to T as argued above, also at low temperatures, barring the cases where the metal becomes superconducting and the electronic heat capacity becomes very low. The different temperature dependences of the vibrations and the electrons mean that

the excitation energy of the electrons in the so-called normal (non-superconducting) state will be higher than that of the phonons at sufficiently low temperatures.

The reason for this somewhat counterintuitive cross-over is the fact that the spectrum of the fundamental vibrations decreases rapidly at low energy, approximately as ω^2, which implies a strong reduction of the effective number of vibrational degrees of freedom at low temperature. Electrons, on the other hand, are found in a strongly degenerate Fermi gas at temperatures below E_F. Around the Fermi energy the single particle level density is high, and there is a large number of unoccupied states available within an energy corresponding to the temperature. It is this fact that allows bulk electrons to be excited at even very low temperatures.

For sufficiently small particles the picture changes again, because the electronic level spacing becomes comparable to and larger than the Debye frequency. For small enough particles the electronic excitations will therefore be suppressed relative to phonons, both at low and high temperatures.

Suppressed only means numerically smaller. It does not mean unimportant. As we will see later, there are several effects associated with the electronic structure that can, and will, be washed out at realistic temperatures, in spite of only a small amount of thermal energy carried by the electrons. In particular, thermally excited electrons will tend to smear out any size specific properties that depend on the electronic ground state energies, such as odd-even effects and electronic shell structure, which will both be treated in detail in this chapter. It is clear from previous chapters that the presence of thermally excited electrons will also have an effect on other properties; The electrons contribute to level densities, canonical partition functions and free energies on an equal footing with all other degrees of freedom in equilibrium, and any process that involves thermal properties will therefore be affected.

The combination of vibrational and electronic excitations may at times give synergistic effects. An example is thermally induced particle shape fluctuations. Such fluctuations can be accompanied by excitations to BO surfaces where the deformation potential is different than in the ground state. And the excitation to such states may be easier or more difficult in the deformed electronic ground state than in the undistorted. A complete mapping of any system requires that such effects are taken into account. This will be a fairly system-specific mapping. We will settle for the more limited goal of describing the effects of electronic excitations in the static approximation, where levels are given and occupation numbers calculated.

Although electrons will often be thermally excited, the temperatures one encounters are very (very!) rarely be high enough to excite all electrons, although it may be possible to approach conditions where a significant fraction have excitation energies on the order of eV (see Chap. 11). The Fermi energy is too large for thermal excitations to probe the bottom of the Fermi sea, as the collection of electrons obeying the Pauli principle is called (this is the origin of the water metaphor, by the way; waves rarely reach the bottom of the sea). This means that most of the time we can calculate the thermal properties of the system as if the number of electrons were infinite.

At the other end of the spectrum, the existence of electronic states is not switched off when the vacuum level is reached at the ionization energy, IE, but their nature changes dramatically. The IE, which for bulk matter is called the work function, is

similar in magnitude to the Fermi energy. Excitation to energies above IE, which is the basic process in thermionic emission, is therefore a strongly suppressed phenomenon, in spite of the not infrequent observation of it in mass spectrometers. For a single, very highly excited electron, one can use the Boltzmann factor to calculate the probability that the electron energy is excited above an energy which corresponds to the IE as $P \sim \mathrm{e}^{-\mathrm{IE}/T}$. The reader is invited to do this calculation with typical values of IE and T and verify that it is indeed very small. For this reason one can pretend that the single particle level scheme extends to infinitely high energies in most processes.

10.1 Electron Number Fluctuations in the Grand Canonical Ensemble

The first attempt to calculate thermal properties of a specific fermionic system in the strongly degenerate regime will invariably use the grand canonical ensemble in which the temperature, volume and chemical potential are fixed. The variable conjugate to the chemical potential is the particle number, and we can calculate the mean particle number in the ensemble as a sum over contributions from the single particle levels i with energy ε_i:

$$\overline{N} = \frac{\partial \ln(Z_{gc})}{\partial(\beta\mu)} = \frac{\partial \sum_i \ln\left(1 + \mathrm{e}^{-\beta(\varepsilon_i - \mu)}\right)}{\partial(\beta\mu)} = \sum_i \frac{1}{\mathrm{e}^{\beta(\varepsilon_i - \mu)} + 1}. \tag{10.13}$$

The first equality follows from the definition of the grand canonical partition function, and the second from the calculation of (the logarithm of) the partition function for fermions. The terms in the last sum are simply the thermally averaged occupation numbers of the individual states, p_i;

$$p_i \equiv 0 \cdot \frac{1}{z_i} + 1 \cdot \frac{\mathrm{e}^{-\beta(\varepsilon_i - \mu)}}{z_i} = \frac{1}{\mathrm{e}^{\beta(\varepsilon_i - \mu)} + 1}, \tag{10.14}$$

with $z_i = 1 + \mathrm{e}^{-\beta(\varepsilon_i - \mu)}$. Hence the total average electron number is simply a sum of thermally averaged populations of the individual single particle levels. This allows us to write it as an integral with the single particle density of states $g(\varepsilon)$, a form which will often prove convenient:

$$\overline{N} = \int \frac{g(\varepsilon)}{\mathrm{e}^{\beta(\varepsilon - \mu)} + 1} \mathrm{d}\varepsilon \equiv \int g(\varepsilon) f(\varepsilon) \mathrm{d}\varepsilon. \tag{10.15}$$

Here, and elsewhere, the integral over states is acceptable, even if the spectrum is discrete, because $g(E)$ can always be expressed as a sum over δ functions. This was discussed in Chap. 1 and this discussion applies equally well here. The function f,

$$f \equiv \frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1} \tag{10.16}$$

is called the Fermi function and is ubiquitous in solid state physics.

The drawback of the grand canonical ensemble in connection with small particles for which the precise number of electrons is important, is precisely that it is not an ensemble with a fixed particle number. The variance of the particle number can be calculated either as the contribution from each level or, in analogy to the calculation of the variance of the energy in the canonical ensemble and shown explicitly in Chap. 2, as

$$\overline{N}^2 - \overline{N}^2 = \frac{\partial^2 \ln(Z_{gc})}{\partial(\beta\mu)^2} = \tag{10.17}$$
$$\sum_i \frac{1}{\mathrm{e}^{\beta(\varepsilon_i-\mu)}+1}\left(1-\frac{1}{\mathrm{e}^{\beta(\varepsilon_i-\mu)}+1}\right) = \sum_i p_i(1-p_i).$$

This result is both in agreement with the alternative calculation (not shown), and reasonable: At low temperatures where the population of a level is either 1 (below the Fermi energy) or 0 (above the Fermi energy), the variance is zero. At very high temperatures all states are populated very sparsely and one can approximate the terms $p_i(1-p_i) \approx p_i$. This makes $\sum_i p_i(1-p_i) \approx \sum_i p_i = \overline{N}$, i.e. the variance is the sum of single particle occupation numbers and consequently equal to the mean number of particles. The standard deviation of the particle number is thus the square root of the number itself.

At intermediate temperatures where a few excitations from around the Fermi level occurs, the variance and the standard deviation are already of order unity. Some properties, like electronic shell structure, ionization energy and electron affinity, may depend strongly on the number of electrons in the particle, and the average over neighboring sizes implicit in the grand canonical ensemble will therefore give misleading results in this interesting temperature range.

The standard deviation of an electron number distribution in the grand canonical ensemble is shown in Fig. 10.1 as a function of the scaled temperature. It is calculated for a simple model with equidistant single particle level spacing, Δ, which we will have occasion to use later, and which we will call a ladder spectrum. There is (unrealistically) no degeneracy of the individual levels. The data illustrate a point which is also relevant for more realistic situations, viz. that the number fluctuation is significant already below $T = \Delta$.

Another example which demonstrates the effect of the lack number conservation is given in Fig. 10.2, which shows the probabilities, $P(1)$, that precisely one electron is excited in the 1000 electron-2000 state ladder spectrum in the canonical and the grand canonical ensembles.

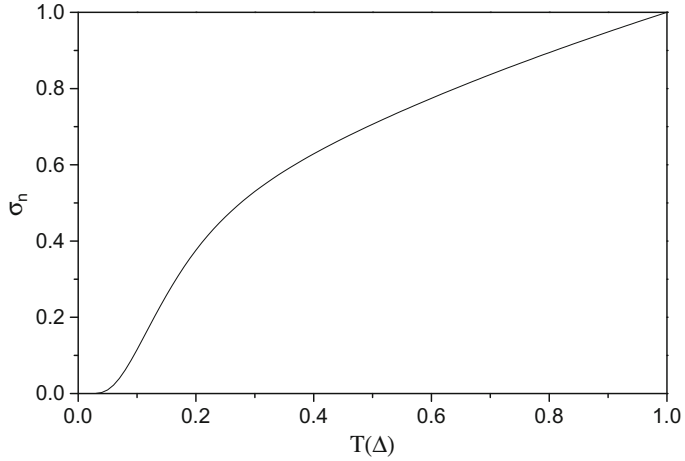


Fig. 10.1 The standard deviation of the particle number in the grand canonical ensemble for a single particle spectrum with equidistant levels, Δ. The number of electrons is 1000 and the number of states 2000, which are both practically infinite at these temperatures

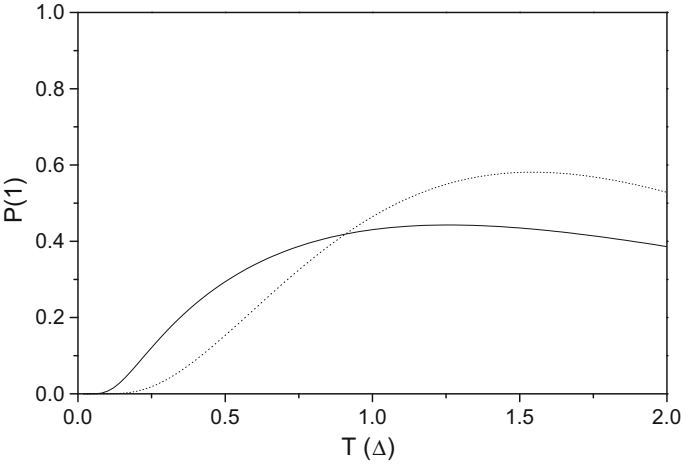


Fig. 10.2 The probability that precisely one electron is excited to a state above the chemical potential in the ladder (equidistant) potential. The full line shows the grand canonical ensemble values; the dotted line gives the canonical ensemble values

The values are calculated for the canonical case as

$$P_c = \frac{\mathrm{e}^{-\beta\Delta}}{\left(1-\mathrm{e}^{-\beta\Delta}\right)^2}\prod_{k=1}^{1000}\left(1-\mathrm{e}^{-k\beta\Delta}\right). \tag{10.18}$$

For the derivation of this expression, material from Sect. 10.4 is needed. The grand canonical curve can be calculated with results available at this point. The population of a state in this ensemble is given by

$$p_k = \frac{\mathrm{e}^{-\beta(\varepsilon_k-\mu)}}{1+\mathrm{e}^{-\beta(\varepsilon_k-\mu)}}, \tag{10.19}$$

and the desired probability is therefore

$$P_{gc} = \sum_{k=1001}^{2000}\frac{p_k}{1-p_k}\prod_j\left(1-p_j\right). \tag{10.20}$$

10.2 Thermal Electronic Properties in the Microcanonical Ensemble

Valence electrons arrive with an atom, and atoms will vibrate and, as discussed, will usually have a thermal energy much larger than the thermal energy carried by electronic excitations. This is therefore the situation of a small system in equilibrium with a bigger system which is described in Chap. 1. We can therefore use the vibrational heat bath to find a simple expression for the contribution of electronically excited states to the thermal properties of a particle at constant total energy. The standard formula for the partitioning of energy between different degrees of freedom is

$$\rho_{tot}(E) = \int_0^E \rho(E-\varepsilon)\rho_{el}(\varepsilon)\mathrm{d}\varepsilon \approx \rho(E)Z_{el}, \tag{10.21}$$

for a system with total energy E, electronic level density ρ_{el} and vibrational level density $\rho(E)$. The last approximate equality that gives the canonical partition function of the electronic excitations is the same one that was used to calculate the influence of electronic degrees of freedom on rate constants in Chap. 5. This result describes to a good accuracy the electronic contribution to the total level density.

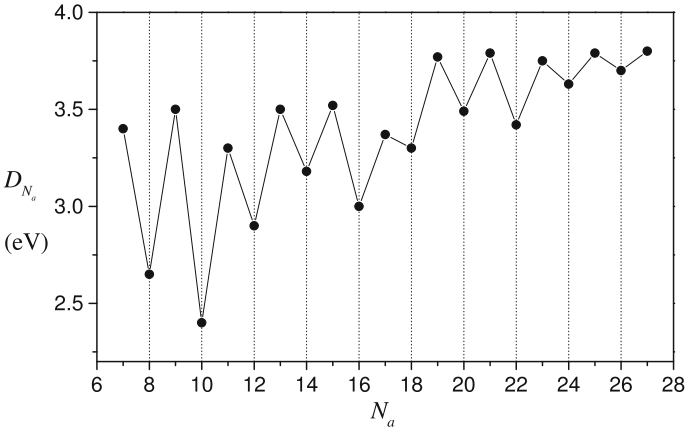


Fig. 10.3 Dissociation energies (evaporative activation energies) for the process $Au_N^+ \rightarrow Au_{N-1}^+ + Au$. Note that N here denotes the number of atoms and not the electron number. The odd number clusters contain an even number of electrons and the D's for these are consequently higher than for the even numbered. Data published in K. Hansen et al., *Phys. Rev. A* **73** (2006) 063202

10.3 The Odd-Even Effect

One of the strongest manifestations of the size dependence of the electronic contribution to particle stability is the odd-even effect. The expression refers to the tendency of properties of metal clusters to oscillate with period two in the number of valence electrons. The oscillations are associated with the spin degeneracy of the electron, and the properties affected are mainly binding and ionization energies. It was discovered in abundance spectra of hot metal clusters, where the clusters with even electron numbers have higher intensities than the odd-numbered. The phenomenon is seen in small clusters of the alkali metals (Li, Na, K, Cs) and particularly pronounced in the coinage metals (Cu, Ag, Au). Figure 10.3 shows the odd-even effect in the dissociation energies of small, positively charged gold clusters.

In the simplest picture the phenomenon reflects a single particle spectrum with equidistant levels with the standard Kramers degeneracy of two due to the electron spin. Often there will be some level bunching in addition to the spin degeneracy that will cause local variations in the gap, Δ, in the single particle spectrum.

The experimental dissociation energies for cationic gold cluster in Fig. 10.3 can be converted to the so-called odd-even energies. We use the standard expansion of quantities that have a size-to-size variation, consisting of a smooth liquid drop dependence with size and a term that varies more or less rapidly with N (see (8.13)). The shell structure appears to be of minor importance for gold clusters and will be ignored in the discussion here. We will then use the equidistant level spacing and calculate the thermal properties for this system with the spacing for doubly degenerate levels,

(10.6). The assumptions therefore represent the ideal case of complete smearing of levels around the Fermi energy.

If the contribution from the ionic cores is smoothly varying with size, we can find the odd-even effect by calculating the electronic ground state energies for both odd and even electron numbers and subtract the smooth part. With the valency z of the material and the net charge q of the particle, the number of electrons is $N = zN_a - q$, where N_a is the number of atoms. For gold we can assume that $z = 1$ as in bulk gold, although bulk valencies are not necessarily representative for values in small particles. Photo-electron studies of mercury clusters, for example, show a transition from van der Waals bound ($z = 0$) to metallic clusters ($z > 0$) at sizes of some dozen of atoms. For gold we expect the bulk values to hold, though, mainly because of the presence of the odd-even effect.

Let's be a little more precise with the meaning of the expression odd-even effect. In the absence of the effect, the Fermi energy is a constant or a smoothly varying function of N. The odd-even effect is defined as the variation with period 2 of the energy of the highest occupied level (HOMO), on top of this smoothly varying Fermi level. We can then calculate the effect in the simplified ladder level N electron systems as follows: For the even numbered clusters N and $N + 2$, the HOMO energy is set to be identical. We can choose it to be zero without any loss of generality. This is achieved by lowering all levels, occupied and unoccupied alike, with a linear interpolation between the two neighboring even-numbered clusters. For the HOMO energy of the odd cluster in between these two, the HOMO level is therefore located at energy $\Delta/2$. Figure 10.4 illustrate the logic.

Defining the average of the odd and even levels as the average, we can then write the electronic ground state energy as a smooth term, $\tilde{E}_N^0$, plus the staggering odd-even contribution as

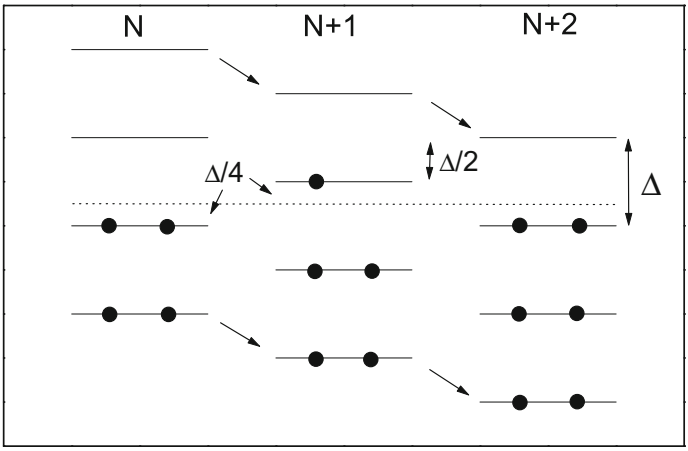


Fig. 10.4 The change of the populated levels with size around the Fermi energy in the simple ladder spectrum used to determine the odd-even energy. Lines represent levels and dots electrons

$$E_N^0 = \tilde{E}_N^0 - \frac{(-1)^N}{4}\Delta. \tag{10.22}$$

Introducing the expression for the number of electrons, which for the cationic gold clusters is $N = N_a - 1$, (10.22) becomes

$$E_{N_a}^0 = \tilde{E}_{N_a}^0 - \frac{(-1)^{N_a - q}}{4}\Delta. \tag{10.23}$$

This gives rise to an electronic contribution to the dissociation energies of

$$E_{a,N_a} \equiv E_{N_a-1}^0 - E_{N_a}^0 = \tilde{E}_{a,N_a} + \frac{(-1)^{N_a-1}}{2}\Delta, \tag{10.24}$$

where the smooth dissociation energy is defined as $\tilde{E}_{a,N_a} \equiv \tilde{E^0}_{N_a-1} - \tilde{E^0}_{N_a}$. As in most other places it is assumed to be equal to the evaporative activation energy.

We can find Δ from the experimentally determined E_{a,N_a}'s from the relation

$$E_{a,N_a} - \frac{1}{2}(E_{a,N_a+1} - E_{a,N_a-1}) \approx (-1)^{N_a-1}\Delta, \tag{10.25}$$

when the change of the smooth term with size is ignored.

The Δ's determined from the experimental data in Fig. 10.3 are shown in Fig. 10.5. The simple Fermi gas prediction derived above is also shown multiplied by an arbitrary factor of two.

Equation (10.24) is the zero temperature value of the effect and the effect will only decrease with temperature for the schematic single particle level structure we have

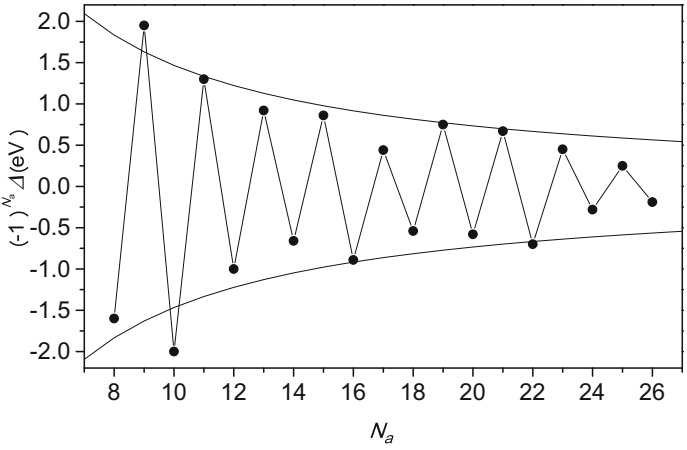


Fig. 10.5 The separation of single particle levels calculated with the Fermi gas approximation (10.6) (line) and the experimental data for gold together with (10.25)

used. The data show that the effect is in fact underestimated relative to experiment. Several mechanisms can change the magnitude, both upward and downward. An effect that will increase the observed gap is the tendency for levels to repel at the Fermi energy. This can be understood with a quasi-explanation as an effect of maximizing the binding energy. A larger gap means lower energies of the filled orbitals and hence a tendency to a larger binding energy. The effect will make the odd-even effect stronger than expected from our simple model. Another possible effect is that the liquid drop energies, which implies that addition energies contain a smooth term in addition to the electronic contribution are not really smoothly varying with size. Indeed, small gold clusters are known to be 2-dimensional and experience a transition to approximately spherical symmetry around the low end of the mass range measured here. Seemingly, these effects did not have any major influence on the magnitude of the dissociation energies in this case.

10.4 Canonical Properties of the Equidistant Spectrum

The canonical partition function can be calculated for the equidistant single particle level system used above to model the odd-even effect. This is a model of a strongly degenerate Fermi gas, and the system is close enough to real Fermi gasses to make a detailed study of it worthwhile, as the data on gold in Fig. 10.5 has demonstrated.

Consider first the situation where there is no spin degeneracy. The partition function for this system can be calculated with a recurrence relation. We denote the partition functions for N electrons in the system by $Z(N)$, dropping the temperature dependence for ease of notation. This partition function can be written as the sum of two contributions, Z_1 and Z_2. One of these contribution, Z_1 say, is the sum over states where the lowest single particle level is occupied, and the other, Z_2, is the sum where it is unoccupied. For Z_1, the sum is over configurations that contain $N-1$ electrons that are all shifted upward by Δ from the lowest single particle state. Z_1 is therefore the $N-1$ electron system with the canonical partition function $Z(N-1)$, by definition, and with the Boltzmann factor $\mathrm{e}^{-\beta\Delta(N-1)}$. The contribution is therefore $Z_1 = Z(N-1)\mathrm{e}^{-\beta\Delta(N-1)}$. Z_2 is the one where the lowest single particle state is unoccupied and all electrons therefore shifted up by Δ. The term for these configurations is therefore $Z_2 = Z(N)\mathrm{e}^{-\beta\Delta N}$. In total we then have

$$Z(N) = Z(N-1)\mathrm{e}^{-\beta\Delta(N-1)} + Z(N)\mathrm{e}^{-\beta\Delta N}. \tag{10.26}$$

This is rewritten

$$Z(N) = Z(N-1)\frac{\mathrm{e}^{-\beta\Delta(N-1)}}{1-\mathrm{e}^{-\beta\Delta N}}. \tag{10.27}$$

The exponential numerator is just the Boltzmann factor corresponding to the change in the ground state energy on adding the last electron because the highest occupied level in the ground state has the state energy $\Delta(N-1)$. If we therefore make the

sensible choice that the zero of energy is the ground state energy of the entire N-electron system, the numerator becomes unity and we get

$$Z(N) = Z(N-1)\frac{1}{1-\mathrm{e}^{-\beta\Delta N}}. \tag{10.28}$$

This is a very simple recurrence relation from which we calculate

$$Z(N) = \prod_{n=1}^{N} \frac{1}{1-\mathrm{e}^{-\beta n\Delta}}. \tag{10.29}$$

As explained above we can, on physical grounds, let N go to infinity in this formula. This is also clear from an inspection of (10.29) if we consider the magnitude of $\exp(-n\beta\Delta)$. Under any realistic temperatures, this is vanishingly small when $n \geqslant N$. There will later be occasion to use the result with another zero of energy. The expression then acquires a Boltzmann factor. With a ground state energy of $E_0(N)$, the partition function becomes:

$$Z(N) = \mathrm{e}^{-\beta E_0(N)} \prod_{n=1}^{\infty} \frac{1}{1-\mathrm{e}^{-\beta n\Delta}}. \tag{10.30}$$

The partition function and quantities derived from the ladder spectrum are identical to the ones of a collection of uncoupled harmonic oscillators with quantum energies $n\Delta$, $n \in Z_+$. The thermal excitation energy of the system is

$$E = -\frac{\partial}{\partial\beta} \sum_{n=1}^{\infty} \ln\left(1-\mathrm{e}^{-\beta\Delta n}\right)^{-1} = \sum_{n=1}^{\infty} \frac{n\Delta \mathrm{e}^{-n\beta\Delta}}{1-\mathrm{e}^{-n\beta\Delta}}. \tag{10.31}$$

High energy approximations for thermal energies and heat capacities will be calculated in Sect. 10.5. The heat capacity, valid for all temperatures, is given by

$$C_v = \sum_{n=1}^{\infty} \frac{(n\Delta\beta)^2 \mathrm{e}^{-n\beta\Delta}}{\left(1-\mathrm{e}^{-n\beta\Delta}\right)^2}. \tag{10.32}$$

The chemical potential is calculated with the discrete version of the derivative of $\ln(Z)$ with respect to particle number.

$$\beta\mu = -\ln\left(\frac{Z(N)}{Z(N-1)}\right), \tag{10.33}$$

One must use a common zero of energy when comparing systems with different electron numbers, and it is convenient to set the single particle ground state energy to zero in this calculation. The chemical potential is then calculated to

$$\beta\mu = -\ln\left(\frac{\mathrm{e}^{-\beta\Delta(N-1)}}{1-\mathrm{e}^{-N\beta\Delta}}\right), \tag{10.34}$$

or

$$\mu = \Delta(N-1) + T\ln\left(1-\mathrm{e}^{-N\beta\Delta}\right). \tag{10.35}$$

The leading term is constant and the last term is an exponentially small negative term for temperatures less than a value which is on the order of the Fermi energy $N\Delta$.

The almost constant value of μ in this case reflects the very symmetric nature of the ladder spectrum with respect to energy. The equation does not represent the general case for electrons quantitatively, but other single particle spectra will show a similar strong suppression of the temperature dependence of the chemical potential if they are symmetric with respect to energy around the Fermi energy.

The general behavior of the chemical potentials can be understood from the Sommerfeld expansion, which is an expansion of the thermal properties of a Fermi gas at low temperatures, $T \ll T_F$. The expression for a property, H, of the Fermi gas which depends on the single particle energy ε, $H(\varepsilon)$, is in this expansion

$$\int_0^\infty H(\varepsilon)\mathrm{d}\varepsilon = \int_0^\mu H(\varepsilon)\mathrm{d}\varepsilon + \frac{\pi^2}{6}T^2\frac{\mathrm{d}H}{\mathrm{d}\varepsilon} + \frac{7\pi^4}{360}T^4\frac{\mathrm{d}^3H}{\mathrm{d}\varepsilon^3} + \cdots, \tag{10.36}$$

if there are no singular points in the spectrum. The chemical potential can be calculated with this expansion (see Exercise 10.4) which gives

$$\mu = E_F - \frac{\pi^2}{6}T^2\frac{\frac{\mathrm{d}g(E_F)}{\mathrm{d}\varepsilon}}{g(E_F)}, \tag{10.37}$$

where $g(\varepsilon)$ is the single particle density of states. The second derivative is non-positive as expected on general grounds (see Exercise 1.10. Note however, that the system here is not included in the class treated in that problem; Paradoxically it is not large in the sense used in the derivations in that problem, because a Taylor expansion in temperature is not possible due to the Boltzmann factor in the last term of (10.35)).

Figure 10.6 shows the excitation energies in the canonical and the grand canonical ensembles for the single particle ladder spectrum. It demonstrates that excitation energies are not negligible even before the temperature reaches the value corresponding to the level spacing Δ. The inset shows the ratio of the two on a linear scale. The heat capacity for the system shown in Fig. 10.7.

Let's now make the problem quasi-realistic by including the double degeneracy that represent the spin degeneracy of each single particle level. This task is made much easier if we use the result for the non-degenerate case we have just derived. Look at the problem as two systems of the non-degenerate type that can exchange particles freely, only constrained by the price of the Boltzmann factor and total particle number conservation. The partition function of the whole system, $Z(N,2)$, is the product of the two partition functions, summed over all possible occupation

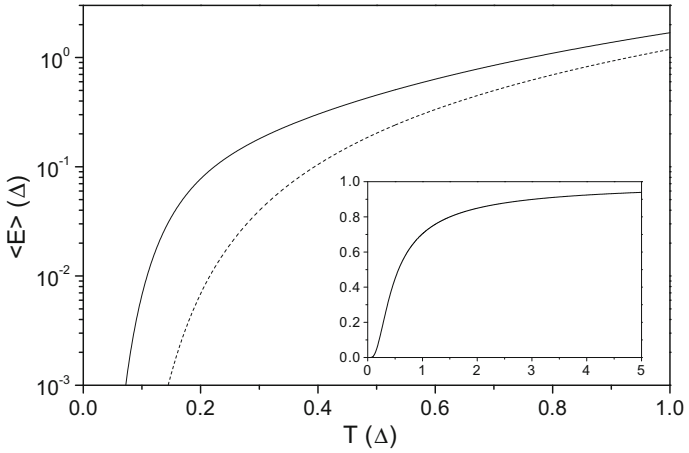


Fig. 10.6 The excitation energy of the ladder spectrum in the grand canonical (full line) and canonical (dotted line) ensemble. The inset shows the ratio of canonical to grand canonical excitation energies vs. temperature. Temperature and energy units are both Δ. Note that the curve also represent the size dependences to the extent that Δ scales with size as in (10.5)

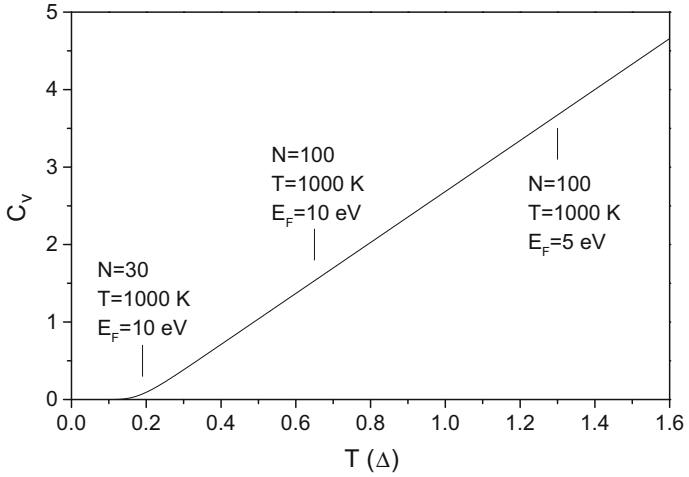


Fig. 10.7 The canonical heat capacity of the fermionic ladder spectrum with spacing Δ. As in Fig. 10.6, the temperature dependence is implicitly also a size dependence through Δ. A few combinations of T, E_F and N are indicated

numbers that conserve the total number of electrons, with the proper Boltzmann factors as weights;

$$Z(N,2) = \sum_{p=0}^{N} Z(N-p)Z(p)$$
$$= \sum_{p} \left[\mathrm{e}^{-\beta E_0(N-p)} \prod_{n=1}^{N-p} \frac{1}{1-\mathrm{e}^{-\beta n\Delta}} \right] \left[\mathrm{e}^{-\beta E_0(p)} \prod_{n=1}^{p} \frac{1}{1-\mathrm{e}^{-\beta n\Delta}} \right]. \quad (10.38)$$

The ground state energies $E_0(N-p)$ and $E_0(p)$ will increase quadratically with the number of electrons in the subsystems, $N-p$ and p, and this makes the Boltzmann factors that involve these energies very strongly energy dependent. The two products, on the other hand, vary much weaker with particle number. The main contributions to the sum are therefore the partitionings where the number of electrons in the two subsystems are close to equal, $p \simeq N/2 - p \simeq N/2$. As before we can then extend the products to infinity without any significant loss of precision:

$$Z(N,2) = \left(\prod_{n=1}^{\infty} \frac{1}{1-\mathrm{e}^{-\beta n\Delta}} \right)^2 \sum_{p=0}^{N} \mathrm{e}^{-\beta(E_0(N-p)+E_0(p))}. \quad (10.39)$$

The sum depends on whether the system contains an even or an odd number of electrons. The even electron numbers first. The ground state energy of the state with the electrons partitioned as $N/2-m$ and $N/2+m$ is, relative to the absolute ground state with $m=0$:

$$E = \Delta \sum_{k=1}^{m} (2k-1) = m^2\Delta, \quad (10.40)$$

because the energy needed for the kth electron to be transferred from one column to the other is $(2k-1)\Delta$. Including the Boltzmann factor for $\Delta/4$ from (10.22) for the odd-even energy, the total partition function for an even electron particle is therefore equal to

$$Z(N=\text{even},2) = \mathrm{e}^{\beta\Delta/4} \left(\prod_{n=1}^{\infty} \frac{1}{1-\mathrm{e}^{-\beta n\Delta}} \right)^2 \sum_{m=-\infty}^{\infty} \mathrm{e}^{-m^2\beta\Delta}. \quad (10.41)$$

When the analogous calculation of the ground state energy of the electron partitionings is made for the odd electron systems, one gets

$$E = m(m+1)\Delta. \quad (10.42)$$

Equation (10.42) reproduces the doubly degenerate ground state, as required because $m = 0$ and $m = -1$ are degenerate. The partition function for this system is

$$Z(N = \text{odd}, 2) = \mathrm{e}^{-\beta\Delta/4} \left(\prod_{n=1}^{\infty} \frac{1}{1 - \mathrm{e}^{-\beta n \Delta}} \right)^2 \sum_{m=-\infty}^{\infty} \mathrm{e}^{-m(m+1)\beta\Delta}. \tag{10.43}$$

At low temperatures, $\beta\Delta > 1$, the sums that appear in (10.41, 10.43) are best calculated directly by explicit summing over the few terms that are important. At high temperatures the sum would seem like a good candidate for a calculation with the Euler-Maclaurin formula. This turns out not to give a very accurate result. In fact, only the first term in that formula, the integral, is non-zero to any finite order of the expansion.

Instead we can calculate the sum with a trick. Consider the sum

$$S(x) \equiv \sum_{m=-\infty}^{\infty} \mathrm{e}^{-m(m+x)\beta\Delta} = \mathrm{e}^{\beta x^2 \Delta/4} \sum_{m=-\infty}^{\infty} \mathrm{e}^{-(m+x/2)^2 \beta\Delta} \equiv \mathrm{e}^{x^2\beta\Delta/4} S'(x). \tag{10.44}$$

The two sums we are after are the special values $S(0)$ and $S(1)$. The sum $S'(x)$ is periodic in x. This means that it can be expressed as a Fourier series. The period in x is 2, and we have

$$S'(x) = \sum_{k=-\infty}^{k=\infty} c_k \mathrm{e}^{ik2\pi x/2}, \tag{10.45}$$

with the coefficients c_k determined as

$$c_k = \frac{1}{2} \int_0^2 \sum_{m=-\infty}^{\infty} \mathrm{e}^{-\beta\Delta(m+x/2)^2 - 2\pi i k x/2} \mathrm{d}x. \tag{10.46}$$

We can add $2\pi i$ times an integer to the argument of the exponential and write $\exp(-2\pi ikx/2) = \exp(-2\pi ik(m + x/2))$. Interchanging summation and integration gives a sum that covers the real axis precisely once. This gives us

$$c_k = \frac{1}{2} \int_{-\infty}^{\infty} \mathrm{e}^{-\beta\Delta(x/2)^2 - 2\pi i k x/2} \mathrm{d}x. \tag{10.47}$$

The integral is calculated by completing the square and integrating the resulting Gaussian. The result is

$$c_k = \left(\frac{\pi}{\beta\Delta} \right)^{1/2} \mathrm{e}^{-\frac{\pi^2 k^2}{\beta\Delta}}. \tag{10.48}$$

Introducing these coefficients into the Fourier series we get

$$S(0) = \left(\frac{\pi}{\beta\Delta}\right)^{1/2}\left(1 + 2\sum_{k=1}^{\infty} \mathrm{e}^{-\frac{\pi^2 k^2}{\beta\Delta}}\right), \tag{10.49}$$

and

$$S(1) = \mathrm{e}^{\beta\Delta/4}\left(\frac{\pi}{\beta\Delta}\right)^{1/2}\left(1 + 2\sum_{k=1}^{\infty}(-1)^k \mathrm{e}^{-\frac{\pi^2 k^2}{\beta\Delta}}\right). \tag{10.50}$$

The temperature appears in the numerator in the (negative) argument of the exponentials in these two sums, and the calculation has therefore provided us with a high temperature expansion of the sums.

We now have both a high and a low temperature expansion. With a sufficient number of terms included, these have an overlap where both are accurate. The required number of terms for either expansion is low because of the square integers and π^2 in the exponential in the sums.

The expression for the partition function for odd electron numbers becomes

$$Z(N = \text{odd}, 2) = \mathrm{e}^{-\Delta/4T}\left(\prod_{n=1}^{\infty}\frac{1}{1-\mathrm{e}^{-n\Delta/T}}\right)^2 \times \tag{10.51}$$
$$\begin{cases} \left(\frac{\pi T}{\Delta}\right)^{1/2} \mathrm{e}^{\Delta/4T}\left(1 - 2\mathrm{e}^{-\frac{\pi^2 T}{\Delta}} + 2\mathrm{e}^{-4\frac{\pi^2 T}{\Delta}} + \cdots\right) & \Delta \lesssim T \\ \left(2 + 2\mathrm{e}^{-2\Delta/T} + 2\mathrm{e}^{-6\Delta/T} + \cdots\right) & \Delta \gtrsim T, \end{cases}$$

and for even electron numbers:

$$Z(N = \text{even}, 2) = \mathrm{e}^{\Delta/4T}\left(\prod_{n=1}^{\infty}\frac{1}{1-\mathrm{e}^{-n\Delta/T}}\right)^2 \times \tag{10.52}$$
$$\begin{cases} \left(\frac{\pi T}{\Delta}\right)^{1/2}\left(1 + 2\mathrm{e}^{-\frac{\pi^2 T}{\Delta}} + 2\mathrm{e}^{-4\frac{\pi^2 T}{\Delta}} + \cdots\right) & \Delta \lesssim T \\ \left(1 + 2\mathrm{e}^{-\Delta/T} + 2\mathrm{e}^{-4\Delta/T} + \cdots\right) & \Delta \gtrsim T. \end{cases}$$

The ground state energies are included explicitly in these expressions, and it is therefore possible to calculate the odd-even difference in free energy at finite temperatures with no further ado. In the low temperature limit the ratio of the two partition functions is:

$$\mathrm{e}^{-(F_{\text{even}} - F_{\text{odd}})/T} = \frac{Z(N = \text{even}, 2)}{Z(N = \text{odd}, 2)} = \mathrm{e}^{\Delta/2T}\frac{1 + 2\mathrm{e}^{-\Delta/T} + 2\mathrm{e}^{-4\Delta/T}}{\left(2 + 2\mathrm{e}^{-2\Delta/T} + 2\mathrm{e}^{-6\Delta/T}\right)}, \tag{10.53}$$

or to next-to-leading order

$$F_{\text{even}} - F_{\text{odd}} \approx -\frac{\Delta}{2} + T\ln(2). \tag{10.54}$$

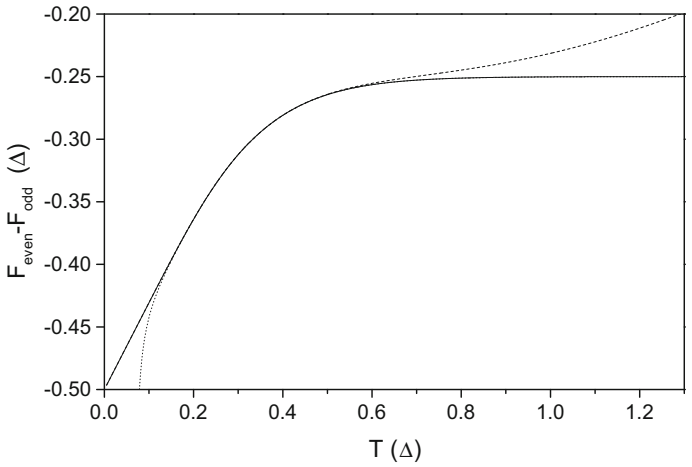


Fig. 10.8 The difference in free energies of two particles with an even and an odd number of valence electrons. The free energies are given in units of the level spacing Δ. The dashed line is the low temperature approximation, (10.53) and the dotted line the high temperature approximation in (10.56). Both the low and high temperature limit include the first two terms in the expansion of the partition functions. The full line is based on the numerically exact summation of the partition functions

For the high temperature limit we have

$$\mathrm{e}^{-(F_{\mathrm{even}}-F_{\mathrm{odd}})/T} = \mathrm{e}^{\Delta/4T}\,\frac{1+2\mathrm{e}^{-\frac{\pi^2 T}{\Delta}}+2\mathrm{e}^{-4\frac{\pi^2 T}{\Delta}}}{1-2\mathrm{e}^{-\frac{\pi^2 T}{\Delta}}+2\mathrm{e}^{-4\frac{\pi^2 T}{\Delta}}} \tag{10.55}$$

and correspondingly

$$F_{\mathrm{even}} - F_{\mathrm{odd}} \approx -\frac{\Delta}{4} - 4T\mathrm{e}^{-\frac{\pi^2 T}{\Delta}}. \tag{10.56}$$

The calculated difference in free energies is shown in Fig. 10.8.

10.5 High Energy Level Density of the Equidistant Spectrum

In the previous section the odd-even effect was calculated as a function of the temperature. The motivation for a calculation of canonical properties for a microcanonical system was given in Sect. 10.2, where it was shown that the microcanonical electronic properties can be converted into canonical with an accuracy which are often sufficient in applications. Occasionally, however, one needs also the level density.

It is even essential when the thermodynamics involve the electrons alone, as in the subject of Chap. 11.

We will therefore calculate the level density of the non-degenerate spectrum with equidistant level spacing, which will continue to stand as the paradigm of a Fermi gas. The high energy part will be calculated here and the low energy limit will be computed later with a numerical procedure which will also be developed.

The term 'high energy limit' means that the level spacing is small compared with the canonical temperature;

$$\Delta \lesssim T. \tag{10.57}$$

We can use the general results in (3.46, 3.47) that express the level density in terms of canonical values. For this purpose we need the canonical partition function, the thermal excitation energy, and the heat capacity. The partition function is already known (10.29), and the other two quantities are calculated from the partition function by the standard procedures.

The logarithm of Z is

$$\ln(Z) = \ln\left(\prod_{N=1}^{\infty} \frac{1}{1-\mathrm{e}^{-\beta N\Delta}}\right) = -\sum_{N=1}^{\infty} \ln\left(1-\mathrm{e}^{-\beta N\Delta}\right), \tag{10.58}$$

where the infinite number of particles indicate an infinitely deep Fermi sea. From this point on the calculation reduces to an application of known operations and is shown in details mainly to give an example of these procedures. The reader who does not need this example and trusts the calculations can proceed to the result in (10.63).

With the help of the Euler-Maclaurin formula we can convert the sum in (10.58) to an integral plus additional small terms:

$$\begin{aligned}\ln(Z) = &-\int_1^{\infty} \ln\left(1-\mathrm{e}^{-x\beta\Delta}\right)\mathrm{d}x - \frac{1}{2}\ln(1-\mathrm{e}^{-\beta\Delta}) \\ &+ \frac{1}{12}\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(1-\mathrm{e}^{-x\beta\Delta}\right)\Big|_{x=1} - \frac{1}{720}\frac{\mathrm{d}^3}{\mathrm{d}x^3}\ln\left(1-\mathrm{e}^{-x\beta\Delta}\right)\Big|_{x=1} + \cdots\end{aligned} \tag{10.59}$$

We will calculate this to first order in $\beta\Delta$. This does not account for all terms up to this order, because the Euler-Maclaurin formula is an asymptotic expansion and not a Taylor series. Some small zero and first order contributions will therefore not be included. The last three terms given in (10.59) are expanded to first order in $\beta\Delta$ to give

$$-\frac{1}{2}\ln(\beta\Delta) + \frac{1}{4}\beta\Delta + \frac{1}{12} - \frac{\beta\Delta}{24} - \frac{1}{360}. \tag{10.60}$$

The integral in (10.59) is calculated by partial integration

$$-\int_1^\infty \ln\left(1-\mathrm{e}^{-x\beta\Delta}\right)\mathrm{d}x = -\left[x\ln\left(1-\mathrm{e}^{-\beta\Delta x}\right)\right]_1^\infty + \int_1^\infty \frac{\beta\Delta x\mathrm{e}^{-\beta\Delta x}}{1-\mathrm{e}^{-\beta\Delta x}}\mathrm{d}x. \tag{10.61}$$

The last integral is written as the difference, $\int_1^\infty = \int_0^\infty - \int_0^1$. The first of these is calculated by expanding the denominator in the exponential function, integrating and resummation of the series with the result $\zeta(2)/\beta\Delta$, where $\zeta(2) = \pi^2/6$ is the value of Riemann's zeta function for the argument 2. The integral from zero to one is calculated with an expansion to next-to-leading order in $\beta\Delta$, which is a meaningful expansion here because we consider the high temperature limit. Expanding also the first term in (10.61) to the same order gives the integral

$$-\int_1^\infty \ln\left(1-\mathrm{e}^{-x\beta\Delta}\right)\mathrm{d}x = \frac{\pi^2}{6}\frac{1}{\beta\Delta} + \ln(\beta\Delta) - 1 - \frac{1}{4}\beta\Delta. \tag{10.62}$$

Adding (10.60, 10.62), the total partition function is then to this degree of approximation equal to

$$\ln(Z) = \frac{\pi^2}{6\beta\Delta} + \frac{1}{2}\ln(\beta\Delta) - \frac{331}{360} - \frac{1}{24}\beta\Delta. \tag{10.63}$$

As the reward for going through this exercise we can now calculate physical quantities. The (canonical) thermal excitation energy is

$$\langle E\rangle = -\frac{\partial\ln(Z)}{\partial\beta} = \frac{\pi^2}{6\beta^2\Delta} - \frac{1}{2\beta} + \frac{\Delta}{24} = \frac{\pi^2}{6}\frac{T^2}{\Delta} - \frac{T}{2} + \frac{\Delta}{24}. \tag{10.64}$$

To find the level density we need to translate this expression into the relation between the microcanonical excitation energy, E, and temperature, $1/\beta$ (see Chap. 3):

$$\langle E(\beta)\rangle = E + 1/\beta, \tag{10.65}$$

which gives us

$$E = \frac{\pi^2T^2}{6\Delta} - \frac{3}{2}T + \frac{\Delta}{24} \Rightarrow T \approx \sqrt{\frac{6\Delta E}{\pi^2}} + \frac{9}{2\pi^2}\Delta, \tag{10.66}$$

and the heat capacity

$$C_v = \frac{\pi^2T}{3\Delta} - \frac{3}{2} \approx \frac{\pi^2}{3}\frac{T}{\Delta}. \tag{10.67}$$

Using (10.63, 10.66, 10.67) in the expression for the level density

$$\rho = \frac{\beta}{\sqrt{2\pi C_v}}\mathrm{e}^{\beta E}Z(\beta), \tag{10.68}$$

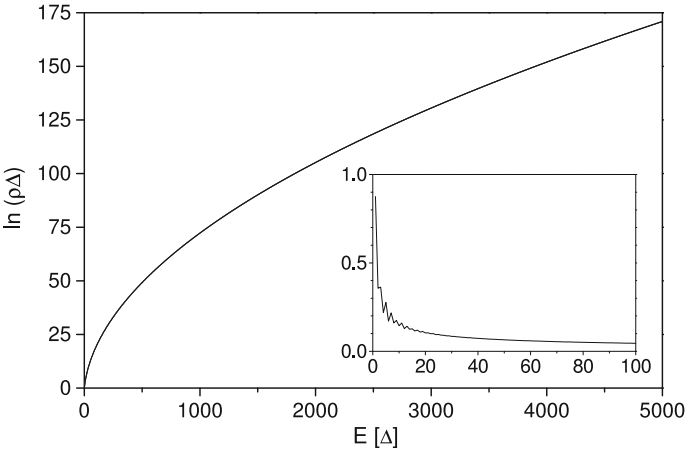


Fig. 10.9 The level density of electrons in a system with equidistant single particle states in units of the level spacing, Δ. The figure shows the numerically exact result from an calculation with (10.84), of which the first few are given in Table 10.3, the approximation in (10.69), and the level density of Bethe, (10.71). These curves all appear as a single line. The curves do not include any spin degeneracy, as is clear from the derivation. The inset shows the error in (10.69) relative to the exact result calculated with (10.84). Note that this is a conservative comparison. Including more terms will give a significant improvement already from the smallest excitation energy (see Table 10.3)

we finally get

$$\rho(E) \approx \frac{c}{E} \mathrm{e}^{\left(\frac{2\pi^2 E}{3\Delta}\right)^{1/2}}, \tag{10.69}$$

where

$$c = \frac{\pi^2}{6}\left(\frac{2}{3}\pi^3\right)^{-1/2} \mathrm{e}^{-331/360} = 0.14426\ldots \tag{10.70}$$

This level density is shown in Fig. 10.9 together with the exact result calculated later in this chapter (10.84). The inset shows the relative difference between the exact result and the one derived here. The error decreases above the energies shown. The result of H. Bethe is also included in the plot. It is very similar to the one calculated here;

$$\rho = \frac{1}{\sqrt{48}E} \mathrm{e}^{\left(\frac{2\pi^2 E}{3\Delta}\right)^{1/2}}. \tag{10.71}$$

In fact, this only differs from the result derived here by the factor $0.14426\ldots \times \sqrt{48} = 0.99949\ldots$

If the simple estimate for Δ in terms of the Fermi energy and the number of electrons in (10.5) is used, one has

$$\rho(E) \approx \frac{c}{E} \mathrm{e}^{\left(\pi^2 \frac{NE}{E_F}\right)^{1/2}}. \tag{10.72}$$

The use of (10.5) for the level spacing instead of the one where spin is taken into account is warranted because we are concerned with the high energy result here, and the total average level spacing is what counts for that limit. One should of course be aware that the approximation is less accurate for the doubly degenerate single level case than for the non-degenerate case calculated here.

10.6 Magnetism: The Superparamagnetic Cluster

The purely academic desire to understand magnetism in all its fascinating manifestations is happily married to its technological importance, which is derived from information technology applications in particular. The continued reduction in size of the transistors in computers, in the form of the often-quoted Moore's law which states that the density of components on chips double each 18 months, is paralleled by the increasingly dense data storage capability of the magnetism-based devices attached to the data processing units. The decreasing size of such components makes the research of the magnetic properties of nano-sized particles more than just interesting. Although applications tend to be thought of as firmly anchored in the solid state, with quantum computing one possible future exception, fundamental properties can with benefit be studied in gas phase.

Magnetic fields are intimately connected with electric fields and moving charges, as demonstrated by H.C. Ørsted and M. Faraday and summarized by J.C. Maxwell in what are now known as Maxwell's equations. With the connection between moving electrical charges and magnetism in mind, it is therefore not surprising that metals with their itinerant electrons can have non-trivial magnetic properties.

The fundamental ingredients of magnetism are the magnetic moments and the interaction between them. The magnetic moments appear classically from circulating electric currents. However, the classical orbital electron motion can not maintain a permanent magnetic moment of any material, as shown by N.Bohr, and the magnetism of materials is fundamentally a quantum phenomenon.

Both orbital motion of the electrons as well as the spins of the electrons will contribute to the magnetic moments. In principle also the nuclei will contribute, although the amount is very small amount compared to the electrons', because the intrinsic magnetic moments of particles are inversely proportional to their mass.

The simplest manifestation of magnetism is probably ferromagnetism. In a ferromagnet, the magnetic moments align and the combined magnetic field from all the moments combine to a macroscopically observable magnetic field, strong enough to keep school schedules and baby pictures fixed to refrigerators.

There are other manifestations of magnetism than ferromagnetism. One is antiferromagnetism, which describes situations where moments are anti-aligned so the total moment averages out to zero. A closely related type is called ferrimagnetism, which is the word used when moments are anti-aligned but the moments in the two directions are of different size. This produces a net macroscopic magnetic field.

Two other types of magnetism arise as the response of materials to an external magnetic field. In the absence of a field, no residual magnetism of these types exists in these materials. One is diamagnetism, which is the word for the magnetism of a material that is repelled by an external magnetic field. The other is paramagnetism, which denotes the state of matter attracted to an external magnetic field. For diamagnetism, the elementary magnetic dipole moments tend to have opposite direction of the external field, and for paramagnetic substances they tend to point in the same direction.

Bulk ferro-, ferri- and antiferromagnetism is an example of spontaneous symmetry breaking, at least in the first, idealized description. In the absence of any internal structure, a macroscopic magnetization can point in any direction relative to the (macroscopic) piece of material it is living in, but it does point in a specific direction, and this direction is not fluctuating. This is spontaneous symmetry breaking. In reality the breaking of the symmetry is not completely spontaneous, because there will be directions that are preferred because of the lattice structure, for example, or the existence of a microcrystal structure in the macroscopic piece of material. Also the shape of the material will usually favor a certain direction of the magnetism, even if all other effects were absent. It is pretty difficult (and not too clever) to use a compass needle magnetized perpendicular to its long direction, because the lowest energy state for the spontaneous magnetization is along the long axis. A compass needle leaves you with two degenerate orientations of the magnetization. Other symmetries of an object may produce more degeneracies. But for a compass needle, once it is magnetized, the direction of the magnetization stays put, North is North and not suddenly South, even if the energy of the reversed field would be the same.

The quantitative description of ferromagnetism is a good deal more complicated than simply placing atomic magnetic moments close together and deriving their interaction energy from the free atom moments. This is easy to see by a comparison of the calculated interaction energy between two magnetic dipoles with the experimentally measured Curie temperatures. The Curie temperature of a material is the temperature at which a spontaneous magnetism appears or disappears when the system is cooled or heated. To calculate the Curie temperature in the simple-minded fashion, consider the fundamental interaction between a magnetic moment μ and an external magnetic field B. It is

$$E = -\overline{\mu} \cdot \overline{B}. \tag{10.73}$$

The magnitude of the magnetic dipole moment, $\overline{m}$, of an atom is on the order of the Bohr magneton, $m \equiv |\overline{m}| \sim \mu_B = 9.3 \cdot 10^{-24}$ J/T $= 0.67$ K/T, and the interaction energy between two such moments with magnitudes m_1, m_2 at a distance r apart is, with μ_0 the vacuum permeability, on the order of

$$V \sim \mu_0 \frac{m_1 m_2}{r^3} \sim \mu_0 \mu_B^2 r^{-3} = \frac{7.9 \text{ K\AA}^3}{r^3}. \tag{10.74}$$

Table 10.2 Measured Curie temperatures of elements

Element	T_C (K)
Co	1394
Dy	88
Er	19
Fe	1043
Gd	292
Ho	20
Ni	627
Tb	221
Tm	32

It is difficult to put two atoms any closer than a few Å, and for interatomic distances above 1 Å, say, this interaction energy will therefore be below 10 kelvin. A more precise estimate requires the inclusion of some factors given by angular momentum quantum numbers. Such quantum numbers are on the order of $\ell \sim Z^{1/3}$, where Z is the atomic number, and will therefore not change the estimate seriously.

Based solely on this estimate of the strength of the dipole-dipole interaction, one would expect that the magnetism will be completely washed out by thermal fluctuations at ambient temperatures. As experience tells us, this is not the case. This is also seen from Table 10.2, where some experimentally measured Curie temperatures are listed, showing that magnetism will in fact survive to temperatures in excess of 1000 K in a few cases, indicating a very strong coupling, high above the above naive estimate. We will set the questions about the origin of such interactions aside and simply use effective magnetic moments as input parameters in the following.

Let's now consider the simplest description of the thermal properties of bulk magnetism, which is given by mean field theory. The Hamiltonian is, in the absence of an external magnetic field, represented by the single parameter expression

$$H = \sum_{i,k} -J s_i s_k \tag{10.75}$$

where the s_i are the atomic spins and the sum runs over all pairs of nearest neighbor spins. All the electron dynamics that change the interaction from the simple values calculated with the Bohr magneton, μ_B, and the Landé factor g is absorbed into the value of the effective coupling constant J, and we only retain the semiclassical two-state description of the orientation of the magnetic moments arising from the nature of the electron's spin, i.e. $s_i = \pm 1/2$.

The mean field theory consists in replacing the pair-wise average $\langle s_i s_k \rangle$ with the factorized average $\langle s_i \rangle \langle s_k \rangle = \langle s \rangle^2$. The thermal average of the orientation of a single spin can then be calculated as the solution to the mean field equation:

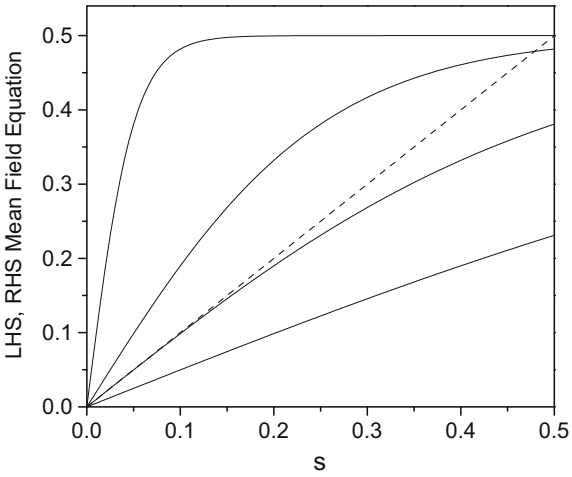


Fig. 10.10 The left (dotted line) and the right hand side (full lines) of (10.76) for different temperatures; From below 0.5, 1, 2, 10 times the critical temperature

$$\langle s\rangle = \frac{1}{2}\frac{e^{\beta 2dJ\langle s\rangle} - e^{-\beta 2dJ\langle s\rangle}}{e^{\beta 2dJ\langle s\rangle} + e^{-\beta 2dJ\langle s\rangle}}, \tag{10.76}$$

where d is the dimensionality of space and $2d$ therefore the number of nearest neighbors in the regular lattice assumed here.

Equation (10.76) always has the solution $\langle s\rangle = 0$, which is the solution found in the high temperature, unmagnetized phase. In addition, for sufficiently low temperatures, it has a solution for some s in the interval $0 < s < 1/2$, which represents the magnetized phase. When the solution exists, the magnetized phase is the stable one, because the free energy of this solution is lower than that of the unmagnetized phase. It has only one of these solutions because the right hand side of (10.76), which is essentially the hyperbolic tangent, is concave when considered a function of the magnetization. The behavior of (10.76) is shown in Fig. 10.10.

We can then find the critical temperature, where magnetized material turns to unmagnetized and vice versa, by equating the derivatives of the left and right hand sides of (10.76) at $\langle s\rangle = 0$. This gives the critical temperature

$$T_c = dJ. \tag{10.77}$$

This is the predicted Curie temperature from mean field theory. The magnetization given by the solution of (10.76) is shown in Fig. 10.11 as a function of the temperature. The onset of this magnetization is not sharp. Below the Curie temperature, there is some amount of thermal fluctuations which will reduce the magnetization from the maximum value. This is one feature that will survive a more sophisticated treatment of the question of bulk magnetism. It is associated with diverging correlation lengths that obey scaling laws and that vary with the difference in temperature to the Curie temperature to some power. The applicability of diverging lengths to small systems is limited and we will leave that subject.

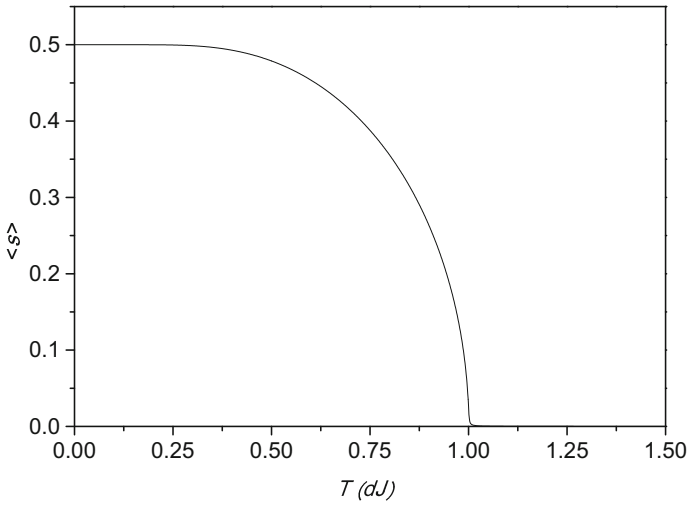


Fig. 10.11 The magnetization as a function of temperature according to (10.76)

In addition to these more or less quantitatively understood fluctuations of individual spins away from the direction of the total moment, adding a little fluctuation to the magnitude of the total moment and its orientation, the total moment itself can also have a fluctuating direction. In the bulk, this is manifested as the propagation of waves in the polarization direction and are called spin waves. They are the so-called Goldstone bosons that always appear when a continuous symmetry is broken.[2] They have a linear dispersion relation which allows arbitrarily small excitations; Just make them with long enough wavelengths and they will be as energetically cheap as you like.

A spin wave in a small particle will experience no restoring force from the surroundings, and it is to a macroscopic spin wave what a pendulum in free fall is to a pendulum stationary on earth; no restoring force, no oscillatory motion. Obviously the pendulum will have to be small and kicked around and damped by random thermal motion, and not by a coherent, externally imposed excitation for the analogy to hold.

It is therefore not surprising that the spontaneous symmetry breaking is different for a finite system; it is simply a lot easier for a spontaneous fluctuation of the direction of the total moment to wander off if the system is small.

This Wanderlust is can be partly inhibited by the existence of domains, which is one of the features of bulk magnetism. In these domains the local magnetic moments point in the same direction which, however, differs from that of other, nearby domains. There is, fortunately, a smallest domain size. It is determined by the balance between the surface tension related to the formation of the domain walls and the bulk term

[2] Well, most of time, at least. The mechanism that gives masses to particles, named after P. Higgs, avoids this.

associated with the electrostatic energy of a homogeneous magnetic field. The result is that domains are created spontaneously with sizes typically between 10 nm and 1 μm, with the average size depending on the material. Smaller particle than this will not be able to accommodate more than one domain. Consequently there are no magnetic domain structure in nanoparticles. This simplifies matters considerably.

We will continue with disregarding not only domains, which is justified, but also the type of structural symmetry breaking that gave rise to the domains, and consider a nanoparticle a spherically symmetric system. Although this may appear like the spherical horse approximation, one can expect it to be reasonably good for small particles. If we proceed with this assumption, we will have a cluster in which the direction of the total magnetic moment can fluctuate thermally or even quantum mechanically. Without an external magnetic field, the direction of the magnetization will diffuse around on the surface of a unit sphere. Effectively, the bulk broken symmetry is restored in small particles.

This description is called the superparamagnetic model. The word paramagnetic refers to the fact that a magnetic orientation is induced by an external magnetic field, and that this orientation is parallel to the field, vs. the anti-parallel diamagnetism. In this sense a magnetic nanoparticle behaves as a giant paramagnetic atom, with two important differences. One is the magnitude of the moment, which can be considerable, but which amounts to a quantitative difference. The other is the presence of excitable internal degrees of freedom in the particle, which is a qualitative difference from an atom. It should also be remembered that, in spite of the para-prefix, the origin is still the intrinsic ferro- or ferri-magnetism of the particle which aligns the magnetic moments inside the particle.

The interaction energy of the magnetic particle in an external field is therefore described by an equation similar to (10.73) with the substitution of the correct magnetic moment,

$$E = -N\langle\overline{\mu}\rangle \cdot \overline{B}, \tag{10.78}$$

where N is the number of atoms in the particle, and $\langle\overline{\mu}\rangle$ is the average magnetic moment of a single atom along the axis of the resulting total moment. With the angle θ between the magnetic field and the total moment of magnitude μ, the energy is

$$E = -N\mu B\cos(\theta). \tag{10.79}$$

This assumes that the moments can be treated classically without concerns for quantization of the projection of the moment. For large moments this should be an acceptable procedure.

The phase space associated with the energy in (10.79) is $\sin(\theta)\mathrm{d}\theta\mathrm{d}\phi$ where ϕ is the azimuthal angle. If the particle has enough internal energy and angular momentum to exchange both of these freely with the magnetic moment, the degree of freedom associated with the magnetic moment in an external field can be described in the canonical ensemble, as discussed above and in Chap. 5 in the context of unimolecular rate constants. The canonical partition function is

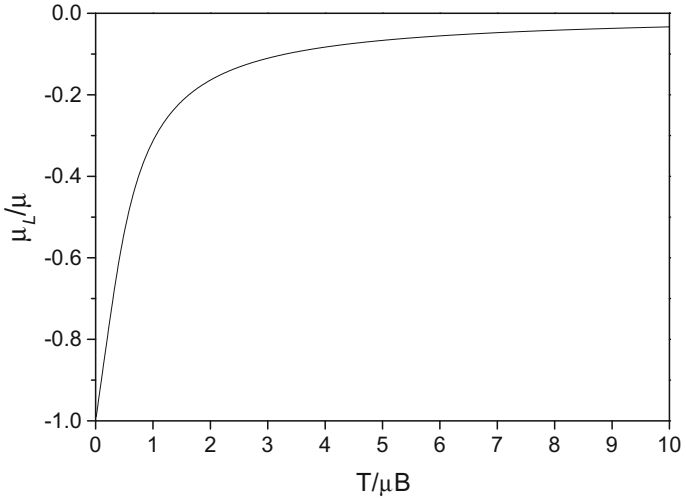


Fig. 10.12 The function $\langle \mu_L \rangle / \mu$ in (10.81) vs. the scaled temperature parameter $T/\mu B$

$$Z = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin(\theta) \mathrm{d}\theta \mathrm{e}^{N\mu B \cos\theta} = \frac{2\pi}{N\beta\mu B} \left(\mathrm{e}^{N\beta\mu B} - \mathrm{e}^{-N\beta\mu B} \right) \tag{10.80}$$

We find the average projection of the moment along the magnetic field (subscript L for laboratory) as

$$\langle \mu_L \rangle = -\frac{1}{B} \frac{\partial \ln Z}{\partial \beta} = \frac{T}{B} - \mu \frac{\mathrm{e}^{\beta\mu B} + \mathrm{e}^{-\beta\mu B}}{\mathrm{e}^{\beta\mu B} + \mathrm{e}^{-\beta\mu B}} = \mu \left(\frac{T}{\mu B} - \frac{\mathrm{e}^{\mu B/T} + \mathrm{e}^{-\mu B/T}}{\mathrm{e}^{\mu B/T} + \mathrm{e}^{-\mu B/T}} \right) \tag{10.81}$$

Figure 10.12 shows the behavior of the term in the brackets of (10.81).

Experimentally, the average moment calculated in (10.81) is measured by sending a beam of the particles through a Stern-Gerlach type of device, where a strongly inhomogeneous magnetic field deflects the particles. The deflections are minute with the available magnetic fields and the deflection angle is to a good approximation given by

$$\Theta \approx \frac{\Delta v_z}{v} \propto \frac{\partial B}{\partial z} \langle \mu_L(B, T) \rangle \frac{l}{m v^2}, \tag{10.82}$$

where l is the length of the Stern-Gerlach magnet, v is the speed of the beam from the source along the direction of the magnet, and m the mass of the particle.

10.7 Numerical Methods

In the calculation of electronic partition functions a direct summation is only possible insofar as one knows the energies and degeneracies of the electronic spectrum. Although this is generally not the case, there are a few cases of interest where the spectrum is actually known. If one is concerned with a reaction like

$$A_N \rightarrow A_{N-\Delta N} + A_{\Delta N}, \tag{10.83}$$

and ΔN is 1 or 2, the electronic excitation spectrum of $A_{\Delta N}$ will often be known from spectroscopy. Then it should be included in the calculation of the product level density, where it replaces the degeneracy factor g in (5.10, 5.12). This, at least, solves half the problem. In other (rare) situations a direct analytical calculation is possible because the spectrum can be realistically approximated with a sufficiently simple model like the equidistant single particle state spectrum, where the partition functions can be given in closed form. Truth be told, these two cases do not cover a lot of ground.

A third approach is to make a direct numerical, quantum mechanical calculation of the excitation spectrum, and use that as the input for a calculation of the level density. This use of excitation spectra is restricted to situations where the quantal structure can be reliably described with single particle states. A description of excited states as combinations of single particle excitations is not rigorously true, but in certain cases the single particle states give a good representation of the ground state and, one must expect, also of excited states. This seems to be the case for pure metals, for which the delocalized electron mean field description reproduces a number of experimental features.

Irrespective of shortfalls of rigour we proceed by using a single particle spectrum to generate the level density. The procedure we will suggest will solve your problems up to fairly high particle numbers and excitation energies. The idea is similar to the one used for calculating the partition function for the equidistant spectrum; Divide up the level density into two parts, one counting states where a given level contains an electron and one where it does not. The total level density is the sum of these two parts, because they are mutually exclusive and at the same time exhaust all the possibilities. The procedure is used recursively.

We label the levels and define a level density, called $\rho(E, n_l, N)$, which is a function of the energy E, the number of electrons N, and of the number of levels, n_l, that have been included in the calculation of that specific level density. The single particle levels are listed and numbered and the value of n_l specify precisely which levels have been included and which not. Levels with indices above n_l do not contribute to the value of $\rho(E, n_l, N)$, irrespective of the excitation energy. The labelling of levels does not need to be in order of increasing energy, although for numerical reasons it is often convenient to do so. We then have the recurrence relation:

$$\rho(E, n_l, N) = \rho(E - \varepsilon_{n_l}, n_l - 1, N - 1) + \rho(E, n_l - 1, N). \tag{10.84}$$

With a proper initialization, the function can be calculated numerically exactly only for excitation energies up to the highest level provided as input. In practice the accuracy is very good to much higher energies, though.

The calculation is initialized with the boundary condition

$$\rho(E, N, N) = \delta\left(E - \sum_{i=0}^{N-1} E_i\right), \tag{10.85}$$

which expresses that there is one way to fit N electrons into N levels, and with

$$\rho(E, n_l, 0) = 1, \tag{10.86}$$

The functions for $N > n_l$ are all zero because the Pauli principle forbids such states. The procedure is illustrated in Fig. 10.13.

The level density of the equidistant spectrum calculated with this method is shown in Fig. 10.9. Table 10.3 gives the degeneracies of the first excited states of the spectrum, which can be useful in calculations where the excitation energy is not sufficiently high to use the high energy approximations derived in Sect. 10.5. Figure 10.14 shows another application of the method, the lowest part of the C_{60} electronic excitation spectrum.

Canonical quantities can be calculated once the level density is known (re-read Chap. 1 if this is not clear), but it is possible to modify the procedure in (10.84) to calculate canonical quantities directly. It goes like this: For a given single particle spectrum, define the canonical partition functions for N electrons distributed into

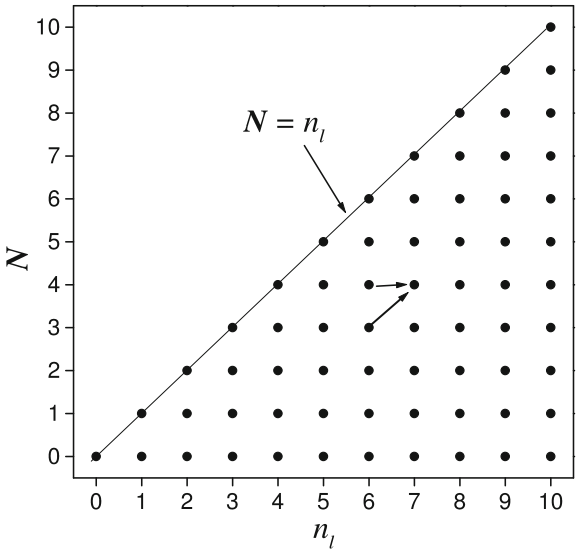


Fig. 10.13 Illustration of the computational flow in the recurrence relation in (10.84). Every dot represents a function of energy, specified with respect to N and n_l. The arrows indicate the two contributions used to calculate the target level density at the endpoint of the arrows. Only points where the level density can take non-zero values are shown

Table 10.3 Degeneracies of the fermionic equidistant spectrum. E/Δ is the scaled excitation energy, $g(1)$ is the degeneracies when all single particle levels are singly degenerate, $g(2, \text{odd})$ is the corresponding number for single particle levels with a degeneracy of two and an odd number of electrons and $g(2, \text{even})$ the same for an even number of electrons. The term in the bracket is the number calculated with the analytical method presented in (10.69), but with all terms shown in the preceding derivation included

E/Δ	$g(1)$		$g(2, \text{odd})$	$g(2, \text{even})$
0	1	(–)	2	1
1	1	(1.111)	4	4
2	2	(1.882)	12	9
3	3	(3.038)	24	20
4	5	(4.720)	50	42
5	7	(7.114)	92	80
6	11	(10.46)	172	147
7	15	(15.07)	296	260
8	22	(21.34)	510	445
9	30	(29.77)	840	744
10	42	(40.98)	1372	1215
11	56	(55.77)	2176	1944
12	77	(75.13)	3424	3059
13	101	(100.3)	5268	4740
14	135	(132.7)	8040	7239
15	176	(174.3)	12072	10920
16	231	(227.2)	17976	16286
17	297	(294.4)	26428	24028
18	385	(379.2)	38564	35110
19	490	(485.6)	55680	50844
20	627	(618.6)	79846	73010

the n_l first single particle levels, $Z(n_l, N)$, in analogy to the definition of the level densities above. We have the relation

$$Z(n_l, N) = Z(n_l - 1, N) + \mathrm{e}^{-\beta\varepsilon_{n_l}} Z(n_l - 1, N - 1), \tag{10.87}$$

that is, either the new level n_l with energy ε_{n_l} is occupied or it is unoccupied. As a technical remark, note that if the summation is performed horizontally in a n_l, N diagram, one reduces the storage needed and one can do with a few arrays of a size equal to the number of levels included in the calculation.

This algorithm is completely analogous to (10.84) if the shift in the energy argument in the equation is replaced with the Boltzmann factor. It was first suggested for the calculation of nuclear level densities where a knowledge of the excitation spectrum is of much greater importance than for most systems containing electrons,

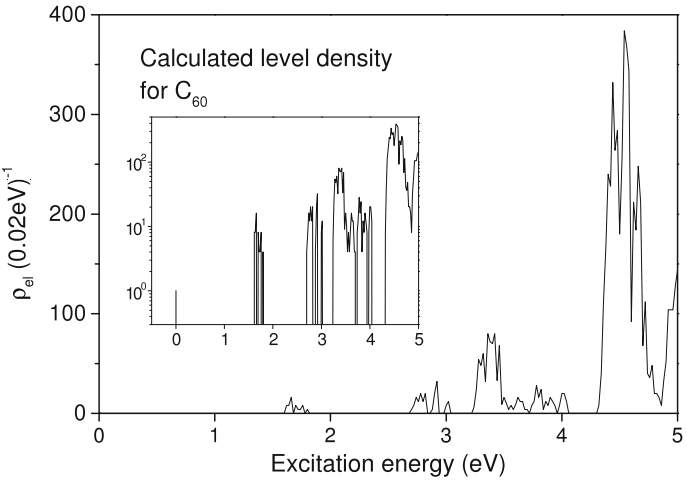


Fig. 10.14 The electronic level density of C_{60}, based on a Local Density Approximation (LDA) calculation, courtesy of K. Yabana. The inset shows the same on a logarithmic scale. At the low energies shown the curve is clearly a non-trivial function. At higher energies (not shown) the strong fluctuations disappear and the curve resembles that of a Fermi gas

because electrons usually come with an atomic nucleus (see Chap. 11 for a short-lived exception) which participate in low quantum energy vibrational excitations whereas all excitation energy in nuclei is located in excitations of the nucleons. The algorithm is occasionally rediscovered.

An equation similar to that for the partition function can be written down for the calculation of the energy integral,

$$\overline{E}Z \equiv \int E\rho(E)\exp(-\beta E)\mathrm{d}E. \tag{10.88}$$

That this integral is really equal to $\overline{E}Z$ can be seen by division by Z. To establish the recurrence relation for this quantity, we consider the recurrence relation for the average excitation energy, $\overline{E}(n_l, N)$. The weighted contribution from the $n_l - 1, N$ configuration is

$$\overline{E}(n_l - 1, N)\frac{Z(n_l - 1, N)}{Z(n_l, N)}, \tag{10.89}$$

and the contribution from the $n_l - 1, N - 1$ configuration is

$$(\overline{E}(n_l - 1, N - 1) + \varepsilon_{n_l})\frac{Z(n_l - 1, N - 1)}{Z(n_l, N)}\mathrm{e}^{-\beta\varepsilon_{n_l}}. \tag{10.90}$$

Adding the two we therefore have

$$\overline{E}(n_l, N) = \overline{E}(n_l - 1, N)\frac{Z(n_l - 1, N)}{Z(n_l, N)} \tag{10.91}$$
$$+ \left(\overline{E}(n_l - 1, N - 1) + \varepsilon_{n_l}\right)\frac{Z(n_l - 1, N - 1)\mathrm{e}^{-\beta\varepsilon_{n_l}}}{Z(n_l, N)}.$$

This can be rearranged to express it in terms of the energy integral $\overline{E}Z$, in a form that may be more convenient:

$$\overline{E}Z(n_l, N) = \overline{E}Z(n_l - 1, N) + \left[\overline{E}Z(n_l - 1, N - 1) + \varepsilon_{n_l}Z(n_l - 1, N - 1)\right]\mathrm{e}^{-\beta\varepsilon_{n_l}}. \tag{10.92}$$

The average canonical thermal energy is then

$$\langle E\rangle = \frac{\overline{E}Z(n_l(\max), N)}{Z(n_l(\max), N)}, \tag{10.93}$$

where $n_l(\max)$ is the number of levels included. From the thermal energy and the partition function, also the entropy can be calculated.

The algorithm is very efficient if only a few temperatures need to be calculated, which you see from the fact that the arrays have one index less than the similar algorithm for the level densities, corresponding to one less loop in a code.

We can also find the average square of the energy, which is used to calculate heat capacities, by defining the analogous function $\overline{E^2}Z$ and replacing E with E^2 in the preexponentials in (10.91). After rearranging we have

$$\begin{aligned}\overline{E^2}Z(n_l, N) = \overline{E^2}Z(n_l - 1, N) + \big[&\overline{E^2}Z(n_l - 1, N - 1) \\ &+ 2(\varepsilon_{n_l} - \varepsilon_N)\overline{E}Z(n_l - 1, N - 1) \\ &+ (\varepsilon_{n_l} - \varepsilon_N)^2 Z(n_l - 1, N - 1)\big]\mathrm{e}^{-\beta(\varepsilon_{n_l} - \varepsilon_N)}.\end{aligned} \tag{10.94}$$

The implementation of both of these algorithms requires only a few additional lines to the code used to calculate Z.

One useful computational feature which is pays to implement should be mentioned. It is related to the ground state energy. This energy changes every time an electron is added to the system. One prefers to keep the ground state energy equal to zero to avoid over- and underflow problems in the computer, which are potentially very severe when exponentials of numerically large numbers are involved. This is accomplished by adjusting (10.87, 10.92) to the following:

$$Z(n_l, N) = Z(n_l - 1, N) + Z(n_l - 1, N - 1)\mathrm{e}^{-\beta(\varepsilon_{n_l} - \varepsilon_N)}, \tag{10.95}$$

and

$$\overline{E}Z(n_l, N) = \overline{E}Z(n_l - 1, N) + \tag{10.96}$$
$$\left[\overline{E}Z(n_l - 1, N - 1) + (\varepsilon_{n_l} - \varepsilon_N)Z(n_l - 1, N - 1)\right]\mathrm{e}^{-\beta(\varepsilon_{n_l} - \varepsilon_N)}.$$

After we have seen that we can calculate the thermally average of the zeroth, first and second moment of the energy in the canonical ensemble, it is clear that we can calculate whatever moment we like. The next extension of the method is to calculate whatever quantity associated with the state of the system, i.e. properties that can be described by the single particle wave functions. Consider a generic quantity q. Every single particle state has the associated value q_i. It can be energy as above, the density at some point in space or the root-mean-square extension of the wave function. The recurrence proceeds as

$$\overline{q}Z(n_l, N) = \overline{q}Z(n_l - 1, N) + \left[\overline{q}Z(n_l - 1, N - 1) + q_{n_l}Z(n_l - 1, N - 1)\right] \mathrm{e}^{-\beta\varepsilon_{n_l}}. \tag{10.97}$$

10.8 Electronic Shell Structure

The picture of metal particles as containers of free, independent electrons is not a sufficient description of levels across the whole metallic section of the periodic table. For some elements, mainly the alkali metals, it is nevertheless a very good starting point for a description, also when one wants to understand the effects of thermal electronic excitations. This is true both for calculations of situations where the equidistant spectrum is relevant and when electronic shell structure is.

Electronic shell structure is manifested experimentally in several ways. One is sawtooth-like variations with size of the ionization energy, another is a systematic variation of the chemical reactivity with size, and yet another are particularly high abundances for particles with certain number of valence electrons. For the alkali metals these high abundances coincide with the shell closings, corresponding to the so-called 'magic numbers' which are 2, 8, 20, 40, 58, 92,…. The increased abundances for clusters with these electron numbers is a direct consequence of the increased binding energy induced by the shell structure, by the effects derived in Chap. 7.

The origin of the shell structure, in turn, is the quantization of the motion of the valence electrons in a spherical confining potential. There is a certain amount of degeneracies associated with a spherical potential. As a start, one has the angular momentum degeneracy one learns about in introductory quantum mechanics courses. But the degeneracy is actually higher. The reasons for the higher degeneracy, which is only approximate, can be understood with reference to semiclassical quantization. This is another of these interesting phenomena we must sadly refer to another day. For our purposes it is sufficient to know that shell structure is associated with a bunching of single particle levels and that a shell can contain a large number of single particle states. When a shell is filled all states in a bunch are occupied and an added electron must to go into a new shell with an energy which is higher than the average HOMO energy. This gives a higher than average binding energy of the latest electron added to the shell, and a lower than average binding energy for the first electron to be added to a shell.

The picture is very similar to the odd-even effect, with the difference that shell structure involves more than two electrons per shell. And for the same reasons that even numbered clusters are more stable toward loss of an atom, i.e. has a higher dissociation energy, closed shell clusters will also have a higher than average dissociation energy, and vice versa for the cluster with one or a few electrons in the next shell.

The shell structure thus causes the binding energies to vary non-monotonically with size. The variation is summarized by the shell energy, $E_{shell,N}$, which is added to the bulk and surface contributions,

$$E_N = -AN + BN^{2/3} + E_{shell,N}, \tag{10.98}$$

as discussed in Chap. 8. By definition, $E_{shell,N}$ oscillates around zero and averages to zero when averaging over a sufficiently wide range of sizes. An odd-even effect may also be present. It does not appear in the calculations used as input in the example given below, though.

The shell energy has an approximately parabolic shape between two shell closings. This is a consequence of the general trend of a single particle energy level to decrease when the size of the confining potential increases and the depth of the potential remains unchanged, described in (10.1). To understand this quantitatively, make the simplification that a shell is a single, highly degenerate level with energy $E_q(n)$. This level decreases smoothly with the total number of atoms in the particle and this size dependence can be linearized as a function of the electron number, n,

$$E_q(n) \approx E_q(0) + \frac{\mathrm{d}E_q(0)}{\mathrm{d}N} n, \tag{10.99}$$

with $dE_q(0)/dN < 0$. The total energy of the electrons in the shell is then the occupation number times the energy, or

$$E_{q,tot} = nE_q(n) \approx nE_q(0) + n^2 \frac{\mathrm{d}E_q(0)}{\mathrm{d}N}. \tag{10.100}$$

Because the second order term is negative, the shell closings have the (locally) lowest binding energy. This gives the shell energy (schematically) as a function composed of a number of parabolas with negative curvature, joined at the shell closings.

This does not reflect reality entirely truthfully, because an unfilled shell will deform away from the spherical shape and remove the degeneracy by the action of the Jahn-Teller effect. The level splitting caused by this deformation may be significant, and for alkali metal clusters the deformation acts to reduce the shell energy by a factor of two, mainly affecting the mid-shell particles. We will stick with the spherical particles here in order to illustrate the finite temperature effects.

These considerations can be made more quantitative and are borne out by calculations, both with Density Functional Theory (DFT) and with the simpler type of calculations where the Schödinger equation is solved in a mean field potential. The result of one such calculation is shown in Fig. 10.15, together with excitation

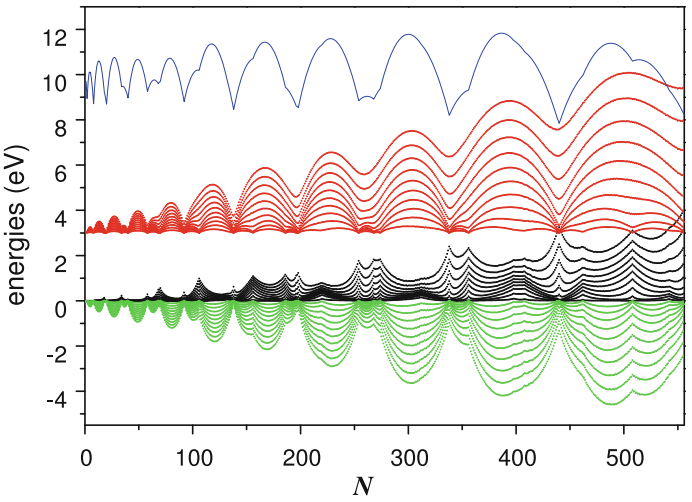


Fig. 10.15 The thermal properties of electronic shell structure of small and medium size sodium clusters. The curves are, from top to bottom, **a** the shell energy as defined in (10.98), shifted upward by 10 eV for display purposes, **b** the entropy times temperature for the temperatures $k \times 100$ K, $k = 1, 2, \ldots, 10$, shifted upward by 3 eV, **c** the thermal energy at the same temperatures and **d** the electronic free energy, $F = E - TS$. The free energy has its zero at the ground state of the particle in question. The shell energy is calculated in H. Nishioka, K. Hansen and B. R. Mottelson, *Phys. Rev. B* **42** (1990) 9377

energies and entropies, calculated with (10.95, 10.96) using the single particle levels calculated with a mean field potential of the Saxon-Woods type. The potential has the form

$$V(r) = -\frac{V_0}{1 + \exp((r - r_N)/a)}, \tag{10.101}$$

where, in this application for sodium clusters, the parameters are $V_0 = 6$ eV, $r_N = r_1 N^{1/3}$, $r_1 = 2.25$ Å and $a = 0.74$ Å. The potential is basically a spherical box with a soft edge that interpolates smoothly from $-V_0$ to zero over the distance a.

The top curve in Fig. 10.15 is the ground state shell energy, $E_{shell,N}$, with the concatenated parabolas clearly visible. The entropy at low temperature and/or small particle sizes is strongly correlated with the shell energy (top and next-to-top curves). At high temperatures and/or large sizes this correlation tends to be washed out. The correlation appears because the open shell particles, those with a number of valence electrons between shell closings, have a degenerate or nearly degenerate highest occupied level; there is room for N_s electrons in the shell but it is only occupied with less than N_s electrons. This gives rise to a non-zero entropy already at zero temperature. If the states in a shell were completely degenerate, which is not the case for this potential, the entropy would be given by binomial coefficients

$$S = \ln \left(\begin{array}{c} N_s \\ n \end{array} \right), \tag{10.102}$$

which peaks at $n = N_s/2$ with the value $S = N_s \ln 2$. Hence the term TS in the free energy is largest mid-shell and almost vanishes at the shell closings at low temperatures.

The excitation energy has a more complicated behavior. At low temperatures the mid-shell particles are the ones where electrons can be excited thermally because they have a spectrum composed of a number of closely spaced levels. When the temperature is increased, the excitation energy of these mid-shell particles will increase slower because the empty states in the shell are already explored and it requires a relatively costly and therefore rare excitation of an electron to a new shell to increase the average excitation energy. At the temperatures where electrons in mid-shell particles begin to be excited, also electrons in the closed shells will begin to get excited. The gain in entropy is higher for closed shell particles than for mid shell particles and the degree of excitation will therefore be higher for electrons in closed shell particles.

At very high energies, the shell structure in the entropy is completely washed out and only the excitation energy has any shell structure. This structure matches the shell energy but has the opposite sign. This gives us a high temperature free energy with a shell structure which is a mirror image of the shell energy. Because this free energy is measured relative to the electronic ground state of the respective particles, the use of a common zero of energy for all particles will then give an almost vanishing shell free energy at high temperatures, as expected intuitively.

10.9 An Excursion into Bose-Einstein Statistics

The properties of strongly degenerate electron systems depend on the spin of the electrons in a crucial way, because the spin determines the statistics the particles obey. Bosons (particles with integer spin) will behave completely differently from fermions because one quantum state can accommodate more than one particle, infinitely many, in fact. We can calculate thermal properties of bosonic systems with a method very similar to the one used for fermions, with similar reservations on the description of the total energy in terms of single particle levels. The results are not at all directly useful for the description of excited valence electrons, but the idea has so many parallels with the fermionic procedure that it makes sense to present it here.

With the label n_l for the number of levels included in the sum and N the number of bosons, we calculate canonical partition functions as

$$Z(n_l, N) = \sum_{p=0}^{N} Z(n_l - 1, N - p) \mathrm{e}^{-p\beta\varepsilon_{n_l}}, \tag{10.103}$$

with $Z(n_l, 0) = 1$. As for fermions, this equation simply expresses that if a boson is not in state n_l, it will be in one of the other single particle states. The relation seems to imply a lot more work for the computer because the summation runs over levels and particle numbers, as for fermions, but in addition also over the partitioning of particles in the latest added single particle state. We can eliminate one of these sums with a rewrite of (10.103):

$$\begin{aligned} Z(n_l, N) &= Z(n_l - 1, N) + \sum_{p=1}^{N} Z(n_l - 1, N - p)\mathrm{e}^{-p\beta\varepsilon_{n_l}} \qquad (10.104) \\ &= Z(n_l - 1, N) + \sum_{p=0}^{N-1} Z(n_l - 1, N - 1 - p)\mathrm{e}^{-p\beta\varepsilon_{n_l}}. \end{aligned}$$

Comparison of the last sum in this equation with (10.103) shows us that it is equal to $Z(n_l, N - 1)$. We therefore have the equivalent but computationally much less demanding recurrence relation

$$Z(n_l, N) = Z(n_l - 1, N) + \mathrm{e}^{-\beta\varepsilon_{n_l}} Z(n_l, N - 1). \qquad (10.105)$$

The flow of this equation and (10.103) in an (N, n_l) diagram is illustrated in Fig. 10.16.

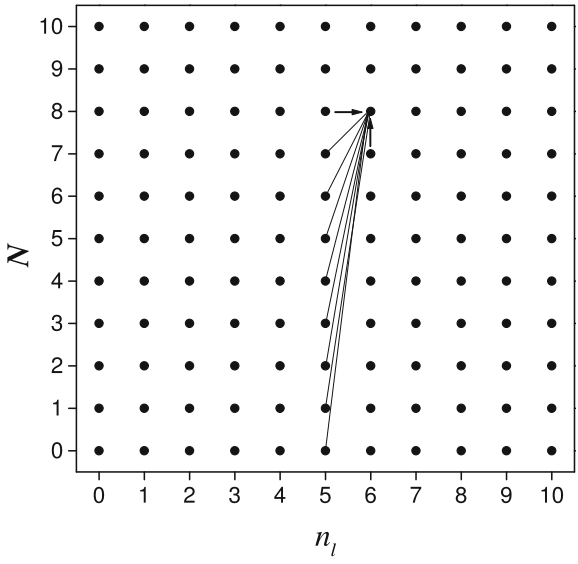


Fig. 10.16 Illustration of the recurrence relation used to calculate bosonic level densities and partition functions. The two arrows represent the flow in the simple relations in (10.105, 10.107). The lines originating at $n_l = 5$ without arrows indicate the flow of computation in (10.103)

A similar calculation for the level density gives

$$\rho(E, n_l, N) = \sum_{p=0}^{N} \rho(E - p\varepsilon_{n_l}, n_l - 1, N - p). \tag{10.106}$$

As for the partition function, terms with $p > 0$ on the right hand side can be summed to a single term,

$$\rho(E, n_l, N) = \rho(E, n_l - 1, N) + \rho(E - \varepsilon_{n_l}, n_l, N - 1). \tag{10.107}$$

Just like for fermions, one can write down recurrence relations that can be used to calculate arbitrary moments of the excitation energy, and anything else that is contained in the single particle wave functions, as we have seen. We will only show how to calculate the thermal energy. The relation reads

$$\overline{E}(n_l, N) = \sum_{p=0}^{N} \overline{E}(n_l - 1, N - p) \frac{Z(n_l - 1, N - p)}{Z(n_l, N)} \mathrm{e}^{-p\beta\varepsilon_{n_l}}. \tag{10.108}$$

Also this relation is similar to the fermionic relation, the difference being the summation over all possible partitioning of atoms into the most recently added state. Use of (10.108) for $\overline{E}(n_l, N - 1)$ and (10.103) for $Z(n_l, N - 1)$ again allows us to collect terms and gives the more compact relation

$$\begin{aligned} \overline{E}(n_l, N) = {} & \overline{E}(n_l - 1, N) \frac{Z(n_l - 1, N)}{Z(n_l, N)} \\ & + \mathrm{e}^{-\beta\varepsilon_{n_l}} \left(\overline{E}(n_l, N - 1) \frac{Z(n_l, N - 1)}{Z(n_l, N)} + \varepsilon_{n_l} \frac{Z(n_l, N - 1)}{Z(n_l, N)} \right). \end{aligned} \tag{10.109}$$

When we multiply with $Z(n_l, N)$ we get the more convenient expression:

$$\overline{E}Z(n_l, N) = \overline{E}Z(n_l - 1, N) + \mathrm{e}^{-\beta\varepsilon_{n_l}} \overline{E}Z(n_l, N - 1) + \varepsilon_{n_l} \mathrm{e}^{-\beta\varepsilon_{n_l}} Z(n_l, N - 1). \tag{10.110}$$

Together with the calculation of Z with (10.105), this permits calculations of excitation energies and entropies for the systems.

An interesting special case exists for which one can calculate the canonical partition function for bosons explicitly. That case is none other than the ladder spectrum with the singly degenerate single particle levels $\varepsilon_k = k\Delta$, $k = 0, 1, 2, \ldots \infty$. The partition function can be decomposed into contributions for no bosons in the single particle ground state, one boson in the ground state and so on, up to all N, in a procedure that should be familiar by now. The partition function for p bosons in the ground state is $Z(N - p) \exp(-(N - p)\beta\Delta)$, which just corresponds to a shift in energy of the $Z(N - p)$ partition function. We therefore have

$$Z(N) = \sum_{p=0}^{N} Z(N-p)\mathrm{e}^{-(N-p)\beta\Delta} = Z(N)\mathrm{e}^{-N\beta\Delta} + \sum_{p=1}^{N} Z(N-p)\mathrm{e}^{-(N-p)\beta\Delta}. \tag{10.111}$$

The last sum can be rewritten as

$$\sum_{p=1}^{N} Z(N-p)\mathrm{e}^{-(N-p)\beta\Delta} = \sum_{p=0}^{N-1} Z(N-1-p)\mathrm{e}^{-(N-1-p)\beta\Delta}, \tag{10.112}$$

If we compare this with the first equality sign in (10.111) we see that this is equal to $Z(N-1)$. Thus,

$$Z(N) = Z(N)\mathrm{e}^{-N\beta\Delta} + Z(N-1). \tag{10.113}$$

This is exactly the same recurrence relation as for fermions for the same single particle spectrum, and because $Z(1)$ is the same for a single fermion and a single boson, also the solution is of course identical to the fermionic case:

$$Z(N) = \prod_{k=1}^{N} \left(1 - \mathrm{e}^{-k\beta\Delta}\right)^{-1}. \tag{10.114}$$

Identical partition functions for all temperatures implies identical level densities for all energies. This seems impossible at first sight, because the population of levels is very different in the two cases, fermions stack on top of each other and bosons collect in the ground single particle state at low temperatures. The explanation is that the level density is that of the excitations, counted from the ground state of the system, whichever of the two different kinds of systems it may happen to be.

Exercises

10.1 Consider a metallic particle as a box potential with a flat bottom and a steep edge for the sake of describing the motion of the electrons. Quantum mechanically the electrons will penetrate the classically forbidden regions of the potential near the edge. Calculate an estimate of the length of this penetration into this region. Use this to estimate corrections to (10.1). Use the example values 10 eV for the depth of the potential, $r_1 = 2$ Å, and levels half way between the bottom of the potential and the vacuum level.

10.2 Calculate the variance of the electron number in the grand canonical ensemble by summation of the contributions from the variances of the individual levels and show that it is identical to the result in (10.17).

10.3 Calculate the thermal energy of a 1d gas of photons in a cavity of length L. Compare with (10.31).

10.4 Use the Sommerfeld expansion in (10.36) to find the leading order temperature dependence of the chemical potential of a Fermi gas in (10.37). Hint: Calculate the number of electrons with an integral of the density of states; next approximate the integral with the value up to E_F and a term linear in the difference $E_F - \mu$. Finally, set the calculated change to zero and solve for μ.

10.5 Show that (10.40, 10.42) are right, to the accuracy stated.

10.6 Equation (10.25) gets a contribution from the variation of Δ with size. Show that this correction is of second order in $1/N$.

10.7 Calculate the high temperature level density of an odd or even electron number of particles in the equal spacing spectrum described in Sect. 10.3. Use the equations that were applied in Sect. 10.4. Compare with the results of that section, taking into account that the two Δ's differ by a factor 2. Compare your result with the exact low energy values in Table 10.3.

10.8 Show that (10.18, 10.20) are correct. For the canonical result you sum the probabilities of all relevant states and divide by the partition function. For the grand canonical calculation you use the fact that the occupation of the individual states are independent. This renders the expression for the sum of relevant states simple. Calculate also the probability that no electron is excited in the two ensembles.

10.9 Integrate the relation

$$\langle E \rangle \propto T^2 \tag{10.115}$$

to show that the canonical partition function for the Fermi gas in the equidistant spectrum, (10.29), is approximately exponentially increasing with temperature.

10.10 Sketch the low temperature values of the excitation energy, Helmholtz' free energy and the entropy of the system with the partition function in (10.29).

10.11 For independent particle motion in spherical potentials, the electrons' energy eigenstates are characterized by a radial and an angular momentum quantum number, and in addition a spin and a projection of the angular momentum. Find the possible combinations of angular momenta that form the shell closings from 8 to 20 valence electrons, from 20 to 40, from 40 to 58 and from 58 to 92. You can assume that a specific ℓ value appears only once in one shell.

10.12 **(a)** Calculate the relative number of excited electrons in a Fermi gas by integration of the Fermi-Dirac distribution. Assume that the chemical potential is equal to the Fermi energy. In the first approximation you can assume that the single particle level density is constant. Give a physical interpretation of the result.
(b) Next consider the situation where the single particle density of states is not constant but still only varies slowly over a range of energies comparable to the temperature. Show that the three-dimensional Fermi gas fulfils this criterion, and show, without an explicit integration, that the number of excited electrons calculated in **(a)** is corrected approximately with the factor $g_0 E_F$, where g_0 is the density of states at the Fermi level.

10.13 Show that the low-temperature, non-zero solution of the mean field equation for the electron spin magnetization in (10.76) is below the maximum value of $1/2$.

10.14 Below the critical temperature, there are two solutions to the mean field equation of magnetization, one magnetized and one unmagnetized. Show that the magnetized state is the stable phase in this situation.

10.15 Show that (10.79) has the expected behavior in the limits of $T \to 0$ and $T \to \infty$.

Chapter 11
Hot Electron Reactions

All reactions described in this book so far have been equilibrium processes, or at least so close to equilibrium that equilibrium relations can be applied without any significant loss of precision. But interest in thermal properties are not restricted to situations of complete equilibrium. After all, complete equilibrium means that nothing happens. The description of non-equilibrium states and processes adds a layer of complications relative to the equilibrium description, complications that often prevent a simultaneously simple and accurate description. Exceptions exist, and for electronic degrees of freedom, non-equilibrium can give rise to the special effects that are the subject of this chapter.

When electrons in metallic particles, defined broadly as a system where valence electrons are quasi-free, are excited on time scales shorter than the vibrational period, the energy may equilibrate internally within the electron system by electron-electron collisions and reside there for some time before it is dissipated into vibrations. For this short time, the particle is in a transient state best described as a free electron gas of hot electrons made up of the valence electrons, coexisting with fairly cool vibrations. In solid state physics this situation is described with the aptly named two-temperature model.

The existence of a hot electron state is possible because the time scales for the motion of the electrons and the vibrations are very different. The typical electronic time scales are given by $\hbar/1\ \mathrm{Ry} = 49$ as (1 as = 10^{-18} s), and vibrational periods by the reciprocal Debye frequency as $2\pi/\omega_D$, which varies with material (see Chap. 4) but is around 10–100 fs, i.e. three orders of magnitude longer than the time scale for electrons. The situations described in this chapter depend crucially on this separation.

The phenomenon can be studied with pump-probe experiments with femtosecond lasers or, for free particles, by thermal electron emission in which the electron kinetic energy is used as a thermometer. Experimentally, temperatures of a couple of tens of thousands of kelvin have been seen in the Boltzmann distribution of electrons emitted from such systems. Electron-phonon coupling times can be measured this

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_11

way and one gets values on the picosecond and subpicosecond time scales, with the precise numbers depending on the material. The values are generally a few times the typical vibrational period.

11.1 The Initial Excitation

The idea of a transient hot electron phase of a free molecule was originally proposed to account for the Penning ionization yields of fullerenes. Penning ionization is the process in which a rare gas atom (Rg) in a metastable but long-lived excited state with well defined energy collides with another molecule (M), usually in a low energy, often thermal collision, and transfers all its electronic excitation energy to the target molecule which then undergoes ionization:

$$Rg^* + M \rightarrow Rg + M^+ + e^-. \tag{11.1}$$

The molecules in the experiments that provoked the model were the fullerene molecules C_{60} and C_{70}, actors that tend to appear whenever a mass-selected, single-atom composition molecule is desired in an experiment. The metastable state energies of the rare gas atoms in the experiment ranged from slightly above the ionization energies of 7.6 to 20 eV (for helium). The excited state energy of helium is enough to both ionize and dissociate a fullerene, the latter process causing the emission of a C_2 molecule if it happens, but the ions were always measured to be unfragmented. This is one signature that will carry over to other means of excitation, as long as they do not impart too high energies. The term 'too high' is in this context several times the dissociation energy (ca. 50 eV vs. the dissociation energy of 11 eV for C_{60}). Figure 11.1 shows the experimental data together with the fitted yields from the hot electron model described in quantitative detail below.

Another mode of excitation is collisions with high speed or highly charged ions. Fast and highly charged ions have the potential to excite the collective motion of valence electrons known as the surface plasmon resonance in flyby encounters. The surface plasmon is the oscillation of the valence electrons against the positively charged ionic cores. The surface plasmon, which can be imagined as being build by single particle-hole excitations, will dephase and decay after a few oscillations into a collection of these electronic excitations. This is the starting point for further electronic equilibration. In one of the many similarities between nanoscale physics and nuclear physics, the plasmon has its nuclear analogue, called the giant dipole resonance. It arises as the oscillation of the positively charged protons and the neutral neutrons against each other.

Most situations relevant for this chapter involve absorption of light. The reason is that when a particle absorbs energy from an electromagnetic wave, the energy usually generates electronic excitations. Radiation can also be absorbed by optically active vibrations, but usually these only carry a small fraction of the oscillator strength of the whole molecule. Visible or UV light absorption is therefore a very efficient method

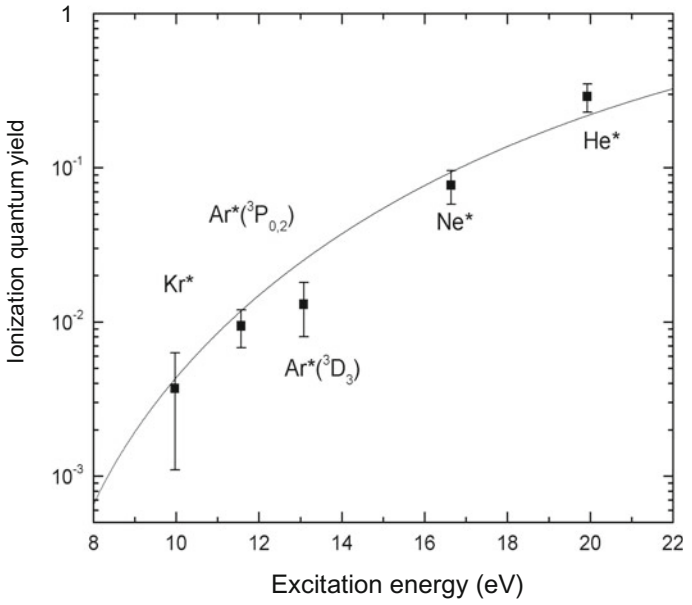


Fig. 11.1 The measured absolute quantum yield in Penning ionization of C_{60}. The identity of the rare gas atoms, and for Ar the quantum states, carrying the specific excitation energies are indicated. The line is a fit with the hot electron model described below. The data are published in J.M. Weber, K. Hansen, M.-W. Ruf and H. Hotop, *Chem. Phys.* **239** (1998) 271

to generate electronic excitations selectively. Depending on the photon energy, one or several photons may be needed to provide enough excitation energy to develop fully the characteristics of the hot electron phase.

Absorption of photons during a short time interval may activate one or more of the several alternative photo-physical reaction paths seen in molecules, such as resonant ionization, photon absorption followed directly by internal vibrational relaxation (IVR), which is the process whereby the electronic energy is dissipated into vibrational motion, or Above Threshold Ionization (ATI), which is a direct multiphoton ionization process seen in atoms and some molecules.

The creation of the hot electron phase is therefore not a given, and the precise mechanism that converts e.g. a single photon to incoherent electronic energy without causing direct ionization is still unknown at the time of writing. For a multiphoton absorption situation it is clear that the photons must arrive within a time interval shorter than the electron-phonon coupling time. But under some conditions photons are absorbed and an electronically excited state with energy above the ionization energy is in fact created in competition with the other processes mentioned.

Let's therefore start from the point where electrons are excited and consider the initial relaxation of the electronic excitations into a hot Fermi gas. Enough is known about the behavior of a Fermi gas to allow some fairly reliable conclusions about this process.

A highly excited electron gas will experience collisions at a rate of

$$\tau_{ee}^{-1} \sim n\sigma_{ee}v_F, \tag{11.2}$$

where n is the concentration, σ_{ee} the electron-electron collision cross section, and v_F the Fermi velocity. For the collision cross section we can use the Thomas-Fermi screening length, λ_{TF}, calculated for a Fermi gas,

$$\left(\frac{2\pi}{\lambda_{TF}}\right)^2 = \frac{e^2}{\varepsilon_0}\frac{3n}{2E_F} \Rightarrow \sigma_{ee} = \pi\lambda_{TF}^2 = \frac{2\pi^2}{3}\frac{1}{\frac{e^2}{4\pi\varepsilon_0}}\frac{E_F}{n}. \tag{11.3}$$

With the Fermi gas relation between density and Fermi energy,

$$E_F \equiv \frac{1}{2}mv_F^2 = \left(3\pi^2 n\right)^{2/3}\frac{\hbar^2}{2m}, \tag{11.4}$$

the electron-electron time constant becomes

$$\tau_{ee}^{-1} \sim \frac{2\sqrt{2}\pi^2}{3}\frac{1}{\frac{e^2}{4\pi\varepsilon_0}}\frac{E_f^{3/2}}{\sqrt{m}}. \tag{11.5}$$

For the moderately small value of 5 eV, the value is

$$\tau_{ee} = 0.033 \text{ fs}, \qquad E_F = 5 \text{ eV}. \tag{11.6}$$

Another estimate of the *e-e* equilibration time is given by what is known as Landau's theory of Fermi liquids. The theory goes beyond the scope of this work, but the bottom line for the question of relaxation time is that the theory gives a value on the order of

$$\tau \sim \frac{\hbar}{E_F}. \tag{11.7}$$

For most elements this time is somewhat longer than the estimate in (11.5). The difference is hardly observable, though, unless you perform a dedicated experiment on the sub-fs time scale, and it will play little role in the situation described in the following.

11.2 Decay of the Hot Electron Phase

The decay of the hot electron phase proceeds via the dissipative coupling to the systems vibrational degrees of freedom, with a coupling time on the order of a vibrational period. For a vibrational quantum energy of 0.03 eV, corresponding to

the Debye energy of a number of crystals (see Table 4.1), this corresponds to a coupling time of

$$\tau_{ep} \sim \frac{\hbar}{0.03\ \text{eV}} 2\pi \approx 140\ \text{fs}. \tag{11.8}$$

A proxy for such time constants are measured in pump-probe (p-p) experiments where a strong probe pulse excites the particle and the signal, typically in the form of an ionization yield, is measured after the application of a variable delay probe pulse. The decrease of the signal vs. the p-p delay can then be fitted with a time constant. Some caution should be exercised in the identification of this time with the coupling time. A straightforward identification of the two is possible when the dynamics involves two levels; the ground state and the decaying excited state, but this is manifestly not the case here. Keeping the caveat in mind, some measured values are 200 fs for C_{60} and several values around or slightly above 1 ps for sodium clusters. The coupling time can be inferred from fits of other types of data, of which some examples are shown below.

The fact that the two consecutive equilibration times are so dissimilar;

$$\frac{\tau_{ee}}{\tau_{ep}} \ll 1, \tag{11.9}$$

is the crucial inequality that allows the existence of the phase of highly thermally (incoherently) excited electrons coexisting with an essentially cold background of vibrations. When the excitation of the electrons is achieved by sequential absorption of multiple photons, τ_{ph} is then the limit of the width of the light pulse or, more precisely, on the time interval during which a sufficient number of photons are absorbed to induce the reactions specific to the hot electron phase. Pulses longer than τ_{ep} will also work at correspondingly higher photon fluxes, but in those cases the hot electrons will be accompanied by warmer vibrational degrees of freedom and an excessive amount of fragmentation.

The electron-phonon collision is the elementary step in the final equilibration of the excitation energy. After dissipation into the vibrational motion, with the occasional excitation into some electronic state or other, no further dissipation or equilibration will take place before the particle interacts with the environment, either radiatively, by colliding with a gas molecule or by hitting a surface in a vacuum chamber.

The time constants calculated for the electron-phonon coupling and the electron-electron coupling in the previous section represent the two elementary time steps in the relaxation. Several of these elementary steps are needed to achieve a, say, 1/e degree of relaxation, and the time constants cited are therefore lower limits. The precise number of these steps will be influenced by the energy transfer efficiency in the collisions.

The time scales for the processes and phases can then be summarized schematically as in Fig. 11.2.

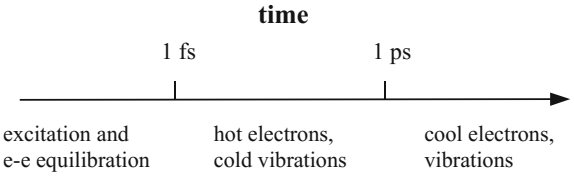


Fig. 11.2 The separate regimes of the hot electron phenomenon, from the unspecified excitation process to the completely equilibrated phase that prevail after a time of about one picosecond. The time markers are illustrative only. Values for different systems will differ

The observable consequences of the hot electron phase are strongly dependent on the degree of excitation; The observable signals of interest here are generated by the high energy regime that exist before a significant part of the energy has leaked into the vibrations. Reactions usually have rates that are strongly dependent on the energy. It is therefore sufficient to describe the decrease of electronic excitation energy in the initial stage of equilibration, and the decrease of energy with time can therefore be approximated with an exponential, irrespective of the precise shape of the curve at lower energies. Hence we write

$$\frac{\mathrm{d}E_e}{\mathrm{d}t} \approx -\frac{E_e}{\tau_{ep}}. \tag{11.10}$$

The equation does not predict correctly the ultimate fate of the energy because even at equilibrium a small amount is still residing in the electrons. We correct (11.10) for this by setting the rate proportional to the difference in energy from the mean value of the equilibrium distribution, E_e^{eq}:

$$\frac{\mathrm{d}E_e}{\mathrm{d}t} = -\frac{1}{\tau_{ep}}\left(E_e - E_e^{eq}\right). \tag{11.11}$$

This is easily solved to

$$E_e = E_e^{eq} + (E - E_e^{eq})\mathrm{e}^{-t/\tau_{ep}}, \tag{11.12}$$

with E is the initial (total) excitation energy. When the excitation energy stored in the electronic motion is small compared with the corresponding energy for vibrations,

$$E_e^{eq} \ll E_{ph}^{eq}, \tag{11.13}$$

which is always the case in applications, (11.11) is well approximated by (11.10) at early times.

An alternative parametrization of the relaxation uses the temperature;

$$\frac{\mathrm{d}T_e}{\mathrm{d}t} = -\frac{T_e}{\tau_{ep,T}}. \tag{11.14}$$

When the Fermi gas description is sufficiently close to the physical situation, the description of the coupling time is essential identical to the parametrization in terms of the energy because $E_e = \alpha T_e^2$:

$$\frac{\mathrm{d}E_e}{\mathrm{d}t} = -\frac{1}{\tau_{ep}} E_e \Rightarrow \frac{\mathrm{d}T_e}{\mathrm{d}t} = -\frac{T_e}{2\tau_{ep}}. \tag{11.15}$$

Hence the two descriptions of the dissipation are identical and the time constants related as

$$\tau_{ep,T} = 2\tau_{ep}. \tag{11.16}$$

To get an idea of the equilibration/dissipation process it is necessary to specify the thermal properties of both the electronic and vibrational subsystems. For the electrons we can use the Fermi gas expression (ignoring the pre-exponential factor of energy to the power of minus one for simplicity):

$$\rho_e(E) = a_e \mathrm{e}^{\sqrt{aE_e}}, \tag{11.17}$$

and for the vibrations the classical harmonic oscillator expression for s oscillators;

$$\rho_{ph} = a_{ph} E_{ph}^{s-1}. \tag{11.18}$$

The two microcanonical temperatures are

$$T_e = 2\sqrt{\frac{E_e}{a}}, \quad T_{ph} = \frac{E_{ph}}{s-1}, \tag{11.19}$$

and the equilibrium electronic energy is found from equating these two temperatures, using $E = E_{ph} + E_e$;

$$\sqrt{E_e^{eq}} = -\frac{s-1}{\sqrt{a}} + \sqrt{E + \frac{(s-1)^2}{a}}. \tag{11.20}$$

For an $N = 20$ sodium cluster with $E_F = 3$ eV, the average gap at the Fermi level is $\Delta = 2E_F/3N = 0.1$ eV, corresponding to $a = 66$ eV^{-1} (see (10.72)), and the high temperature harmonic oscillator heat capacity is $s - 1 = 53$. This gives $E_e^{eq} = 5.9 \cdot 10^{-3}\mathrm{eV}^{-1}E^2$. The equilibration between electrons and vibrations is illustrated for this cluster in Fig. 11.3 for the initial excitation energy of 20 eV and the excitation energies and temperatures given above. Also the entropies are shown. They are calculate for the electrons as

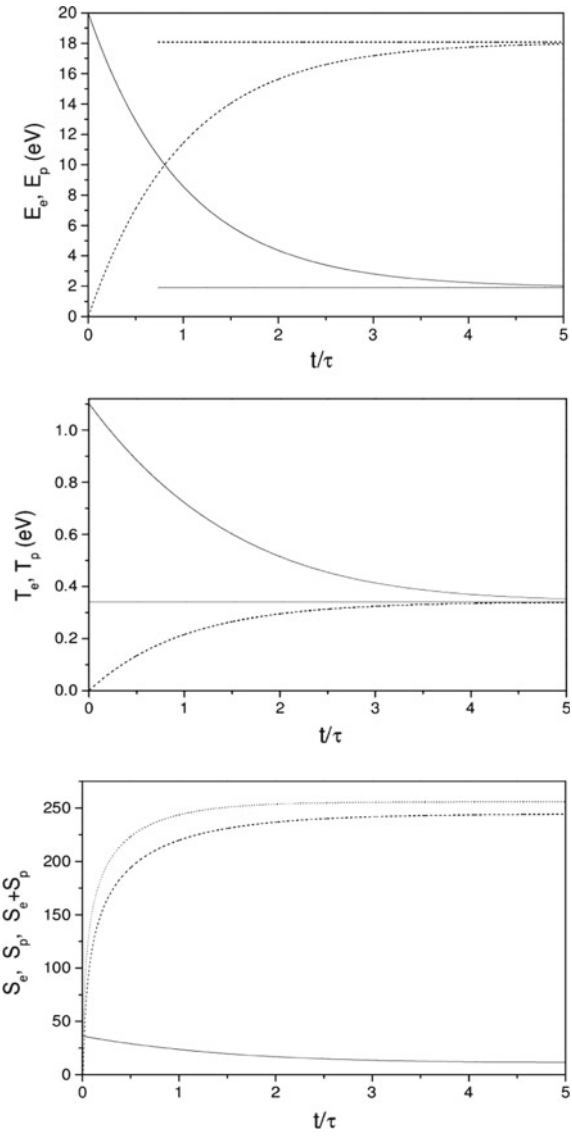


Fig. 11.3 Excitation energies for electrons and vibrations as a function of time (top frame). The electronic value is given with the full and the vibrational by the dashed line, here and in the two other frames. The asymptotic values are indicate with horizontal lines. Middle frame: The temperatures corresponding to the excitation energies. Bottom frame: The entropies of the two subsystems and the total entropy

$$S_e = \int_0^T \frac{C_v}{T'} \mathrm{d}T' = \frac{a}{2} T, \tag{11.21}$$

where the microcanonical heat capacities and temperatures are used. The vibrational entropy is calculated as the canonical value for an Einstein crystal with $\hbar\omega = 0.01$

eV. This is the easiest way to conform to Nernst's theorem.[1] In summary:

$$S_e = \frac{1}{2} a T_e, \qquad S_{ph} = s \frac{\hbar\omega}{T_{ph}} \frac{1}{\mathrm{e}^{\hbar\omega/T_{ph}} - 1} - s \ln\left(1 - \mathrm{e}^{-\hbar\omega/T_{ph}}\right). \tag{11.22}$$

11.3 Hot Electron Spectra

The thermal properties of the hot electron phase can be treated with the tools developed in Chap. 10, as already indicated in numerical estimates made above. One should keep in mind that the electronic temperature exceeds usually encountered values by an order of magnitude or more. It is possible to reach very high electronic temperatures with even moderate amounts of excitation energy due to the low heat capacity of the electronic subsystem.

One of the effects which is easiest to observe for a hot electron system is emission of thermal electrons. The process is so similar to the the thermionic emission described in Chap. 5 that one just need to plug in different parameters. The main difference is that of the level density of the emitting system, which does not change the fundamental equations. The change of level densities from vibrational to electronic increases the emission rates for hot electrons compared to an equilibrated system by a very large factor for identical excitation energies, reaching picosecond and sub-picosecond rates for even moderate excitation energies. Figure 11.4 shows the comparison between the hot electron and the equilibrated emission rate constants for C_{60}.

As for thermionic emission the kinetic energy distributions are given by the product of the electron attachment cross section, $\sigma(\varepsilon)$, a phase space factor of energy, ε, and the product level density, $\rho_p(E - \varepsilon)$:

$$P(E;\varepsilon) \propto \sigma(\varepsilon)\varepsilon\rho_p(E - \varepsilon). \tag{11.23}$$

Given the low heat capacity and the high energies of the emitted electrons, it is not obvious that the usual expansion of the level density to give a quasi-Boltzmann factor is justified. It is. Exercise 11.4 asks you to show that the leading order term is indeed sufficient in almost all cases. Hence we have

$$P(E;\varepsilon) \propto \sigma(\varepsilon)\varepsilon\mathrm{e}^{-\beta\varepsilon}. \tag{11.24}$$

The attachment cross section is specific to the material/molecule and the charge state of the particle. If we are happy with schematic values, which should actually be good approximations at high energies, we can refer to the results in Sect. 5.6.

Of particular interest for comparison with experiments are the cross sections referring to a cationic product, whether singly or multiply charged. Given the high

[1] Nernst's theorem: 'The entropy at zero kelvin is zero.'

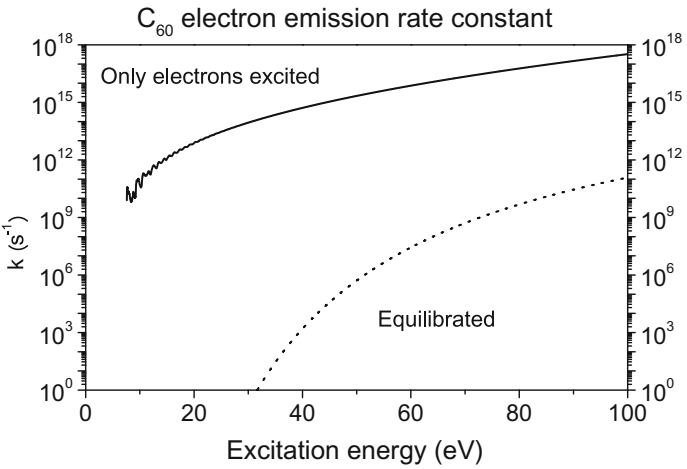


Fig. 11.4 The C_{60} electron emission rate constants for thermionic emission (dotted line) and hot electron emission (full line), as a function of excitation energy

excitation energy of the cation and the high kinetic energy of the electron, which provides empty states to settle in for an impinging electron in the otherwise densely populated Fermi gas, it is not unreasonable to assume that that electrons stick to the particle on reaching some distance from the center of mass. If so, the cross section is given by (5.74) which, multiplied with the phase space factor ε, gives

$$P(E;\varepsilon) \propto \left(\varepsilon + \frac{Ze^2}{4\pi\epsilon_0 r_0}\right) \mathrm{e}^{-\beta\varepsilon}, \qquad (Z > 0). \tag{11.25}$$

Often the second term in the brackets dominates and one will see an effective Boltzmann factor in measured distributions. Figure 11.5 shows an example for C_{60} excited with a short pulse laser. Apart from some low energy features related to the finite laser duration (see figure caption), the curves are well behaved exponentials with slopes that decrease with increasing laser pulse energy, as expected.

A similar behavior is observed for several PAH (polycyclic aromatic hydrocarbon) molecules. Examples of spectra for coronene are shown in Fig. 11.6.

Spectra for sodium clusters are shown in Fig. 11.7. The spectra are fitted with both pure exponentials and with (11.25). The excitation energy is calculated from the fitted temperatures assuming that the electrons can be described as a Fermi gas. The time dependence of the emitting clusters is not included in the fit with (11.25), and neither are the photon number distributions given in (11.28). This causes the calculated dip at the high energy end of the calculated spectra.

The interpretation of the exponential slope as the reciprocal temperature of the product is only approximate. The full expression for the kinetic energy spectra for a given initial excitation energy requires an integration over time to account for the spectra produced at different excitation energies:

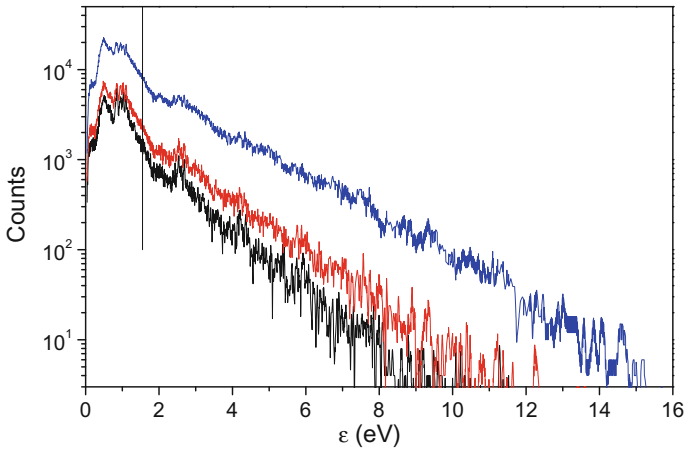


Fig. 11.5 Electron kinetic energy spectra for C_{60} after exposure to a focused 180 fs laser pulse of 1.55 eV photons. The curves are (from below) for the laser fluences 1.5, 2, and 3 J/cm^2. The vertical line indicates the photon energy. The spectra below this energy are heavily influenced by one-photon ionization of the hot electron gas in a thermally assisted ionization process. The small peaks discernible around 3 eV are the remnants of the ATI ionization one observes at shorter pulse durations. The data were recorded with a Time-of-Flight spectrometer. They are published in K. Hansen, K. Hoffmann and E.E.B. Campbell, *J. Chem. Phys.* **119** (2003) 2513

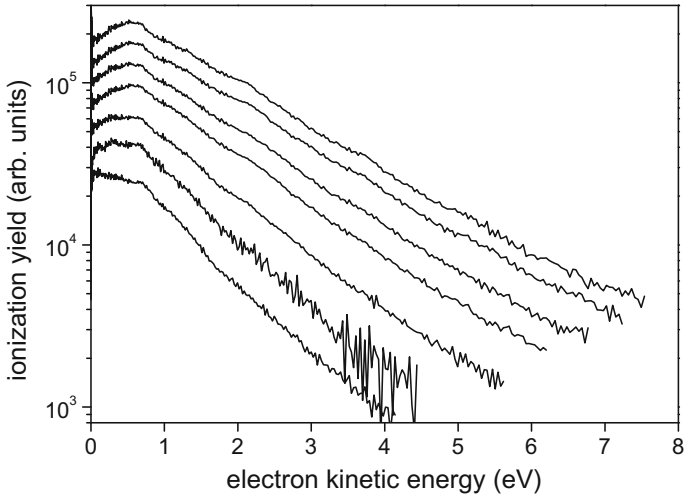


Fig. 11.6 Electron kinetic energy spectra for the coronene molecule ($C_{24}H_{12}$). Coronene is a flat molecule of seven carbon rings with the outer bonds of the outer six terminated by hydrogen. The laser fluence is, from bottom to top, 0.54, 0.74, 1.0, 1.2, 1.4, 1.6, 1.8 J/cm^2. Integrated intensities are not directly comparable. The data are recorded with a Velocity Map Imaging (VMI) spectrometer and are published in M. Kjellberg, A.V. Bulgakov, M. Goto, O. Johansson, and K. Hansen, *J. Chem. Phys.* **133** (2010) 074308

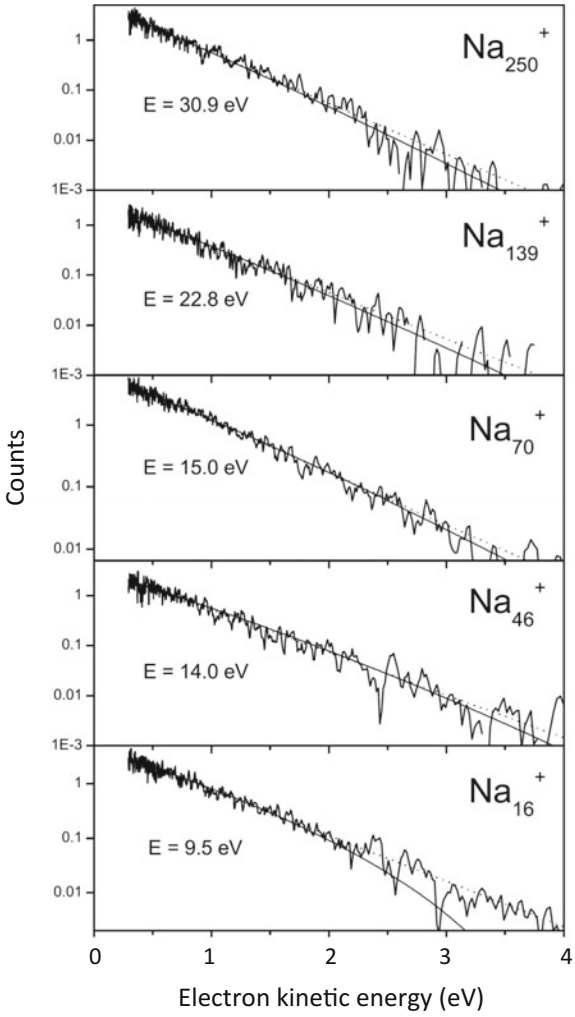


Fig. 11.7 Photo-electron spectra for sodium cluster of the sizes indicated in the frames. The dotted lines are simple exponential fits, and the full line are fits with the detailed balance equation in (11.25) with the image charge included in the calculated electron attachment cross section. The fitted temperatures are between 0.52 eV for $N = 16$ and 0.43 eV for $N = 250$. The clusters were initially singly positively charged, as required to perform the size selection prior to exposure to the laser pulse. The photon energies were 3.1 eV, a little above the sodium surface plasmon resonance. The energies given in the frames are the excitation energies calculated from the fitted temperatures. The data are published in M. Maier, M. Schätzel, G. Wrigge, M. Astruc Hoffmann, P. Didier, B. von Issendorff, *Int. J. Mass Spectrom.* **252** (2006) 157–165

$$f(\varepsilon) = \int_0^\infty P(t)\mathrm{d}t \propto \int_0^\infty \sigma(\varepsilon)\varepsilon \mathrm{e}^{-\varepsilon/T(t)} k(E(t)) \exp\left(-\int_0^t k(E(t'))\mathrm{d}t'\right)\mathrm{d}t. \tag{11.26}$$

The first three factors in the integrand give the energy-specified emission rate constants from (11.24), the fourth is the weight from the instantaneous rate constant, and the last is the surviving population. The expression is often well approximated by using the energy $E(0)$ because the rate constant is so strongly dependent on energy that contributions from later times are small, but observable details will differ with and without this integration.

The distributions in (11.26) refer to a single initial excitation energy. If a beam is exposed to different degrees of excitation resulting in a different initial excitation

energy, these also need to be integrated over. For excitations with laser, for example, one needs to sum over different numbers, n, of photons absorbed. For energy independent photon absorption cross sections, σ_p, the distribution for a constant laser fluence, F, will be given by the Poisson distribution,

$$P_n = \frac{(\sigma_p F)^n}{n!} \mathrm{e}^{-\sigma_p F}. \tag{11.27}$$

This is usually not the whole story, however. Laser fluences are not constant in space, and as a rule one needs to integrate over space dependent fluences:

$$P_n \propto \int \frac{(\sigma_p F(\overline{x}))^n}{n!} \mathrm{e}^{-\sigma_p F(\overline{x})} \mathrm{d}\overline{x}. \tag{11.28}$$

The resulting expressions rapidly become rather involved with little chance of even approximate analytical solutions and they are best solved numerically.

11.4 Ionization Yields

One of the interesting features of the hot electron phase is the possibilities it opens for fragment-free ionization. This possibility was already realized in the Penning ionization yields from fullerenes. It is therefore of interest to have an understanding of the ionization efficiency. Figure 11.8 shows the total ionization yields after exposure of C_{60} molecules to a 1.55 eV photon energy, 180 fs laser pulse as a function of the fluence.

All ions produced by the laser pulse are added up to give the points shown, including fragment ions and multiply charged ions. For high fluences, equivalent to high initial excitation energies, the molecules can ionize several times;

$$\mathrm{C}_{60} \rightarrow \mathrm{C}_{60}^{+} + e^{-} \rightarrow \mathrm{C}_{60}^{2+} + 2e^{-} \rightarrow \ldots \tag{11.29}$$

The ion yield is therefore a safer measure for the quantum yield than the electron yield, although fission processes may still distort the curve at high laser fluences by adding ion counts with the process

$$\mathrm{C}_N^{q+} \rightarrow \mathrm{C}_{N-M}^{(q-p)+} + \mathrm{C}_M^{p+}. \tag{11.30}$$

At high fluences fragment ions are produced. These are usually products generated in normal unimolecular reactions during the time between the establishment of the equilibrium, which takes a ps or so, and the time it takes to mass select the ions in a mass spectrometer, typically about 100 ns. This vast stretch of time is sufficient to host a number of unimolecular decays.

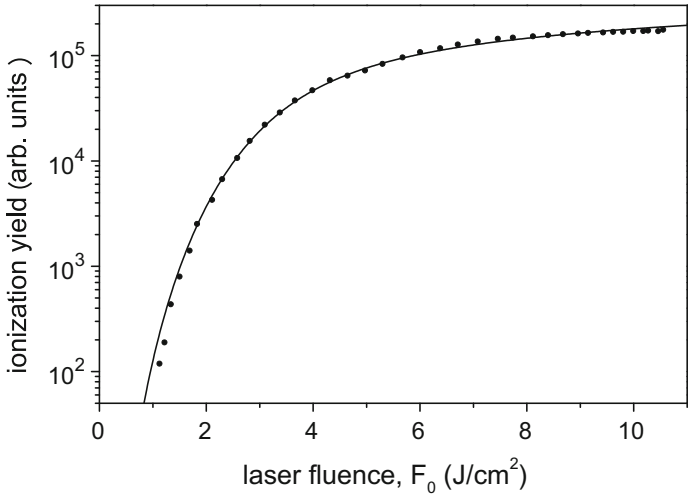


Fig. 11.8 The yield of all ions produced after exposure of C_{60} to a 180 fs laser pulse. The line is a fit determined with (11.31)

The data shown in Fig. 11.8 are fitted with a curve that takes into account both the time dependence, the photon absorption Poisson statistics, the depletion of the molecular beam and the variation of the laser fluence over space, assuming a Gaussian distribution. The grand total expression is

$$Y_{obs} \propto \int_0^\infty \mathrm{d}\varepsilon \int_{r=0}^\infty \mathrm{d}r\, r \sum_{n=0}^\infty \int_0^\infty \left(\varepsilon + \frac{e^2}{4\pi\epsilon_0\, 4\text{Å}}\right) \frac{\rho_p(nh\nu e^{-t/\tau} - \Phi - \varepsilon)}{\rho_r(nh\nu e^{-t/\tau})} \times \frac{\left(\sigma_p F_0/h\nu\, e^{-(r/r_0)^2}\right)^n}{n!} e^{-\sigma_p F_0/h\nu\, e^{-(r/r_0)^2}} \exp\left(-\int_0^t k\left(nh\nu e^{-t'/\tau}\right)\mathrm{d}t'\right)\mathrm{d}t. \tag{11.31}$$

The integral over ε sums the different kinetic energy contributions to give the total yield, the integral over r accounts for the spatial variation of the laser fluence, the summation over n with the Poisson statistics gives the photon absorption statistics for a given laser fluence, and the integral over time the yield as the initial excitation dissipates into vibrational motion. The quantities are as follows: ρ_p is the electronic level density of the product C_{60}^+ and ρ_r that of the neutral reactant, $h\nu = 1.55$ eV is the photon energy, $\Phi = 7.6$ eV is the ionization energy, F_0 is the laser fluence in the center of the beam (in units of photons per area), r_0 is the laser beam waist, and σ_p is the photon absorption cross section, as before, and finally τ is the energy based coupling time. The level densities were calculated as described in Sect. 10.7 with single particle levels provided by a Density Functional Theory (DFT) calculation.

The coupling time for the molecule is 240 fs, fitted from the Penning ionization data. The Penning ionization data have the great advantage that the excitation energy

is known and identical for all molecules in the beam, which removes one integral, one sum and one parameter from (11.31), and therefore gives a safer determination of the coupling time. The photon absorption cross section was assumed constant in the fit of the data. After the coupling time is entered, the fit with (11.31) contains only two parameters. One is for the photon absorption cross section which was fitted to $\sigma_p = 0.06$ Å^2, and the second an overall molecular beam intensity.

At the other extreme of the scale of accuracy and complications, one may simply use the rate constants as a proxy for the yield and approximate the rate constants by the simplified expression

$$Y \sim c \exp\left(-\Phi/T\right), \tag{11.32}$$

where the temperature is approximated by the value derived from the equidistant spectrum;

$$E = F_0\sigma_p = \frac{\pi^2}{4}\frac{N}{E_F}T^2 \Rightarrow T = \frac{2}{\pi}\left(\frac{E_F}{N}\right)^{1/2}\left(F_0\sigma_p\right)^{1/2}. \tag{11.33}$$

Sticking with C_{60}, the number of valence electrons is 240 and the Fermi energy 29.7 eV$-\Phi$ = 22.1 eV. Figure 11.9 shows the logarithm of the ionization yields vs. the reciprocal square root of the laser fluence. Equations (11.32, 11.33) are likely to be best at low pulse energies where the molecular beam depletion is less severe.

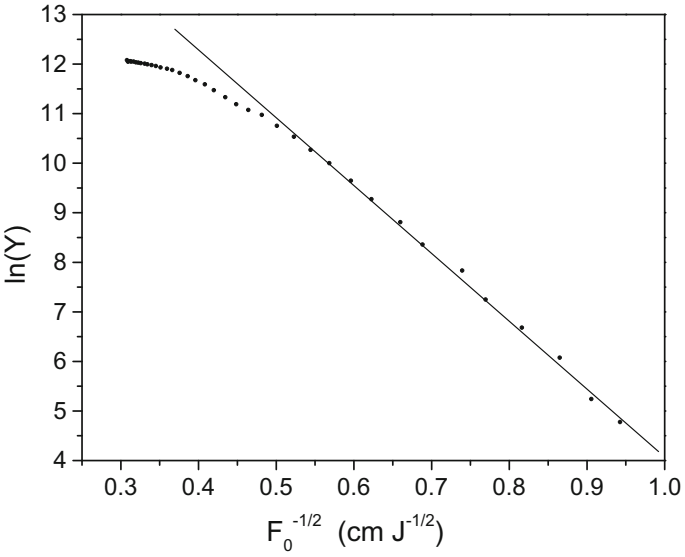
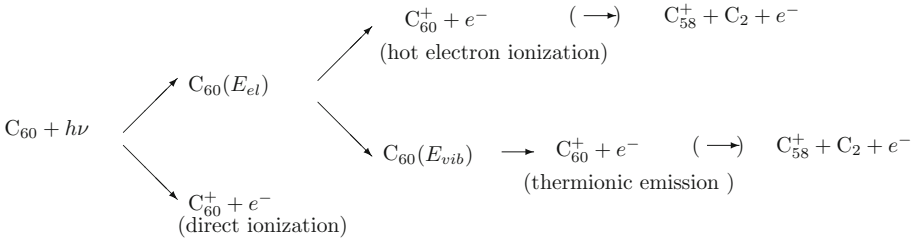


Fig. 11.9 The fit of the C_{60} ionization yields plotted according to the simple analysis in (11.32, 11.33). The slope of the line is 13.7 J$^{1/2}$/cm

The photo-absorption cross section is calculated from the fitted slope to be 0.0013 Å^2, a factor of five less than the value obtained with the much more elaborate procedure.

11.5 Single Photon Excitation

Somewhat surprisingly it is also possible to observe the hot electron phenomenon in single photon ionization. The excitation process is less clear than for multiphoton excitation, where one can imagine that sequential absorption of photons followed by fast dissipation into electronic degrees of freedom will effectively distribute the photon energies into incoherent excitations. From that state the particle's future destiny is hot electron thermal electron emission or further dissipation into the vibrational entropic wilderness. For single photon hot electron ionization a direct ionization is added to the reaction scheme. The possible reactions for C_{60} are the following:



where E_{el} and E_{vib} indicate that the excitation energy is predominantly electronic or vibrational.

There are three processes that produce an electron. They are distinguished by the electron energies. The direct ionization process has a fairly strong presence in single photon ionization. But the electron energy in these processes is given by $E_k = h\nu - E_b$, where E_b is the binding energy of the electron in the state from which it is ionized. It is therefore possible to distinguish these electrons from the hot electrons with their Boltzmann-like spectra with a characteristic energy of a few eV.

Also the thermionic emission spectra can be distinguished from the hot electrons, although they also appear with a Boltzmann-like distribution, because the emission temperatures in these processes are around $T_e = 3000 - 3500$ K $\simeq 0.3$ eV for C_{60}. The identification of the emission process is made easier by the fact that thermionic electrons do not appear in the experimental spectra. This is seen from the time profile of the produced ions measured in the Time-of-Flight mass spectrometer sitting back-to-back with the electron spectrometer and used to measure ions and electrons in coincidence. A thermionic emission contribution would manifest itself as a tail toward longer times in the mass spectrum. None was observed. This is readily understood as the effect of depletion of the beam; At excitation energies where thermionic emission occurs on the ns time scale, the hot electron rate is so high that practically no neutral molecules will survive.

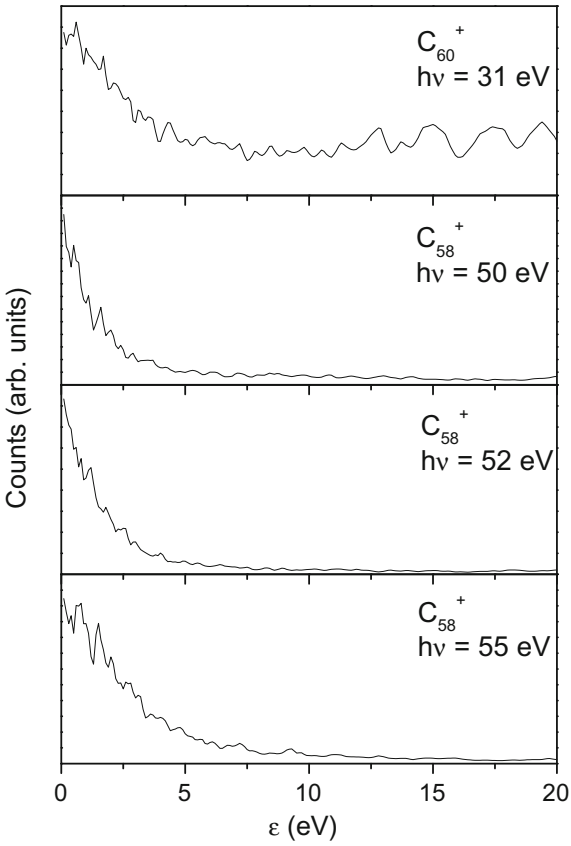


Fig. 11.10 Single photon electron spectra of C_{60} measured at the synchrotron ring Elettra. The electron spectra were recorded in coincidence with the produced ion. The presence of C_{58}^+ is due to post-ionization fragmentation before acceleration in the Time-of-Flight mass spectrometer, as indicated in the top line of (11.5). The spectra show the appearance of hot electron spectra with increasing photon energy and the approximate exponential shape of the spectrum at the low electron energies. The peak structure between 10 and 20 eV in the $h\nu = 31$ eV spectrum is due to direct ionization. These peaks move up in energy with the photon energy and are out of reach of the detector for the other photon energies. The data are published in K. Hansen et al., *Phys. Rev. Lett.* **118** (2017) 103001

Figure 11.10 shows the development of single photon hot electron spectra recorded with the VMI technique with increasing photon energy.

The gradual appearance of an exponential low energy tail and the simultaneous disappearance of the structured spectra at higher energies signals the onset of the hot electron ionization.

Exercises

11.1 The surface plasmon resonance exhausts a large part of the valence electron oscillator strength of the fullerenes. It peaks around 20 eV and has a width which is comparable with the centroid energy. Find the lifetime of the resonance.

11.2 Calculate the electron-electron collision time with (11.5, 11.7) for the Fermi energies 2 eV, 10 eV and 30 eV and compare them.

11.3 Show with (11.22) that the entropy grows during the electron-vibration equilibration process, irrespective of the values of a and q. This result is a cornerstone of thermodynamics, for whichever ensemble, but is not a given when calculated with the microcanonical temperature and heat capacity.

11.4 Show, using typical values of T of 1 eV and the Fermi gas approximation with an E_F of several eV, that the heat capacity of the electron system is large enough to terminate the expansion of the exponential in (11.23) at the first order term for all but the smallest particles. You can set the detection limit on the electron count to 10^{-3} of the maximum.

Chapter 12
He Droplets

Helium is the lightest of the rare gases. As a rare gas it is chemically inactive and is only bound to other atoms, other helium atoms included, by weak and non-directional forces. The potential energy of two rare gas atoms of the same type is often described by a sum of two high powers of the reciprocal of the interatomic (internuclear, to be exact) separation. The most common is called the Lennard-Jones potential and has powers -12 for the repulsive part and -6 for the attractive part;

$$V(r) = 4\varepsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right). \tag{12.1}$$

The Lennard-Jones parameters for the rare gases are given in Table 12.1.

The potential is shown in Fig. 12.1. It reproduces the expected $1/r^6$ long distance behavior of the interaction of two neutral particles. Such a distance dependence is a general behaviour for neutral particles and is due to the polarizability caused by quantum mechanical fluctuations in the electronic charge distributions. The $1/r^{12}$ term represents the strong repulsion between the atoms at short distances. The precise value of the power is dictated more by convenience than fundamental reasons.

The equilibrium distance, r_0, and the binding energy for a helium dimer in a classical ($\hbar = 0$) description of the bonding in the potential in (12.1) is

$$\frac{\mathrm{d}V}{\mathrm{d}r} = 0 \Rightarrow r_0 = 2^{\frac{1}{6}}\sigma = 1.122\sigma, \tag{12.2}$$

and

$$V(r_0) = -\varepsilon. \tag{12.3}$$

The main reason helium attracts special interest is that the interaction with other helium atoms is weak. Combined with the small mass of the atom, this causes so

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_12

Table 12.1 Lennard-Jones parameters of the rare gases. The masses are rounded averages over the isotopic compositions

	He	Ne	Ar	Kr	Xe
ε (K)	10.2	36.7	120	164	231
σ (Å)	2.6	2.9	3.4	3.6	4.0
m (u)	4	20	40	84	131

large quantum delocalization in the positions of the atoms that they are not bound to lattice points, even in the bulk. This prevents liquid helium from crystallizing at low temperatures, as all other elements do (at atmospheric pressure, that is. Above 24 atmospheres it does crystallize). The low temperature phase is a quantum liquid. Precisely what type of quantum liquid is determined by the isotope. Helium has two stable isotopes, ^{3}He and ^{4}He, of which the first is a fermion, with a natural relative abundance of only $1.4 \cdot 10^{-6}$. Most interest is therefore focused on the bosonic ^{4}He, and we will deal exclusively with that species here.

One can get a semi-quantitative understanding of the fact that helium doesn't crystallize, although not of the properties of the liquid, which is a much more involved and interesting problem, by calculating the zero point energy of the vibrational motion of the dimer with the potential in (12.1). Approximating the bottom of the potential with a harmonic oscillator, the vibrational frequency is

$$\mu\omega^2 = \frac{\mathrm{d}^2 V}{\mathrm{d}r^2}\Big|_{r=r_0} = 62 \cdot 2^{2/3} \frac{\varepsilon}{\sigma^2}, \tag{12.4}$$

with the reduced mass $\mu = m/2$. We then have the ratio of zero point energy to dimer binding energy

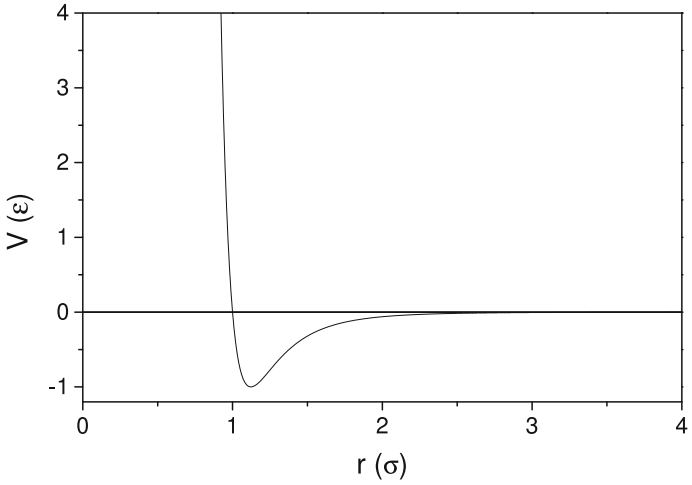


Fig. 12.1 The potential energy of two atoms that interact with the Lennard-Jones potential

Table 12.2 The zero point energy of the rare gas dimers relative to the Lennard-Jones energy

	He	Ne	Ar	Kr	Xe
$\frac{\frac{1}{2}\hbar\omega}{\varepsilon}$	2.9	0.62	0.21	0.12	0.070

$$\frac{\frac{1}{2}\hbar\omega}{\varepsilon} = \sqrt{\frac{62}{2^{1/3}}}\frac{\hbar}{\sigma\sqrt{m\varepsilon}} \approx 7.01\frac{\hbar}{\sigma\sqrt{m\varepsilon}}. \tag{12.5}$$

Values of this ratio for all noble gases are given in Table 12.2. Clearly He stands out with a zero point motion energy that exceeds the binding energy.

This suggests why helium doesn't crystallize, although we should be aware of the limitations of the calculation. It can not, for example, be inferred that the helium dimer does not exist. It does, just about, but it has a very long bond, about 50 Å, and is bound by the minute but finite energy of 10^{-7} eV, or 10^{-3} K. For comparison, the radius R of a liquid He drop composed of N atoms and with size independent density D is

$$\frac{4\pi}{3}R^3 = N\frac{m_a}{D} \Rightarrow R = 2.22\text{Å}N^{1/3} \tag{12.6}$$

where m_a is the mass of the atom and the density is $D = 0.145$ g/cm^2 at 0.4 K.

Small helium droplets will be described with the liquid drop model, but it is worthwhile to keep a few facts in mind when applying this description to helium droplets.

First, the density of the droplet is not a step function at the edge. Given the delocalization of helium atoms in the liquid, one must expect that also the surface densities vary over a finite length. Scattering experiments show that the atomic density varies from 90% of the central value to 10% over a distance of 6 Å of the surface. The surface diffuseness can also be calculated theoretically with Density Functional Theory. It is found to be 7 Å, for the bulk surface and slightly larger for finite sizes, defined in the same way as the experimental value.

Whether the value is 6 or 7 Å, the number of atoms in the surface will be a considerable fraction for even large droplets and renders the dimensions of droplets partly a matter of definition. With the radius of $r_N = 2.22\text{Å}N^{1/3}$ from (12.6), the ratio of the surface diffuseness to the radius is 0.68 for $N = 100$, 0.15 for $N = 10^4$, and 0.03 for $N = 10^6$. Figure 12.2 shows the fraction of atoms in the surface layer vs. size, for an average density of the layer of 0.5 time the bulk density and a thickness of 6.5 Å. It would be a conservative estimate to set the ratios in the figure equal to, e.g., the relative errors in the energy of the surface modes we will calculate below, but the numbers do tell us that the simple estimate for the quantum energy of the excitations needs to be taken with a grain of salt.

The bosonic nature of ^{4}He combined with the delocalization means that liquid helium may be described as a Bose-Einstein condensate (BEC) at sufficiently low temperatures, although it should be mentioned that this classification has been questioned. BEC or not, a phase transition occurs at 2.17 K for bulk liquid helium, called the lambda transition after the graphical similarity between the curve depicting the

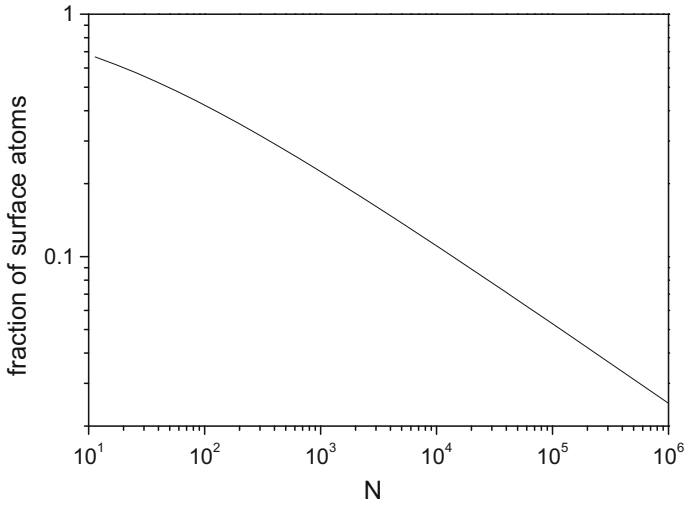


Fig. 12.2 The fraction of surface atoms in helium droplets as a function of the number of atoms in the droplets

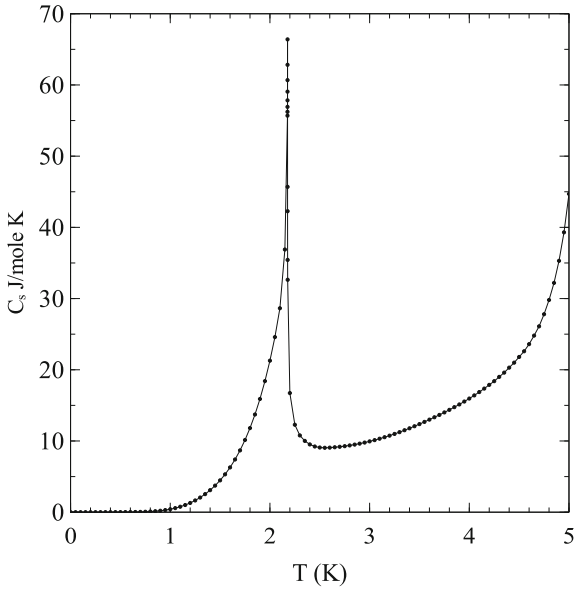


Fig. 12.3 The heat capacity of liquid helium under its saturated vapor pressure, C_s. The data are the recommended values from R.J. Donnelly and C.F. Barenghi, *J. Phys. Chem. Ref. Data*, **27** (1998) 1217–1274

heat capacity vs. temperature around the transition temperature and the Greek letter. Figure 12.3 shows measured low temperature ^{4}He heat capacities.

Below the lambda temperature the liquid has a number of interesting properties which have been studied intensely. One is that the liquid is superfluid, i.e. moves with no friction. This also holds for interactions with foreign objects, provided the

relative speed of the object and the helium does not exceed a value called the Landau critical velocity, and which is 60 m/s in the bulk. Above this relative speed, pairs of quasi-particles called rotons are created at the interface of the liquid and a embedded surface, costing a minimum energy of 8.7 K per roton pair and thus creating an effective friction in the motion.

For bulk helium the binding energy per atom is 0.86 meV, or 10 K. In a beam of freely evaporating helium droplets the temperature is a small fraction of this energy. In Chap. 7 it was found to be on the order of 3–4 % for typical clusters, and the value is similar for helium. Evaporative cooling therefore generates internal energies corresponding to temperatures of $\approx$0.4 K and helium droplets are used experimentally to thermalize dopant molecules to this temperature. Temperatures of 1 K and much lower are routinely obtained in conventional cryostats for pieces of macroscopic matter, but is an unreachable temperature for gas phase molecule and cluster beams without the helium droplet doping technique. With the temperature well below the lambda temperature of 2.17 K, one must expect that all but the smallest droplets will be superfluid. And molecules have indeed been observed to rotate freely inside helium droplets with rotational constants close to the vacuum values, which is only possible for superfluid droplets.

Experimentally, the droplets are formed during expansion of pre-cooled He gas with a stagnation pressure, i.e. the pressure in the source before it starts to flow, close to the saturated vapor pressure. The method can produce droplets containing very large numbers of He atoms. For experiments with doped droplets, the droplets produced in the expansion are passed through a chamber where they pick up one or more molecules or atoms. The dopant molecules collected in such pickup processes may aggregate to form clusters. Studies, for example of spectroscopic nature, can then be performed downstream in a molecular beam machine. The pickup chamber may be heated to high temperatures to get a workable vapour pressure of the dopants without destroying the droplets, because helium is almost transparent to black body radiation at even very high temperatures.

Absorption of photons by dopants are monitored by the loss of mass in the droplet, caused by the relaxation of the excitation energy and dissipation into the droplet which results in He atoms boiling off. The method has a high sensitivity, because the absorption of even low energy photons will cause the loss of a large number of He atoms. An infrared photon with an energy of 500 cm^{-1}, say, equal to 62 meV, will cause the loss of typically 500 cm^{-1}/0.86 meV = 72 atoms from the droplet. Cycling through the process of photon absorption, relaxation and cooling evaporatively several times can give a mass loss that can be measured experimentally.

12.1 The Excitation Spectrum

The calculation of the thermal properties of a helium droplet is greatly facilitated by the fact that the two most important contributions are separable to a

good approximation. The effects of dopants is less known and we will consider the pure helium droplet.

The ground state of a droplet is spherical because that shape reduces the surface energy to the minimum. Waves (ripples) on the surface of this sphere are the lowest energy excitations of the droplet shape. The quantized excitations are called ripplons, and they come with an associated angular momentum, $\ell \geqslant 2$. The spectrum starts at angular momentum 2 because zero angular momentum is spherically symmetric and does not represent a wave, and a putative unit angular momentum wave would correspond to a displacement of the whole droplet and is therefore a translation.

We can estimate the energies of the ripplons with a simple calculation. The kinetic energy of one of the ripplons is

$$E_k = \frac{\hbar^2 k^2}{2M} = \frac{\hbar^2 \ell^2}{R^2 2M}, \tag{12.7}$$

where k is the wave vector, M the effective mass participating in the motion, ℓ the angular momentum quantum number, and R the radius of the droplet. The mass is given by the amplitude of the wave, A, as

$$M = A4\pi R^2 D, \tag{12.8}$$

where D is the mass density of the droplet. Inserting this into (12.7) we have

$$E_k = \frac{\hbar^2 \ell^2}{R^4 A8\pi D}. \tag{12.9}$$

With the surface tension γ, the potential energy associated with the wave is estimated as

$$V = \gamma ARL. \tag{12.10}$$

The energy is then on the form, with a and b known coefficients,

$$E = a/A + bA, \tag{12.11}$$

and we can find the minimum energy for a excitation with a specific angular momentum by minimizing with respect to the amplitude of the wave,

$$\frac{\mathrm{d}E}{\mathrm{d}A} = 0 \Rightarrow A = (a/b)^{1/2}. \tag{12.12}$$

The second derivative is positive, so this is a minimum, as desired.

With the surface tension $\gamma = 3.54 \cdot 10^{-4}$ N/m and the liquid drop atomic size $r_1 = 2.22$ Å the amplitudes of the waves become

$$A = \frac{\hbar}{6m_a r_1^2 \gamma D} \frac{\ell^{1/2}}{N^{5/6}} \approx 0.38\text{Å} \frac{\ell^{1/2}}{N^{5/6}}, \tag{12.13}$$

which does not look unreasonable and which gives us confidence in our guess of how the parameters enters the expression for the energy of a ripplon. Inserting the optimal value of A into (12.11) it becomes

$$E = \hbar \ell^{3/2} \left(\frac{\gamma}{2\pi D R^3} \right)^{1/2}. \tag{12.14}$$

Given the somewhat cavalier treatment of the potential energy contribution, this is a remarkably good result. A more serious calculation of the quantum energies of the ripplons gives

$$E_\ell = \hbar\omega_0 \left(\ell(\ell-1)(\ell+2) \right)^{1/2}, \tag{12.15}$$

with the frequency ω_0 given by

$$\omega_0 = \sqrt{\frac{\gamma}{DR^3}}. \tag{12.16}$$

Each mode has a degeneracy of $2\ell + 1$, caused by the angular momentum. If we use the assumption of constant density, disregarding its shortcomings associated with the diffuseness of the surface, the denominator in the square root is equal to $3m_a N/4\pi$, where m_a is the mass of the atom and N is the number of atoms in the droplet. This gives

$$\omega_0 = \sqrt{\frac{4\pi\gamma}{3m_a N}}, \tag{12.17}$$

or

$$\hbar\omega_0 \approx 3.6\,\text{K} N^{-1/2}. \tag{12.18}$$

The $2\ell + 1$ degeneracy of the modes is consistent with the absence of internal structure in the droplet. The degeneracy is counting the different possible projections of the angular momentum vector on a space-fixed axis, as it is also known from the quantum mechanical treatment of other spherically symmetric systems. If the orientation of the droplet could be defined by an internal structure, one more degeneracy factor of $2\ell + 1$ would describe this orientation. Consistent with the absence of inner structure in the droplet, this degeneracy is absent here, analogously to the situation for e.g. atoms.

The only other type of excitation of importance for the thermal properties of free droplets are phonons. As phonons in solid matter, they are the quantized sound waves and involve atoms in the entire volume of the droplet. They arise as solutions to the wave equation that describes compression waves in the interior of the droplet. This description requires three coordinates per point and there are consequently three quantum numbers associated with the quantization of this motion. The spherical

symmetry of the droplets selects these quantum numbers to be the number of radial nodes in the wave function, and the angular momentum quantum numbers, ℓ and ℓ_z, as for ripplons. Their dispersion relation, i.e. the relation of frequencies to wave vectors, is

$$\omega_{n,\ell} = ck_{n,\ell} \tag{12.19}$$

where where c is the speed of sound and k is the wavenumber.[1] Hence phonon quantum energies are:

$$\varepsilon_{n,\ell} = \hbar ck_{n,\ell} \tag{12.20}$$

The wavenumbers $k_{n,\ell}$ are determined by the boundary condition at the surface. For a spherical particle and a free surface, the k's are given by

$$k_{n,\ell} = a'_{n,\ell}/R, \tag{12.21}$$

where $a'_{n,\ell}$ is the n'th root of the derivative of the spherical Bessel function, j'_ℓ.

Like for ripplons, it is convenient to express these energies in scaled units. We could use the highest frequency in the spectrum, the Debye frequency, as the energy scale. The Debye temperature for phonons in liquid ^{4}He is ≈ 20 K. Bulk helium boils at 4 K under atmospheric pressure and we will be looking at droplets below 1 K, so 20 K is a very high energy in this connection. We will therefore express phonon energies more conveniently in units of the energy of the longest wavelength,

$$\tilde{\varepsilon} = \hbar c\pi/R = \frac{\hbar c\pi}{2.22\,\text{Å}} N^{-1/3} = \tilde{\varepsilon}_0 N^{-1/3} = 25.8\,\text{K}N^{-1/3}, \tag{12.22}$$

where the 0.40 K values for the density (0.145 g/cm^2) and the speed of sound (238 m/s) were used. This energy scale decreases slower than that of the ripplons with N. Nevertheless, phonons turn out to be more important that ripplons for large sizes, because the number of phonons are proportional to N, whereas the number of ripplons is proportional to $N^{2/3}$. We will return to a quantitative treatment of these questions later, when more numbers are available.

Finally, for completeness we recall the existence of rotons, the third type of excitations. They were already mentioned in connection with the Landau velocity. However interesting they are, the threshold for their excitation is at rather high energy and consequently they contribute little to the thermal properties of helium droplets, and they will therefore not be considered further.

[1] The term wavenumber doubles as the length of the wave vector, as here, and an energy unit.

12.2 Ripplon Thermal Properties

He droplets are almost always observed flying freely through a molecular beam machine. The relevant ensemble for helium droplets is therefore the microcanonical ensemble, which means that we need to calculate the level density of the species to describe their thermal properties. In Chap. 4 different formulae were developed to calculate level densities directly with the vibrational spectra as input. One was the Beyer-Swinehart algorithm, and another was the inversion of the canonical partition function. We want a closed analytical form and use the latter procedure here.

We begin with a calculation of the leading order contribution from the ripplon excitation to provide a quick and portable result, followed by a more accurate calculation where higher order terms are included.

The calculation requires that we have the canonical partition function available. The ripplon elementary excitations are bosons and we therefore sum over all possible 'occupation numbers', which in this case are excitation energies:

$$\ln Z = -\sum_{\ell=2}^{\ell_{max}} (2\ell+1) \ln\left(1 - \mathrm{e}^{-\beta\varepsilon_\ell}\right). \tag{12.23}$$

To the desired precision, we can replace the sum in (12.23) by an integral from zero to infinity, and approximate the energy eigenvalues (12.15) by $\varepsilon_\ell \approx \hbar\omega_0 \ell^{3/2}$. The integral then evaluates to

$$\ln Z = \Gamma\left(7/3\right) \zeta\left(7/3\right) \beta^{-4/3} = 1.685 \left(\frac{T}{\hbar\omega_0}\right)^{4/3}, \tag{12.24}$$

where Γ is the gamma function, ζ is the Riemann zeta function and the energy scale was reintroduced in the last equality. The energy is then

$$E = -\frac{\partial \ln Z}{\partial \beta} = 2.25 \frac{T^{7/3}}{(\hbar\omega_0)^{4/3}} = 0.407 N^{2/3} \frac{T^{7/3}}{\mathrm{K}^{4/3}}. \tag{12.25}$$

The temperature and entropy can be expressed in terms of the excitation energy and the partition function. Use of (3.47) gives

$$\rho(E) = \frac{\mathrm{e}^S}{\sqrt{-2\pi(\partial E/\partial\beta)}}, \tag{12.26}$$

where $S = \beta E + \ln Z$ is the entropy. We then get the leading order ripplon level density

$$\rho_{rip}(E) \approx 0.31 \left(\frac{E}{\hbar\omega_0}\right)^{-5/7} \exp\left(2.48 \left(\frac{E}{\hbar\omega_0}\right)^{4/7}\right). \tag{12.27}$$

This expression is sufficient for a number of purposes, but we have the tools to calculate higher order contributions and it is worth doing to check of the limits of validity of the approximation made in leaving out these contributions.

The first improvement in the calculation is to use the relation between the microcanonical energy and temperature from Chap. 3:

$$E = -\frac{\partial \ln Z}{\partial \beta} - \beta^{-1}, \tag{12.28}$$

where Z is the canonical partition function at the microcanonical temperature β^{-1}. The next is to evaluate the partition function in (12.23) more accurately by using the Euler-Maclaurin formula. We use the first three terms in the expansion:

$$\begin{aligned} -\ln Z = &\int_2^\infty (2\ell+1)\ln\left(1-\mathrm{e}^{-\beta\varepsilon_\ell}\right)\mathrm{d}\ell + \frac{5}{2}\ln\left(1-\mathrm{e}^{-\beta\varepsilon_2}\right) \\ &- \frac{1}{12}\frac{\mathrm{d}}{\mathrm{d}\ell}\left[(2\ell+1)\ln\left(1-\mathrm{e}^{-\beta\varepsilon_\ell}\right)\right]\Big|_{\ell=2} \end{aligned} \tag{12.29}$$

The upper limit of integration of the first term has been set to infinity. The actual value is on the order of $\ell_{max} \approx 2\pi R/\lambda_{min} \approx 2\pi R/(2d)$, where λ is the wavelength and d is the interatomic distance. In the liquid drop approximation ($R = N^{1/3}d/2$) one has $\ell_{max} \approx \pi N^{1/3}/2$. In view of (12.15, 12.16) this yields a size-independent ripplon Debye temperature of $\varepsilon_{max} \approx 7.1$ K. Using this value to estimate the error in $\ln Z$, we find that the neglected terms are on the order of $(\beta\varepsilon_{max}/4 - 7\ell_{max}/6)\exp(-\beta\varepsilon_{max})$. For $T = 1$ K this is a relative contribution to $\ln(Z)$ of less than $10^{-2}/N^{1/3}$. For lower temperatures it is even less and we can consequently set the upper limit of the integral to infinity without any major loss of precision.

In the third step the relation between energy and angular momentum, (12.15), is inverted to express the angular momentum as a function of the energy. This allows a calculation of the integral in (12.29) by substitution. Finally, with the expansion of the exponential $\exp(-\beta E_{\ell=2})$ in $\beta E_{\ell=2}$, which is an acceptable approximation for not too small droplets, we get the result

$$\ln Z \approx c_2\beta^{-4/3} + \frac{c_1}{3}\beta^{-2/3} - \frac{349}{96} + \frac{7}{3}\ln(2\sqrt{2}\beta). \tag{12.30}$$

The numerical constants $c_1 = 1.917$ and $c_2 = 1.685$ that appear in this expression are given by

$$c_n = \sum_{j=1}^{\infty} j^{-2n/3-1}\int_0^\infty x^{2n/3}\mathrm{e}^{-x}\mathrm{d}x = \zeta(2n/3+1)\Gamma(2n/3+1). \tag{12.31}$$

Inserting the numerical values of c_1 and c_2 into (12.30) and reintroducing the physical energy scale gives

$$\ln Z = -\beta E + S = 1.685\left(\frac{T}{\hbar\omega_0}\right)^{4/3} + 0.639\left(\frac{T}{\hbar\omega_0}\right)^{2/3} - \frac{349}{96} - \frac{7}{3}\ln(2\sqrt{2}T/\hbar\omega_0). \tag{12.32}$$

The first term coincides with (12.24), as it should. The other terms vary slower with N than the leading order term and are finite-size corrections.

The result in (12.32) is a very good approximation to the exact result. The comparison with a numerical summation of the partition function for ripplons in (12.23) is shown in Fig. 12.4. Already at temperatures where T is equal to the lowest excitation energy $E_2 = \hbar\omega_0\sqrt{8}$ is the free energy well represented by the above expression. At higher energies the agreement improves monotonically.

We proceed by finding the caloric curve, or energy-temperature relation from (12.28) and (12.30) or (12.32) for the free energy:

$$\begin{aligned} E &= 2.247\frac{T^{7/3}}{(\hbar\omega_0)^{4/3}} + 0.426\frac{T^{5/3}}{(\hbar\omega_0)^{2/3}} - \frac{10}{3}T \\ &= 0.407T^{7/3}N^{2/3}\mathrm{K}^{-4/3} + 0.181T^{5/3}N^{1/3}\mathrm{K}^{-2/3} - \frac{10}{3}T. \end{aligned} \tag{12.33}$$

If we invert this relation and rewrite it slightly as

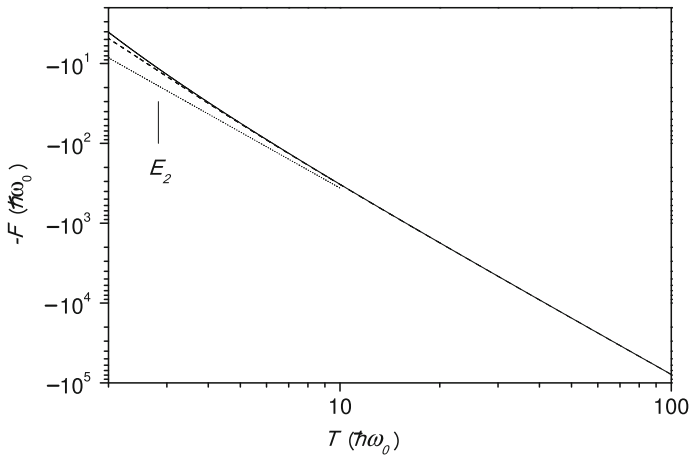


Fig. 12.4 The ripplon free energies, calculated with (12.32) (dashed line) and the summation in (12.23) (full line) which is exact apart from setting the upper summation limit to infinity. The dotted line is the first term of (12.32). The position of the lowest excitation energy, E_2, is indicated. The data were published in K. Hansen, M.D. Johnson and V.V. Kresin, *Phys. Rev. B* **76** (2007) 235424

$$T = \left(\frac{1}{2.247}\right)^{3/7} \left(E(\hbar\omega_0)^{4/3} - 0.426T^{5/3}(\hbar\omega_0)^{2/3} + \frac{10}{3}T(\hbar\omega_0)^{4/3}\right)^{3/7}, \tag{12.34}$$

we can calculate $T(E)$ by iterating with the procedure known as successive approximations of the right hand side. The result is that $T = E^{3/7}(\hbar\omega_0)^{4/7}$ times a power series in $(\hbar\omega_0/E)^{2/7}$ which we need not write down here.

We now have all quantities that enter (12.26). Keeping the leading three terms in the exponential, the total ripplon level density becomes

$$\rho_{rip}(E) = 0.205(\hbar\omega_0)^{5/7}E^{-12/7}\exp\left(2.48\,(E/\hbar\omega_0)^{4/7} + 0.507\,(E/\hbar\omega_0)^{2/7}\right), \tag{12.35}$$

where the energy scale is still given by $\hbar\omega_0 = 3.6\,\mathrm{K}N^{-1/2}$.

The calculation has given us the level density for a given total energy. It is also possible to calculate the angular momentum-specified level density. This function is relevant because excitations come with angular momentum and angular momentum is a conserved quantity. The calculation will therefore be done here.

The degeneracy of the ripplon energy levels, $2\ell + 1$, indicates that we can not use a classical description for this calculation, even for large quantum numbers. A degeneracy proportional to ℓ gives a canonical thermal rotational energy of T in the high temperature limit. In contrast, a classical situation would give a degeneracy factor proportional to ℓ^2, corresponding to the rotational degeneracy of a stiff body (see (8.30)). The canonical thermal energy of this is $3T/2$ (see Chap. 2). An error in thermal energy of 50% is indicative of a serious error in the level density. Happily, it is possible to handle the quantum mechanical problem without any semiclassical approximation.

To do so, we start with calculating the distribution of projections of the angular momentum on a fixed axis. We add these projections, m_ℓ, from all the individual degrees of freedom. The sum of thermal averages is zero:

$$\langle M_{rip}\rangle \equiv \sum_{\ell=2}^{\infty}\langle m_\ell\rangle = 0, \tag{12.36}$$

because there is no preferred direction in space and therefore all the contributions from different angular momenta are zero.

The sum of squares is obviously not zero. It is calculated as the sum over all ℓ's and all projections of that angular momentum, m_ℓ. A single of these states excited with n quanta and consequently with energy nE_ℓ is populated with the probability $\exp(-n\beta E_\ell)/Z$, where $Z^{-1} = 1 - \exp(-\beta E_\ell)$, and the length of the angular momentum projection is nm_ℓ where $-\ell \geqslant m_\ell \geqslant \ell$. For a given ℓ this gives the thermal average

$$\langle m_\ell^2 \rangle = \sum_{m_\ell=-\ell}^{\ell} \sum_{n=0}^{\infty} Z^{-1} m_\ell^2 n^2 \mathrm{e}^{-n\beta E_\ell} \tag{12.37}$$

$$= \left(\frac{2}{3}\ell^3 + \ell^2 + \ell/3 \right) (1 - \mathrm{e}^{-\beta E_\ell}) \sum_{n=0}^{\infty} n^2 \mathrm{e}^{-n\beta E_\ell}.$$

The sum over m_ℓ^2 will be represented by the leading term $2\ell^3/3$ in the following. The sum over n can be done with the result

$$\langle m_\ell^2 \rangle = \frac{2}{3}\ell^3 \left[\frac{\mathrm{e}^{-\beta E_\ell}}{1 - \mathrm{e}^{-\beta E_\ell}} + 2\frac{\mathrm{e}^{-2\beta E_\ell}}{(1 - \mathrm{e}^{-\beta E_\ell})^2} \right]. \tag{12.38}$$

The last term is rewritten by expanding the denominator as

$$\frac{1}{(1 - \mathrm{e}^{-\beta E_\ell})^2} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \mathrm{e}^{-(n+k)\beta E_\ell} = \sum_{n=0}^{\infty} (n+1) \mathrm{e}^{-n\beta E_\ell} \tag{12.39}$$

to give

$$\langle m_\ell^2 \rangle = \frac{2}{3}\ell^3 \left[\frac{\mathrm{e}^{-\beta E_\ell}}{1 - \mathrm{e}^{-\beta E_\ell}} + 2\mathrm{e}^{-2\beta E_\ell} \sum_{n=0}^{\infty} (n+1) \mathrm{e}^{-n\beta E_\ell} \right]. \tag{12.40}$$

The total ripplon contribution to the variance is the sum of such contributions from all ℓ:

$$\langle M_{rip}^2 \rangle = \sum_{\ell=2}^{\infty} \langle m_\ell^2 \rangle. \tag{12.41}$$

We can approximate the sum over ℓ with an integral and set the lower integration limit to zero, which is acceptable for high temperatures. With the substitution $u = \beta E_\ell = \beta\hbar\omega_0 \ell^{3/2}$ we get

$$\langle M_{rip}^2 \rangle \approx \frac{4}{9} \left(\frac{T}{\hbar\omega_0} \right)^{8/3} \int_0^{\infty} u^{5/3} \left(\frac{\mathrm{e}^{-u}}{1 - \mathrm{e}^{-u}} + 2\mathrm{e}^{-2u} \sum_{n=0}^{\infty} (n+1) \mathrm{e}^{-nu} \right) \mathrm{d}u. \tag{12.42}$$

The first part of the integral is calculated by expanding $1/(1 - \mathrm{e}^{-u})$ in e^{-u} and integrating term by term, which gives $\Gamma(8/3)\zeta(8/3)$. The second requires some reshuffling but is otherwise also straightforward. The final result for the integral is $I \equiv \Gamma(8/3)\,[2\zeta(5/3) - \zeta(8/3)] = 1.981\ldots$ and thus

$$\langle M_{rip}^2 \rangle \approx \frac{4}{9} \left(\frac{T}{\hbar\omega_0} \right)^{8/3} I = 1.981\ldots \left(\frac{T}{\hbar\omega_0} \right)^{8/3} = 0.785\ldots \left(\frac{E}{\hbar\omega_0} \right)^{8/7}, \tag{12.43}$$

where the leading order term in the caloric curve, (12.34), was used. With $\hbar\omega_0 = 3.6N^{-1/2}$ K, this can be expressed as

$$\langle M_{rip}^2 \rangle = 5.65 \cdot 10^{-3} N^{4/3} \left(\frac{T}{0.4\,\mathrm{K}} \right)^{8/3} \tag{12.44}$$

This is the mean of the square of the total angular momentum projection on a fixed axis. We still need to find the distribution of this quantity, and do so by invoking the Central Limit Theorem which says that if enough stochastic variables are added, the sum will be normally distributed with a variance which is the sum of the variances of the individual terms. This is exactly what we have calculated for M, and with the fact that the average of M itself is zero, we therefore have the (normalized) distribution of the projection of the total angular momentum:

$$\rho(E, M) = \rho(E) \frac{1}{\sqrt{2\pi}\sigma_M} \mathrm{e}^{-M^2/2\sigma_M^2}, \tag{12.45}$$

with $\sigma_M^2 = \langle M_{rip}^2 \rangle$ given in (12.43).

This distribution can be used to find the distribution of the total angular momentum, J. The number of states of a given M, denoted by $n(M)$, is the number of states of a given angular momentum, $n(J)$, with the projection M. For every J, there is one projection with the value M, provided $J \geqslant M$. For $J < M$ there is zero. In other words:

$$n(M) = \sum_{J=M}^{\infty} n(J). \tag{12.46}$$

Taking the difference of this relation for M and $M + 1$ we get

$$n(M) - n(M+1) = \sum_{J=M}^{\infty} n(J) \; - \sum_{J=M+1}^{\infty} n(J) = n(J = M), \tag{12.47}$$

Replacing the difference on the left hand side with a derivative, and the number of states, n, with the density of states, both of which are permissible procedures because the spacing of J is unity, one gets

$$\rho(E, J) = -\frac{\partial \rho(E, M)}{\partial M}\Big|_{M=J+1/2} = \rho(E) \frac{J + 1/2}{\sqrt{2\pi}\sigma_M^3} \mathrm{e}^{-(J+1/2)^2/2\sigma_M^2}. \tag{12.48}$$

With the size dependence of the energy scale from (12.18), the calculated σ_M and by setting $T = 0.4$ K, the angular momentum distributions for different droplet sizes are found to be,

$$\rho(0.4\,\mathrm{K}, J) \propto N^{-2} (J + 1/2) \exp\left(-88 (J + 1/2)^2 N^{-4/3}\right). \tag{12.49}$$

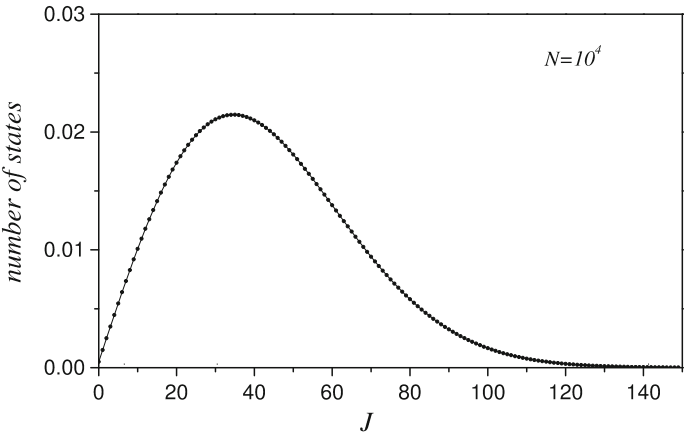


Fig. 12.5 The angular momentum distribution from the ripplon degrees of freedom at a temperature of 0.4 K for $N = 10^4$. The maximum for other droplet sizes varies as $N^{2/3}$

The distribution is shown in Fig. 12.5 for the droplet size $N = 10^4$. On inspection of (12.49) one sees that the distribution scales and has the same form for different sizes, with a scale factor that varies as $N^{2/3}$. The maximum of the distribution is reached at $J = N^{2/3}/\sqrt{88 \cdot 2} = 0.075N^{2/3}$ for the 0.4 K distribution. For other temperatures the values for these parameters can be found by the scaling of σ_M, which varies with temperature as $T^{4/3}$.

12.3 Molecular Beam Temperatures

We will now estimate the previously announced temperature for droplets flying freely through vacuum. The derivation will include only the ripplon thermal properties because these are the most important for droplets that are not extremely big. The droplets cool by evaporating atoms and is an example of the ensembles treated in detail Chap. 7. Following that derivation we set $k = C/tG^2$, where t is the time since the droplets started free flight. We know the evaporation rate constant of helium atoms from a droplet from (5.15). With zero electronic degeneracy and a geometric cross section it is

$$k(E) = \frac{m}{\pi^2\hbar^3}\pi r_1^2 N^{2/3}\, T_d^2\, \frac{\rho_{N-1}(E - E_a)}{\rho_N(E)} = \nu \frac{\rho_{N-1}(E - E_a)}{\rho_N(E)}, \tag{12.50}$$

where $r_1 = 2.22$ Åand $E_a = 10$ K are the atomic radius and the evaporative activation energy, respectively, and T_d is the product (daughter) temperature. With $T_d = 0.4$ K the pre-exponential factor is calculated to

$$\nu \equiv \frac{m}{\pi^2\hbar^3}\pi r_1^2 N^{2/3}\, T_d^2 = 1.8 \cdot 10^9 N^{2/3}\,\mathrm{s}^{-1}. \tag{12.51}$$

This is significantly smaller than corresponding factors for other elements. The difference is in part due to the small mass of the emitted particle but mostly caused by the very low kinetic energy. The cross section, represented by $N^{2/3}$, is a compensating factor, as helium droplets are often produced very large.

The level density of the product of size $N-1$ is close enough to that of droplet N to initially ignore the difference, i.e. we set $\rho_{N-1}(E) \approx \rho_N(E)$ (see Exercise 12.2), and the ratio of level densities is then $\rho_{N-1}(E-E_a)/\rho_N(E) \approx \rho_N(E-E_a)/\rho_N(E)$. The leading order finite heat bath correction is $E_a/2C$ (see Chap. 3). With $E_a = 10\,\mathrm{K}$ and the heat capacity from the ripplons, we get a value relative to the temperature of

$$\frac{E_a}{2CT} = 5.3 N^{-2/3}\, T^{-7/3}\,\mathrm{K}^{7/3}, \tag{12.52}$$

which is $45N^{-2/3} = (300/N)^{2/3}$ for $T = 0.4\,\mathrm{K}$. This all adds up to a rate constant of

$$k(E) = 1.8 \cdot 10^9 N^{2/3}\,\mathrm{s}^{-1} \exp\left(-\frac{E_a}{T\left(1-\left(\frac{300}{N}\right)^{2/3}\right)}\right). \tag{12.53}$$

For precise estimates, N needs to be above 10^4 before the finite heat bath correction can be ignored. Above that number the beam temperature, which is at the same time approximately both parent and daughter temperature, is

$$T = \frac{E_a}{\ln\left(1.8 \cdot 10^9 N^{2/3}\,\mathrm{s}^{-1} t\right)}. \tag{12.54}$$

The precise value depends weakly on the droplet size and on the observation time. The droplets are neutral, their speed thermal and typical flight times after creation will be between 100 μs and 10 ms, say. Sizes are not easy to measure for neutral particles but a good deal of effort has been devoted to this question and there are tools available that will give mean sizes with reasonable reliability. If we limit ourself to numbers between 10^4 and 10^8 atoms, we get for the extreme cases that $T = 0.55\,\mathrm{K}$ (for 100 μs, $N = 10^4$) and $T = 0.35\,\mathrm{K}$ (for 10 ms, $N = 10^8$). These numbers are a lot closer to each other than one may have guessed, given the wide range of the parameters, but they are not identical. The value of 0.36 K often quoted in the literature refers to relatively large droplets measured after a relatively long time.

12.4 Phonon Level Density

The leading-order behavior of the phonon density of states can be determined using the expression for the Debye heat capacity of bulk phonons calculated in the chapter on vibrational thermal properties (4.44), with some modifications. If the droplet is superfluid when it coasts through the vacuum, which experimentally seems to be the case, it can only support waves in the direction of their propagation. The two perpendicular modes are not supported by a superfluid, and the low temperature Debye heat capacity in (4.44) must therefore be divided by a factor 3:

$$\frac{\partial E}{\partial T} = C_{bulk} \approx \frac{2\pi^2 V}{15}\left(\frac{T}{\hbar c}\right)^3 = \frac{8\pi^3}{45}R^3\left(\frac{T}{\hbar c}\right)^3 = \frac{8\pi^6}{45}\left(\frac{T}{\tilde{\varepsilon}}\right)^3, \tag{12.55}$$

where c is the speed of sound, and the energy scale, $\tilde{\varepsilon} = \hbar c\pi/R$ from (12.22) was used. The experimental data for bulk helium, where practically all modes are phonons, agree very well with this equation, as shown in Fig. 12.6.

We find the energy content by integration with respect to T:

$$E = V\int_0^T C_{bulk}(T')\mathrm{d}T' = \frac{2\pi^6}{45}\frac{T^4}{\tilde{\varepsilon}^3} = 0.92N^{2/3}\frac{T^4}{\mathrm{K}^3}. \tag{12.56}$$

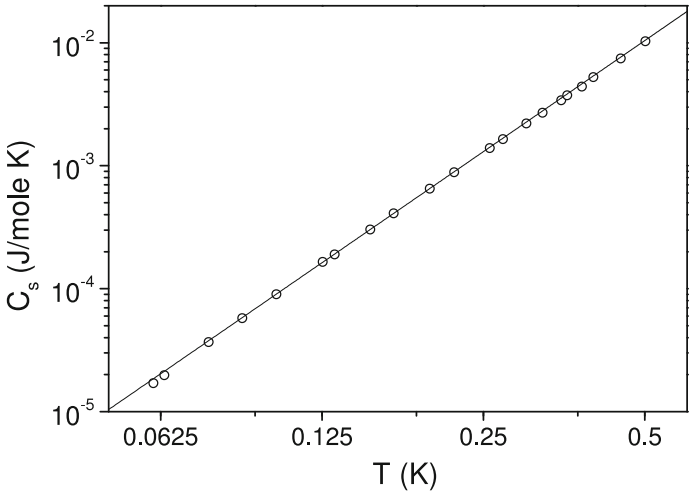


Fig. 12.6 The experimentally measured ^{4}He bulk heat capacity at the saturated vapor pressure (circles) and the predicted 1/3 Debye heat capacity (line), calculated with (12.55) with the experimentally measured density of 0.1451 g/cm^3 and the speed of sound $c = 238$ m/s. There are no adjustable parameters in the theoretical curve

With the caloric curve in (12.56) we can find the entropy in terms of the excitation energy by integration with the factor $1/T$. From the general result in (12.26), the level density is found to be

$$\rho_{ph}(E) \approx A\varepsilon^{-3/8}E^{-5/8}\exp\left(3.409\left(\frac{E}{\tilde{\varepsilon}}\right)^{3/4}\right), \tag{12.57}$$

where A is 0.32.

This is the first estimate of the level density, and depending on the size of the particle it may be precise enough. Higher order contributions are mainly generated by the deviations of the roots of the Bessel functions from an equidistant spectrum. They occur at low quantum numbers and can therefore not be captured by the experimentally measured macroscopic heat capacity, in particular considering that they are usually not measured on macroscopic drops and the detailed behavior of Bessel functions is irrelevant for their description.

A calculation of the thermal properties that takes this kind of systematics into account does not seem like an easy task. Fortunately there exists an expansion of the smoothed density of states in a finite cavity, covering the type relevant here. When applying these corrections, which we will not derive, we get the heat capacity (per unit volume)

$$C = C_{bulk} + \frac{9\zeta(3)}{4\pi}\frac{1}{\hbar^2c^2}\frac{T^2}{R} + \frac{1}{6}\frac{1}{\hbar c}\frac{T}{R^2}. \tag{12.58}$$

As one can see from the R dependence, the corrections give rise to a surface and a curvature term. The entropy and the caloric curve can be found by integration, as done above. The rest of the procedure is then similar to the one used for ripplons. We have gone through this procedure for other degrees of freedom a couple of times and just give the result:

$$\rho_{ph}(E) \approx A\frac{\tilde{\varepsilon}^{3/8}}{E^{5/8}}\exp\left(3.409\left(\frac{E}{\varepsilon}\right)^{3/4} + 0.908\left(\frac{E}{\varepsilon}\right)^{1/2} + 0.482\left(\frac{E}{\varepsilon}\right)^{1/4}\right). \tag{12.59}$$

This expression is accurate enough to warrant a comparison with the level density calculated with the Beyer-Swinehart algorithm (see Chap. 4), with the roots of the Bessel functions as input quantum energies. It must be expected that there are corrections that are unaccounted for because the argument of the exponential is an expansion in $E^{1/4}$, and the fourth term in (12.58), which we don't know, will be a constant and will therefore contribute to the pre-exponential. The calculation with the Beyer-Swinehart algorithm gives a leading order correction which can be fitted with the expression $\exp(-0.62(E/\varepsilon)^{0.2})$. An effective value of $A \approx 0.05$ can be used for phonon energies below 400 $\tilde{\varepsilon}$.

Phonons also carry angular momentum, and one of the two indices of the spherical Bessel functions gives the angular momentum (the other counts the number of nodes in the radial motion). For the calculation of the phonon angular momentum-resolved

density of states, which now be done, we will be satisfied with the leading order contribution.

For the free surface boundary condition, the roots of the Bessel functions are asymptotically $(n + \ell/2 - 3/4)\pi \approx (n + \ell/2)\pi$. If we use the phonon energy scale in (12.22), the quantum energies are thus $n + \ell/2$. In analogy with the calculation for ripplons, we sum the squares of projections of angular momentum as

$$\langle M_{ph}^2 \rangle = \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m_\ell=-\ell}^{m_\ell=\ell} \frac{\sum_{k=0}^{\infty} m_\ell^2 k^2 \mathrm{e}^{-k\beta E_{n,\ell}}}{\sum_{k=0}^{\infty} \mathrm{e}^{-k\beta E_{n,\ell}}} \tag{12.60}$$

There is one more summation here compared with the ripplon case, corresponding to the existence of the radial nodes, consistent with the three-dimensional nature of phonons vs. the two-dimensional nature of ripplons.

The sum over m_ℓ^2 gives $2\ell^3/3$, as before, and the sum over k is also identical to the ripplon case. We therefore have

$$\langle M_{ph}^2 \rangle \approx \frac{2}{3} \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \ell^3 \left[\frac{\mathrm{e}^{-\beta E_{n,\ell}}}{1 - \mathrm{e}^{-\beta E_{n,\ell}}} + 2 \frac{\mathrm{e}^{-2\beta E_{n,\ell}}}{(1 - \mathrm{e}^{-\beta E_{n,\ell}})^2} \right] \tag{12.61}$$

We perform one of the remaining two sums over constant energy surfaces, $n + \ell/2$ equal to a constant, in the n, ℓ plane. For the first term in (12.61), this gives an approximate value of

$$\frac{2}{3} \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\beta E_{n,0}}}{1 - \mathrm{e}^{-\beta E_{n,0}}} \sum_{l=0}^{2n} \ell^3 \approx \frac{8}{3} \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\beta E_{n,0}}}{1 - \mathrm{e}^{-\beta E_{n,0}}} n^4 \approx \frac{8}{3} \int_0^{\infty} \frac{\mathrm{e}^{-\beta n}}{1 - \mathrm{e}^{-\beta n}} n^4 \mathrm{d}n. \tag{12.62}$$

This is calculated with the same method used for ripplons; expand the integrand in $\exp(-\beta n)$, interchange integration and summation and shuffle some limits of the sums. The second term is calculated analogously. We get

$$\langle M_{ph}^2 \rangle \approx \frac{8}{3} \Gamma(5)(2\zeta(4) - \zeta(5)) T^5 = 72.2 \left(\frac{T}{\tilde{\varepsilon}} \right)^5, \tag{12.63}$$

where the energy scale has been reintroduced. Inserting the value of $\tilde{\varepsilon}$, the mean square M is, in terms of droplet size,

$$\langle M_{ph}^2 \rangle \approx 6.3 \cdot 10^{-6} N^{5/3} \left(\frac{T}{\mathrm{K}} \right)^5, \tag{12.64}$$

which, for $T = 0.4\,\mathrm{K}$, is

$$\langle M_{ph}^2 \rangle \approx 6.5 \cdot 10^{-8} N^{5/3}. \tag{12.65}$$

Once $\langle M_{ph}^2 \rangle$ is calculated, the distributions are derived following in all detail the procedure for ripplons, and (12.48) can be used also for phonons, simply by replacing σ_M with the phonon value.

12.5 Thermal Properties of the Combined Excitation Spectrum

The ratio of the leading orders of the ripplon and the phonon excitation energies is

$$\frac{E_{ph}}{E_{rip}} = \frac{\frac{2\pi^6}{15}\frac{T^4}{\tilde{c}^3}}{2.25\frac{T^{7/3}}{(\hbar\omega_0)^{4/3}}} = 0.006N^{1/3}\left(\frac{T}{\mathrm{K}}\right)^{5/3}. \tag{12.66}$$

For $T = 0.4\,\mathrm{K}$ we have equal contributions when $N = 4 \cdot 10^8$. Below this fairly large but not outrageously oversized droplet, the excitation energy of the surface modes (ripplons) exceed that of the bulk modes (phonons).

When the droplet size approaches the crossover size, it will have both ripplon and phonon oscillations excited at the same time and the thermal properties will be determined by the total level density obtained by combining the ripplon and phonon functions. One can calculate the joint level density as the convolution,

$$\rho_{total}(E) = \int_0^E \rho_{rip}(E-\epsilon)\rho_{ph}(\epsilon)\mathrm{d}\epsilon. \tag{12.67}$$

As it turns out, this is not trivial to do with the same precision as with which the two input functions are known. Instead of performing the convolution, we can calculate the thermal properties of the system at the energy partition where the distribution, i.e. the integrand in (12.67), has its maximum. This will often be what we need in any case. This maximum point distribution corresponds to identical temperatures in the two subsystems (see Chap. 1) and makes microcanonical thermal properties such as energy, entropy and heat capacity of the two subsystems additive,

$$E = E_{rip} + E_{ph} = 2.247\frac{T^{7/3}}{(\hbar\omega_0)^{4/3}} + \frac{2\pi^6}{45}\frac{T^4}{\tilde{c}^3} + \cdots \tag{12.68}$$

In the limit where the phonons carry only a relatively small amount of energy, the convoluted level density in (12.67) can be calculated approximately. The procedure is similar to the one used to define the microcanonical temperature in Chap. 1 and to account for excitations of degrees of freedom that only carry marginal amounts of excitation energy, as e.g. electronic excitations. The result is that the total level density is the product of the ripplon level density and the phonon canonical partition function;

$$\rho_{total}(E) \approx \rho_{rip}(E)Z_{ph}(T(E)), \tag{12.69}$$

where T is the microcanonical temperature of the ripplons. The simplest way to calculate the phonon partition function here is to start with the level density and multiply with the Boltzmann factor and the pre-exponential factor as given in (3.47). The calculation of the expression is left as an exercise to the reader.

By the same logic as used in (12.69) one can calculate the level density of the joint system when it is the phonons that dominate the thermal properties. The result is symmetric in the subscripts relative to (12.69);

$$\rho_{total}(E) \approx \rho_{ph}(E) Z_{rip}(T(E)). \tag{12.70}$$

Also the angular momenta of the two different types of excitation add. The ratio of the ripplon and phonon contributions to the width of the angular momentum distributions is, to leading order,

$$\frac{\sigma_{ph}}{\sigma_{rip}} \equiv \sqrt{\frac{\langle M_{ph}^2 \rangle}{\langle M_{rip}^2 \rangle}} \sim 0.0098\,(T/\mathrm{K})^{7/6}\, N^{1/6}. \tag{12.71}$$

This ratio is small compared to unity up to very large droplet sizes. For our 0.4 K favourite example, unity is reached for a droplet size with a radius of 19 μm! For smaller sizes than this, one can ignore the contribution from phonons to the angular momentum.

Exercises

12.1 In a pick-up process a molecule is absorbed into a droplet in a soft collision. One may suspect that the potential dopant will just move through the droplet because it is superfluid and hence provides no resistance to the molecule in its motion. Explain (quantitatively) why this is usually not the case.

12.2 Show that for ripplon level densities,

$$\rho_{N-1}(E) \approx \rho_N(E) \exp\left(-\frac{\alpha}{N^{1/3}\left(\frac{T}{3.6\mathrm{K}}\right)}\right), \tag{12.72}$$

where α is on the order of unity.

12.3 Find the droplet size for which the second term in the exponential in (12.35) is 10% and 1% of the first term for a temperature of 0.4 K. Calculate the absolute magnitude of the correction for these two value.

12.4 Show that the distribution in (12.48) is normalized.

12.5 Complete the steps between (12.61) and (12.63)

12.6 Calculate the joint ripplon-phonon level density both for the situation when ripplons carry most of the excitation energy and when phonons do. Insert the scaling factors to get an explicit expression.

12.7 Show that the smallest size for which the average thermal excitation energy at 0.4 K exceeds the lowest excitation energy is around 100 atoms. Discuss the result in terms of the experimental desire to control the cooling of dopant molecules.

12.8 Consider the convolution of the ripplon and phonon level densities in (12.69). The validity of the result is restricted to situations where the resulting energy in the phonons does not change the temperature of the ripplon sector significantly, i.e. we require

$$T_{rip}(E) \approx T_{rip}(E - \overline{E}_{ph}(T_{rip})). \tag{12.73}$$

Otherwise energy will not be conserved to a sufficient extent.
(a) Find a criterion for the acceptable change in ripplon temperature when the phonon energy is subtracted. Use for example that the phonon partition function should not change more than a factor of two when calculated with the left and right hand side of (12.73).
(b) The criterion you have established in **a)** translates into a limit on droplet size. Find this size, using $T = 0.4$ K and the leading order terms in the expression for level densities.
(c) Repeat the calculation with the role of ripplons and phonons reversed.

Chapter 13
Phase Transitions

Melting and the inverse process of freezing are some of the best known examples of a phase transition. It occurs when a liquid or a solid (a phase, in short), is converted into the other by a change in pressure or temperature. The term solid phase is a collective term that may, for a specific material, umbrella several distinctly different structures, defined by their crystal structure and distinguished by neutron scattering, for example, if not by more mundane means, such as color, density and similar. A phase diagram maps all phases of a substance as function of pressure and temperature. These two external parameters define the phase uniquely for bulk matter. For a free nanoparticle, a more relevant set of parameters is the temperature and the number of atoms in the system. Volumes and pressures are, as discussed previously, of little interest for the properties of a free particle.

A macroscopic solid is usually an ordered phase with a periodic atomic structure, known as a crystal. This gives long range order for the position of the atoms. Liquids have unordered atomic structures, although with some short range order. Macroscopically, the difference between a liquids and a solid is the ease with which the shape of a piece of matter can change. Liquids fill the bottom of containers in the presence of gravity. If in doubt, tilt the container. Characterization by shape does not even require gravity to work. Liquids deform easily and even in zero gravity situations will one observe waves on the surface on a droplet that deform it from the spherical lowest energy shape. Consider the ripplons of the chapter on helium droplets. Shape changes are so good indicators of melting that they can be used to determine melting points of very small but still macroscopic amounts of matter very precisely. This has been used in the past by chemists to help identify chemically pure substances.

Another difference between the liquid and solid phase is that the specific volume is usually different for the two phases. Most often the density is lower for liquids, but there are exceptions, of which the most well known is water.

For nanoparticles, the definition of a crystal or a liquid is less trivial than for the bulk. What is long range order and what is short range in a particle five atoms wide?

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9_13

If one allows for modifications of the structure of the surface relative to the interior, one will have a hard time distinguishing a crystalline structure from a glassy structure where disorder is frozen in.

When particles become small enough they will be invisible to the naked eye, and fluidity will become an impractical tool for phase determination. The use of powerful microscopes is one way to push the boundary of visibility and they have been applied extensively, in particular electron microscopes that can probe much deeper than visible light microscopes. Useful as they are, microscopes have their limitations, both with respect to resolution but also because their operation tend to distort the structure of particles, in particular at the small end of the size spectrum.

Different alternative tools have been developed to study phase diagrams of very small particles. Theoretical studies can make use of numerical simulations of the type described in Chap. 9. Experimentally some of the cleanest methods are based on measurements of heat capacities. Peaks in heat capacities as a function of temperature indicate that something has happened with the collective state of the particle. If this occurs at a temperature where one expects melting and with an expected magnitude, this may actually be melting, or at least the finite size analogue of melting.[1]

A peak in the heat capacity itself is not a sufficient signature for melting because other types of transitions that give rise to these features may occur, and one prefers to have other parameters to study. In the experimental technique known as ion mobility studies, the effective geometric size of a charged particle is measured by dragging it through an inert gas with an electric field. Volume and shape changes vs. temperature can be measured this way and melting identified in complete analogy to the macroscopic technique mentioned above, provided changes are big enough.

The heat capacity peak associated with melting can be integrated over temperature to give a heat of transition, also known as the latent heat or heat of fusion and, more correctly, the enthalpy of fusion, ΔH_m. It is the difference in enthalpy of the liquid and the solid phase, usually per unit mass or mole;

$$\Delta H_m = H_l - H_s = (\overline{E}_m + PV_m) - (\overline{E}_s + PV_s), \tag{13.1}$$

where P is the pressure and V the volume of the system, and the subscripts refer to the respective phases. ΔH_m is an extensive and positive quantity which correlates with the melting temperature of the material. Data for the elements are shown in Fig. 13.1, from where one sees that the melting entropy is around unity, within a factor of two, for most elements.

A number of effects exist that will modify the simple picture of a single melting temperature. Some appear macroscopically and some only for microscopic amounts of matter. A number of materials can be supercooled, i.e. stay liquid when cooled below their usual freezing or melting temperatures, which is the temperature where the liquid material normally exists in equilibrium with the solid phase. Most materials will crystallize when they reach a critical degree of supercooling, if not before, but never below that characteristic temperature. Exceptions are glasses, as in

[1] Or freezing. Let's use the term melting generically in this discussion.

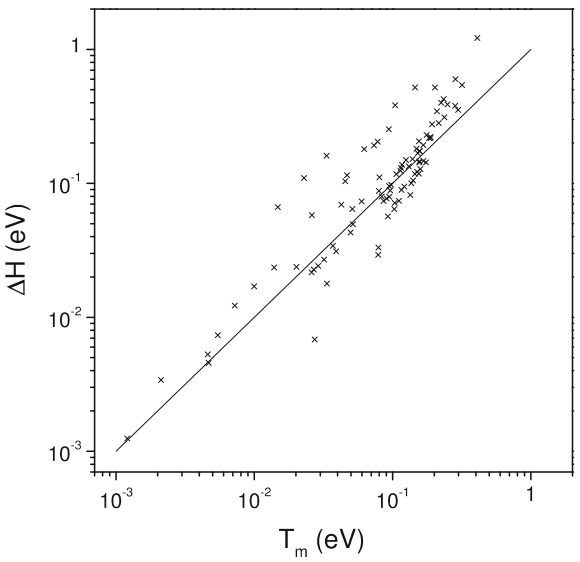


Fig. 13.1 The heat of fusion of the elements, ΔH_m per atom, vs. the melting temperature. The line is $\Delta H_m = T_m$. The values are scattered around this line, indicating that the melting entropy $\Delta H_m / T_m$ per atom is on the order of unity. This is known as Richards' rule

window panes and drinking glasses, which appear solid at room temperatures but never undergo solidification transitions from the unordered liquid state into a solid crystalline state when cooled. The usual glasses are composed of silicates, but a number of other materials can form this phase and are also called glasses.

The reason that a supercooled state can exist is very similar to the reasons a supersaturated gas can exist (see Chap. 8) viz. there is a free energy barrier for the formation of a sufficiently large nucleus or seed of the thermodynamically most stable solid phase. As soon as the seed is present, freezing happens very rapidly. The phenomenon has its analogue in supersaturation of solutions, for which concentrations can exceed the equilibrium values by large factors. Precipitation of crystallized solutes can also be induced by addition of a very small impurity. Another analogue is super-heating of liquids beyond their boiling point, a phenomenon one can experience with clean cups of water heated in microwave ovens.

However interesting these phenomena all are, only the ones that have any ramifications for the behavior of phase transitions of finite systems will be considered in the following.

13.1 Surface Melting

When the temperature of a bulk solid approaches the melting point from below but has not yet reached it, it happens for a number of materials that a thin molten layer forms at its surface. The thickness of the layer grows as the temperature approaches T_m, and it diverges at T_m, causing the whole volume to melt. Because small particles

are mainly surface, this experimentally well established phenomenon suggests they may have melting temperatures that are reduced relative to the bulk values. The phenomenon, called premelting or surface melting, is therefore of obvious interest for the study of nanoparticles.

It can be understood by comparing the free energies of a molten surface layer and the alternative completely frozen surface. Just as it costs energy to create a surface separating matter and vacuum, or vapor, the addition of a liquid-solid surface is associated with an increase in energy of the system. This surface tension is in general not equal to the liquid-vapor or the solid-vapor surface tension. A necessary condition for the existence of a liquid surface layer below T_m is that the surface tensions for the solid-liquid, liquid-vapor (or vacuum), and solid-vapor phases obey the inequality

$$\gamma_{sv} > \gamma_{sl} + \gamma_{lv}, \tag{13.2}$$

c.m.n. The solid-liquid surface tension, multiplied by the surface area, is also known as the interface energy. If the inequality in (13.2) is fulfilled, the creation of a liquid surface layer may lower the total (free) energy, making this the stable state. But it will not necessarily do so. The reduction in surface tension expressed in (13.2) is accompanied by an increase in the free energy of the volume containing the liquid layer created in the process relative to the solid phase. This increase expresses that, in the absence of surfaces, the supercooled liquid layer is not the thermodynamic ground state. The difference between the solid and liquid phase free energies can be calculated from one of the Maxwell relations

$$\left(\frac{\partial(F_s - F_l)}{\partial T}\right)_V = -(S_s - S_l), \tag{13.3}$$

or, in the linearized form, which can be used because temperatures are close to T_m,

$$F_l - F_s \approx \Delta H_m \frac{T_m - T}{T_m}, \tag{13.4}$$

using $S_l - S_s = \frac{\overline{E}_l - \overline{E}_s}{T_m} + \mathcal{O}\left(\frac{T_m - T}{T_m}\right) \approx \frac{\Delta H_m}{T_m}$. The equation reproduces the facts that the free energies of the bulk liquid phase and bulk solid phase are equal at the melting point: Below T_m the liquid phase free energy is higher than the solid's and above the roles are reversed, all as expected.

The free energies, entropies and the melting energy, ΔH_m, in the above equations are the values per volume. Since the difference in the specific volumes of a solid and its liquid is usually small, the difference of the PV-term in the enthalpies are small and in applications one can to a good approximation use tabulated values for enthalpies instead of energies, as done here.

If the thickness of the liquid layer is denoted d, the total free energy difference per unit surface area, Δf, between the situations with and without a supercooled liquid layer is therefore

$$\Delta f = (F_l - F_s)d + \gamma_{sl} + \gamma_{lv} - \gamma_{sv} = \Delta H_m \frac{T_m - T}{T_m} d + \Delta\gamma, \tag{13.5}$$

with

$$\Delta\gamma \equiv \gamma_{sl} + \gamma_{lv} - \gamma_{sv}. \tag{13.6}$$

The minimum of this expression is found for $d = 0$. This corresponds to the creation of a liquid layer, rewarded with the gain in surface energy, without paying the price associated with a volume of supercooled liquid. This is unphysical because the liquid phase cannot be infinitely thin and the two interfaces, solid-liquid and liquid-vapor, can therefore not get arbitrary close. As a minimum one needs a separation of the two interfaces of one atomic diameter to have a liquid phase. The way out of this is to introduce an interpolation between the zero and the infinite supercooled layer thickness situation of the form

$$\Delta\gamma \to \Delta\gamma\ g(d), \tag{13.7}$$

where the function $g(d = 0) = 0$ and $g(d \to \infty) \to 1$. This corresponds to an interface energy for the liquid phase of γ_{sv} for $d = 0$ and $\gamma_{sl} + \gamma_{lv}$ for sufficiently large d. The precise functional form of g is a matter of the detailed properties of the material. We will use an exponential interpolation,

$$g(d) = 1 - \exp(-d/\ell), \tag{13.8}$$

expecting that ℓ is on the order of the liquid phase atomic position correlation length. The free energy difference then becomes

$$\Delta f = \Delta H_m \left(1 - \frac{T}{T_m}\right) d + \Delta\gamma \left(1 - \mathrm{e}^{-d/\ell}\right). \tag{13.9}$$

When $\Delta H_m \left(1 - \frac{T}{T_m}\right) \ell < -\Delta\gamma$, the free energy difference in (13.9) has a negative slope at small d's and the free energy difference therefore a minimum with a negative value for some finite value of d, as required for the formation of the liquid phase. The minimum is located at the d given by

$$d = \ell \ln \left(\frac{-\Delta\gamma}{\ell \Delta H_m \left(1 - \frac{T}{T_m}\right)} \right). \tag{13.10}$$

This is the thickness of the supercooled/premelted layer (recall that $\Delta\gamma$ is negative). The numerical estimate of this length for different temperatures is left as an Exercise 13.1.

Equation (13.10) also predicts that there is a specific temperature, T_{crit}, below which the surface is completely frozen. It is found by setting the argument of the logarithm equal to one, with the result

$$T_{crit} = T_m \left(1 + \frac{\Delta\gamma}{\ell \Delta H_m}\right). \tag{13.11}$$

13.2 Melting Point Depression

A number of different types of particles have been observed to melt at lower temperatures than the bulk melting temperature of the material, reductions being roughly inversely proportional to the radii of the particles. An early observation made on supported gold clusters is shown in Fig. 13.2. The melting point depression is often described with the Gibbs-Thomson equation in terms similar to those we have already encountered. We will derive a slightly different size dependent melting point.

In analogy to the surface melting free energy, (13.5), we can write the difference of the free energies of a particle with a radius r, a solid core at radius r_s and a liquid layer between r_s and r, and a completely solid particle as

$$\Delta F = \Delta H_m \left(1 - \frac{T}{T_m}\right) \frac{4\pi}{3} \left(r^3 - r_s^3\right) + 4\pi\gamma_{sl} r_s^2 + 4\pi\gamma_{lv} r^2 - 4\pi\gamma_{sv} r^2, \tag{13.12}$$

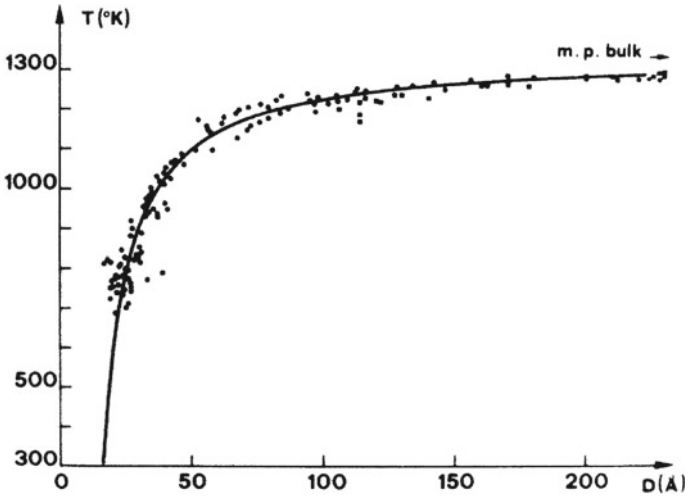


Fig. 13.2 The measured melting temperature vs. diameter for gold particles. The line is a fit to an expression where the melting point reduction is proportional to the inverse radius of the particles. Reprinted from Ph. Buffat and J.-P. Borel, *Phys. Rev. A* **13**, (1976) 2287, http://pra.aps.org/abstract/PRA/v13/i6/p2287_1. Copyright (1976) by the American Physical Society

where r_s is the radius of the solid core and r the radius of the entire particle. This assumes that melting occurs from the surface. From (13.12) we find that the free energy difference has a maximum for a radius given by the relation

$$\Delta H_m \left(1 - \frac{T}{T_m}\right) r_s = 2\gamma_{sl}, \tag{13.13}$$

and a negative second derivative with respect to r_s. The minimum is therefore located at either $r_s = 0$ or $r_s = r$. The two values are

$$\begin{aligned} \Delta F(r_s = 0) &= \Delta H_m \left(1 - \frac{T}{T_m}\right) \frac{4\pi}{3} r^3 + 4\pi r^2 (\gamma_{lv} - \gamma_{sv}), \\ \Delta F(r_s = r) &= 4\pi r^2 (\gamma_{sl} + \gamma_{lv} - \gamma_{sv}). \end{aligned} \tag{13.14}$$

Provided a solid phase exists at all (which is not a trivial condition, see Chap. 12), it will be realized at low temperatures where the first term in the expression of $\Delta F(0)$ is highest and acts to suppress the liquid phase corresponding to this minimum. As the temperature is increased, the term decreases until it reaches zero at the bulk melting temperature. This makes $\Delta f(0) < \Delta f(r)$ at T_m, which tells us that the molten particle is the most stable for some range of temperatures below T_m. The precise melting point is found by solving the equation

$$\Delta F(0) = \Delta F(r). \tag{13.15}$$

It gives the radius-dependent melting point $T_m(r)$ determined by:

$$\Delta H_m \left(1 - \frac{T_m(r)}{T_m}\right) r = 3\gamma_{sl}, \tag{13.16}$$

or

$$T_m(r) = T_m \left(1 - \frac{3\gamma_{sl}}{r \Delta H_m}\right). \tag{13.17}$$

One notes that the length ℓ introduced to derive the equation that describes surface melting is absent from (13.17) because the modification of the interface energy represented by the function g in (13.9) is not implemented here. Leaving it out does not give rise to any unphysical limits here. Inclusion of this factor is left as an exercise. Equation (13.17) is close to but not identical to the Gibbs-Thomson equation, which has a factor two instead of three multiplying γ_{sl} here.

In addition to the smooth, inverse radius dependence in (13.17), small particles composed of several tens or hundreds of atoms tend to show irregular variations in melting temperature with size. These can often be correlated with the geometric packing shell structure seen in for example rare gas clusters, but the correlation is not complete and a lot remains to be understood about this matter.

13.3 Measurements of Heat Capacities as Signatures of Melting

One of the most successful measurements of thermal properties of small, free particles is the measurement of the heat capacity of gas phase clusters. The idea behind the measurement is to use a process that depends on the energy content of the particle as an uncalibrated thermometer or, more correctly, as a calorimeter. A practical implementation proceeds like this: Particles are send through a heat bath where they thermalize to a known temperature, T_{hb}. The temperature is known because it is measured with a thermometer attached to a macroscopic heat bath. After exiting the heat bath, the particles and the thermalizing gas are separated and the clusters, that must be produced charged, are accelerated in an electric field. Acceleration separates the clusters in space according to mass (strictly speaking their mass-to-charge ratio, but the clusters are singly charged) and a single size is selected by pulsed electric fields. The selected cluster is then exposed to the light of a tunable laser, and a countable number of photons is absorbed. The resulting metastable decay is monitored with a mass dependent deflection in the electric fields of a so-called reflectron.

The end result is a measurement of the number of evaporated atoms from a given cluster size with a given initial temperature, after exposure to a flux of photons with a given energy. When photon energies are high enough to cause evaporation of several atoms, it is easy to distinguish between clusters that have absorbed n and $n+1$ photons.

If now the photon energy is changed, so will the average number of evaporated atoms. One can therefore compensate a small increase in the temperature of the heat bath, δT_{hb}, with a small decrease in photon energy, imparting a total of $n\delta h\nu$ less energy into the cluster, and end up with the same number of evaporated atoms. This means that the final energies or temperatures after equilibration and right after photon absorption are identical, and hence

$$n\delta h\nu + \delta T_{hb} C_v = 0, \tag{13.18}$$

from which the heat capacity, C_v, at T_{hb} is found. Figure 13.3 shows some of the results from a series of experiments done in Freiburg, where the method was invented. The evaporation of one atom for sodium costs about 1 eV, so different photon numbers are distinguishable with available photon energies.

The method outlined is not the only variation that will produce such thermal properties. The photo-ionization can be replaced with collision-induced dissociation (CID), for example performed in equipment designed for ion mobility studies. One advantage of that variant is that the initial temperature is somewhat better controlled. A disadvantage is that the excitation energy transferred in collisions is less well defined compared with photon energies.

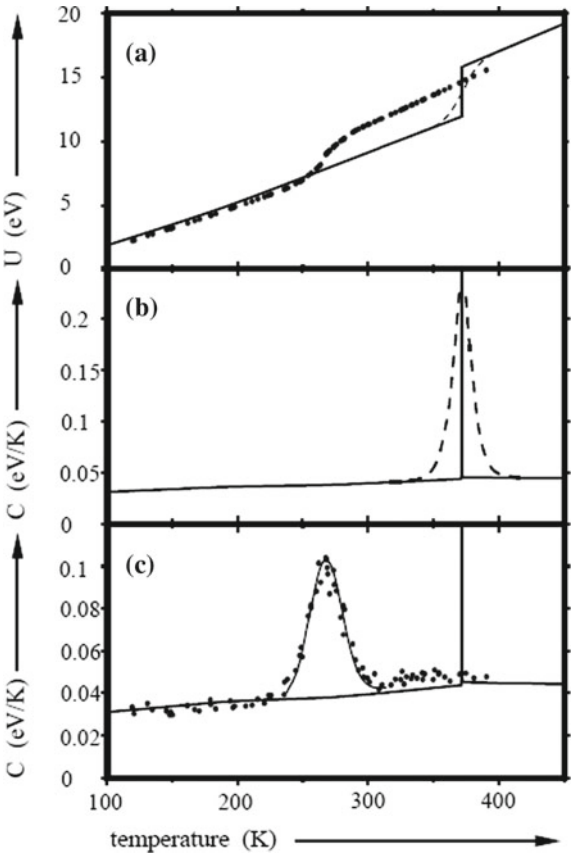


Fig. 13.3 Heat capacities of Na^{+}_{139} determined as described in the text. The lower frame shows measured heat capacities. The broad peak is associated with melting of the cluster. The middle frame shows the canonical heat capacity calculated with bulk parameters, scaled to the $N = 139$. The finite particle size smears the δ function of the bulk melting transition in the canonical representation in that frame. It is, however, still more narrow than the experimentally observed peak in the lower frame, which is microcanonical with an underlying canonical energy distribution. The integral of the peak, i.e. the latent heat, is reduced compared to the bulk value. The top frame shows the thermal energy contents of the cluster in the three different situations. Reprinted from M. Schmidt et al., *Phys. Rev. Lett.* **79** (1997) 99, http://prl.aps.org/abstract/PRL/v79/i1/p99_1. Copyright (1997) by the American Physical Society

13.4 The Lindemann Index

On the microscopic level, the solid and molten states are distinguished by the different degrees of atomic mobility. One measure of this is the Lindemann index, δ. For atom i it is defined as

$$\delta_i \equiv \frac{1}{N-1} \sum_{j \neq i} \frac{\sqrt{\langle r_{i,j}^2 \rangle - \langle r_{i,j} \rangle^2}}{\langle r_{i,j} \rangle}, \tag{13.19}$$

where $r_{i,j}$ is the distance between atoms i and j, and N is the number of atoms in the system. Thus δ_i is the root-mean-square of the relative atomic displacement. For statements about the collective state of a particle it should be averaged over all atoms;

$$\delta = \frac{1}{N} \sum_i \delta_i. \tag{13.20}$$

The averages in (13.19) are over a suitable time. Suitable time means less than astronomically long (this is easy) but also not so short that the atoms have not had the chance to move. An absolute lower limit for the sampling time is a typical vibrational period.

The reason the Lindemann index is useful is that deviations from the mean positions for a solid are basically vibrations around equilibrium positions. The amplitudes of these vibrations are much smaller than the mean distance between atoms and hence $\delta \ll 1$. In the liquid state the atoms are free to move, apart from transient caging effect of neighboring atoms. Ultimately, thermal fluctuations will produce an opening in any cage and an atom can move in to fill the void or move out to create one and thus move far beyond the typical vibrational displacement.

We can estimate the value of the Lindemann index for the solid by using harmonic oscillator properties for the vibrations. The displacement from the equilibrium position of an atom, $\langle x^2 \rangle^{1/2}$, is given by

$$\langle V \rangle = \frac{\overline{E}}{2} = \frac{T}{2} = \frac{1}{2} m \omega_D^2 \langle x^2 \rangle, \tag{13.21}$$

where $\overline{E}$ is the thermal energy, m is the mass of the atom, and the Debye frequency ω_D has been taken as a representative frequency. This gives

$$\langle x^2 \rangle^{1/2} = \sqrt{\frac{T}{m \omega_D^2}}. \tag{13.22}$$

We estimate the bond length, r, from the number density of the bulk, ρ, as

$$\frac{4\pi}{3} r^3 = \rho^{-1}. \tag{13.23}$$

Figure 13.4 shows the value of $\langle x^2 \rangle^{1/2}/r$ for a number of elements at their respective melting temperatures. The number 0.1 is seen to be a reasonable upper limit for the solid state. In spite of the somewhat hand-waving argument used to obtain (13.22), this value has turned out to be a good indicator for the phase of the system, and it is

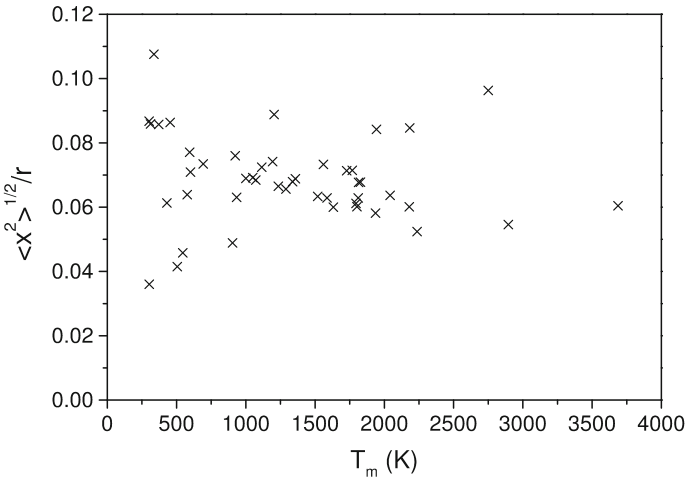


Fig. 13.4 The estimate of the Lindemann index for the solids at the melting point for a number of elements, based on (13.22)

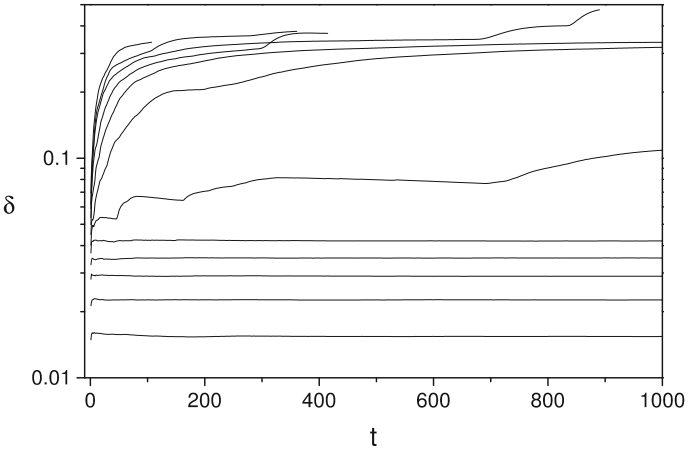


Fig. 13.5 The Lindemann index, δ, for a Lennard-Jones cluster of 55 atoms, which has a closed geometric structure. The inter-atomic potential is given in (13.24). The curves are calculated for excitation energies $E = 10\,\varepsilon$ through $120\,\varepsilon$ in steps of $10\,\varepsilon$. Several of the high energy curves are terminated prematurely because an atom evaporates. The transition from solid to liquid seems to occur between $E = 50\,\varepsilon$ and $60\,\varepsilon$

generally accepted that a value of δ below approximately 0.1 indicates a solid state, and a value above a liquid. This is known as the Lindemann criterion for melting.

To demonstrate how this works in simulations, Fig. 13.5 shows the calculated values of the Lindemann index in a MD simulation of a cluster composed of atoms interacting with the spherically symmetric pairwise Lennard-Jones potential:

$$V_{i,j} = 4\epsilon \left(\left(\frac{\sigma}{r_{i,j}} \right)^{12} - \left(\frac{\sigma}{r_{i,j}} \right)^{6} \right), \tag{13.24}$$

where i, j are labels of the atoms (see Chap. 12 for details of the potential). The lowest excitation energies in Fig. 13.5 are safely in the solid phase and the highest in the liquid, but a few of the intermediate excitation energies have less obvious designations in terms of molten or frozen. One may compare with the results in Fig. 13.9.

13.5 A Simple Model of Melting

One can model the melting by constructing a level density that gives liquid-like properties in the high energy regime and solid-like in the low energy regime. The idea is depicted schematically in the drawing in Fig. 13.6. The solid state, the narrow and deep well in Fig. 13.6, can be represented by a component in the level density with an energy dependence similar to vibrations in a solid (see Chap. 4).

$$\rho_s(E) = a_s E^{C_s - 1}, \tag{13.25}$$

where C_s is the heat capacity of the solid phase. For the sake of argument we can take the solid phase level density not just similar but identical to a harmonic vibrational

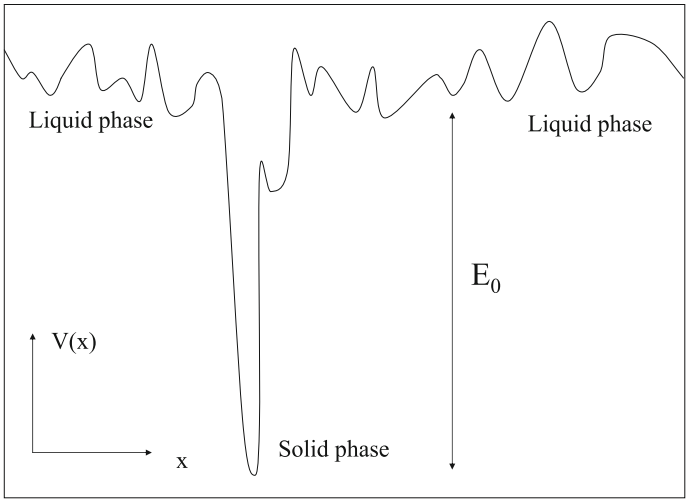


Fig. 13.6 A schematic potential energy landscape that accommodates a solid-liquid transition. The solid state is associated with the single deep potential energy well and the liquid phase with the manifold of high energy local minima

level density. The factor a_s then includes the vibrational quantum energies and a factorial, as derived in Chap. 4.

The liquid state can be described with a similar level density, with two changes. One is that there is an offset in the energy. This reflects that it takes some energy, which we will call E_0, to elevate the system to the liquid phase. The second is that the number of accessible vibrational states is much higher in the liquid phase. This must be present to compensate for the higher price paid by the energy offset for the liquid state to become the dominant. What this boils down to is a liquid state that can be described with the level density

$$\rho_l(E) = \Theta(E - E_0)a_l(E - E_0)^{C_l - 1} \tag{13.26}$$

where $a_l \gg a_s$ and Θ is the step function (see Appendix C).

In principle, melting could be caused by C_l being sufficiently much larger than C_s. If we want to have only one explanation for melting, valid for all chemical compounds, this suggestion can be ruled out. Experimentally, the liquid state heat capacity is in general not larger than the solid state of the same material. A compilation for the elements is shown in Fig. 13.7.

The difference between the two phases must therefore be found in the a-coefficients. The similarity of the solid and liquid heat capacities suggest a similar number of oscillators in the liquid phase metastable phase as in the solid, because this is what C_s and C_l count. If also the vibrational frequencies of the solid and of the metastable minima of the liquid are similar, the ratio of a_l and a_s has a straightforward interpretation as the number of metastable minima in the liquid phase. In fact, we need not assume that the vibrational frequencies of the solid and the liquid states are the same; If they are different, the ratio of the two coefficients is still the exponential of the entropy difference between the solid and the liquid state,

$$R \equiv \frac{a_l}{a_s} = \mathrm{e}^{N\Delta s}, \tag{13.27}$$

where Δs is the melting entropy per atom/monomer.

We will set $C_l = C_s = C_0$ and write the level density as

$$\rho(E) = a_s E^{C_0 - 1} + \Theta(E - E_0)a_l(E - E_0)^{C_0 - 1}. \tag{13.28}$$

It is implicit in this model that the whole volume of the particle is in the same state, liquid or solid. The relative magnitude of the two contributions to the level density in (13.28) should be considered as the probability of observing the system in either of the two phases at some sampling time. For a sufficiently large particle, melting proceeds via the creation of a small region of liquid phase somewhere in the particle (and surface melting makes the surface the most likely place), which then spreads until all solid matter has been converted into liquid. We will not go into details with this phase separation.

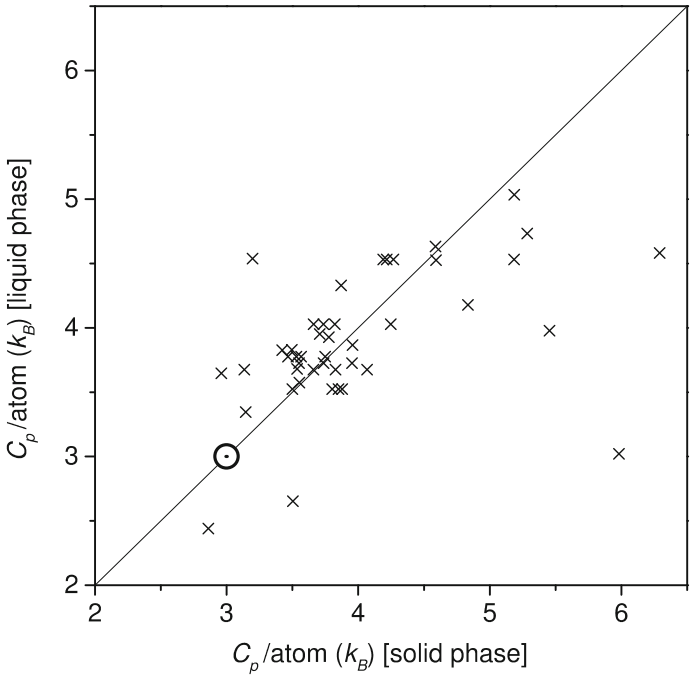


Fig. 13.7 The correlation between the solid and liquid phase heat capacities for a number of elements at the melting point. The line indicates identical values in the two phases. Clearly there is no general rule that the liquid value is above the solid. The dotted circle indicates the Dulong-Petit heat capacity

We start the analysis of (13.28) by calculating canonical properties,

$$Z = \int_0^\infty \rho(E)\mathrm{e}^{-\beta E}\mathrm{d}E = a_s T^{C_0}\Gamma(C_0) + a_l T^{C_0}\mathrm{e}^{-\beta E_0}\Gamma(C_0). \tag{13.29}$$

The thermal energy is

$$\overline{E} = C_0 T + \frac{a_l E_0 \mathrm{e}^{-\beta E_0}}{a_s + a_l \mathrm{e}^{-\beta E_0}} = C_0 T + E_0 \frac{R\mathrm{e}^{-\beta E_0}}{1 + R\mathrm{e}^{-\beta E_0}}. \tag{13.30}$$

In the low temperature limit when $R\exp(-\beta E_0) \ll 1$, the thermal energy reduces to $\overline{E} \approx C_0 T$. In the high temperature limit, $R\exp(-\beta E_0) \gg 1$, (13.30) gives the energy $\overline{E} \approx C_0 T + E_0$. We can therefore identify E_0 with the latent enthalpy of the transformation (ignoring the PV term). It is an extensive quantity in the sense that it grows with the size of the system, but for small particles it is not necessarily a linear nor even a monotonic growth.

The next question is the transition temperature. We can estimate the crossover temperature with the relation $R\exp(-\beta_m N E_0) \approx 1$, which corresponds to a thermal

energy of $\overline{E} = C_0 T_m + E_0/2$ (T_m is the melting temperature of the particle here and below, and not the bulk melting value as used above). The melting temperature is then

$$\beta_m^{-1} = T_m = \frac{E_0}{\ln R}. \tag{13.31}$$

This should not come as a surprise. The logarithm of R is the melting entropy and what the equation says is that $\Delta S = \Delta E / T_m$.

We can use the model to understand a general feature of the melting of small particles, viz. that the temperature interval over which the transition occurs is finite. Bulk matter can be assigned a single melting temperature, aptly named the melting point, apart from the surface melting effect. This is actually only an approximation, albeit a very good one when particle numbers are comparable to Avogadro's constant, and as such it gets worse the smaller the particle becomes. We can use the width of the peak in the heat capacity as a measure of the temperature interval over which the particle melts. The heat capacity in our model is

$$C \equiv \frac{\partial \overline{E}}{\partial T} = C_0 + \left(\frac{E_0}{T}\right)^2 \frac{R\mathrm{e}^{-\beta E_0}}{\left(1 + R\mathrm{e}^{-\beta E_0}\right)^2}, \tag{13.32}$$

which has a maximum of

$$C_{max} = C_0 + \frac{1}{4}\left(\frac{E_0}{T_m}\right)^2. \tag{13.33}$$

The width of the transition region, δT_m, is estimated by approximating the heat capacity peak with a triangular shape and integrating to get the latent heat:

$$\frac{1}{2}\delta T_m \frac{1}{4}\left(\frac{E_0}{T_m}\right)^2 \approx \Delta H_m. \tag{13.34}$$

Using $\Delta H_m \approx E_0$, this gives

$$\delta T_m \approx 8\frac{T_m^2}{\Delta H_m}. \tag{13.35}$$

If the data if Fig. 13.1 are taken to be representative, we can set $\Delta H_m \sim N T_m$ and get

$$\delta T_m \sim \frac{8}{N} T_m. \tag{13.36}$$

This relation is a little more reliable that than suggested by the $\sim$sign. It predicts a fairly broad transition region, e.g. 25% of the melting temperature for an $N = 32$ particle. If we use that the width of a melting transition must be less than the melting temperature itself in order to identify it as a transition at all, (13.36) predicts that

particles below $N = 8$ will not have a melting transition. Other estimates in the literature give a lower limit of $N = 6$.

Another very interesting consequence of the model, and indeed for any realistic description of melting, occurs for microcanonical systems. The conserved total energy means that when the particle is excited into a high-lying local potential minimum, the energy must be supplied by the vibrational degrees of freedom. The excitation into a molten state therefore reduces the vibrational energy, including the kinetic energy. This lowers the temperature of the particle. And because the promotion to the molten state occurs more frequently with higher total energy, one ends in the seemingly paradoxical situation that an increase in excitation energy causes a drop in temperature, i.e. the heat capacity becomes negative. The paradox is only apparent. Freezing processes where the temperature increases just reflect the fact that the heat of fusion is deposited in the system itself. It has nowhere else to go in a microcanonical system. Reversing the process makes similar sense; a melting where the temperature decreases simple reflects that the heat of fusion needs to be supplied from somewhere.

Because the phenomenon appears only in the microcanonical ensemble, one may suspect that the effect is an artifact of the definitions of T and C. It turns out that negative heat capacities is a robust effect, insensitive to precise definitions of both T and C. We can therefore demonstrate the presence of the negative heat capacity with a calculation of the microcanonical temperature using our favourite expression without any qualms about this choice. The relevant energies will be above E_0 and to simplify of notation the Θ function can be dropped from the expression. We get:

$$T^{-1} = \frac{\partial \ln(\rho(E))}{\partial E} = (C_0 - 1)\frac{E^{C_0-2} + R(E - E_0)^{C_0-2}}{E^{C_0-1} + R(E - E_0)^{C_0-1}}. \tag{13.37}$$

A few manipulations give

$$T = \frac{E - E_0}{C_0 - 1}\frac{E_0}{C_0 - 1}\frac{1}{1 + R\left(1 - \frac{E_0}{E}\right)^{C_0-2}}, \tag{13.38}$$

The reciprocal of the heat capacity is calculated to be

$$C^{-1} = \frac{\partial T}{\partial E} = \frac{1}{C_0 - 1}\left(1 - \frac{E_0}{E}\frac{E_0}{E - E_0}(C_0 - 2)\frac{R\left(1 - \frac{E_0}{E}\right)^{C_0-2}}{1 + R\left(1 - \frac{E_0}{E}\right)^{C_0-2}}\right). \tag{13.39}$$

The potential for a negative value is signalled by the negative sign in the bracket in the right hand side of this equation. The minimum of C^{-1} is reached at an energy given approximately by

$$R\left(1 - \frac{E_0}{E_{min}}\right)^{C_0-2} \approx 1 \Rightarrow E_{min} \approx \frac{E_0}{1 - R^{-1/(C_0-2)}}. \tag{13.40}$$

This is above E_0 as required for the simplification made to avoid the x-erei of the step function. With this energy the reciprocal heat capacity is

$$C^{-1} = \frac{1}{C_0 - 1}\left(1 - \frac{C_0 - 2}{4}\left(R^{1/(C_0-2)} - 1\right)\left(1 - R^{-1/(C_0-2)}\right)\right). \tag{13.41}$$

For sufficiently large R and fixed C_0 this will be negative, and hence will also C.

We can find the limiting value of R_{min} for which C is negative. It occurs at

$$1 - \frac{C_0 - 2}{4}\left(R_{min}^{1/(C_0-2)} - 1\right)\left(1 - R_{min}^{-1/(C_0-2)}\right) = 0, \tag{13.42}$$

or

$$R_{min} \approx \left(1 + \frac{2}{\sqrt{C_0 - 2}}\right)^{C_0-2}. \tag{13.43}$$

For the numerical example of the model shown in Fig. 13.8 the first appearance of a negative slope in the $T(E)$ curve is calculated from (13.43) to occur when R reaches the value $3.5 \cdot 10^8$ for $C_0 = 100$, in good agreement with the more imprecise number obtained from inspection of the figure, which shows the caloric curves calculated also with $C_0 = 100$ and several different values of R.

The values of R used for this example may seem unphysically big. They are not. The highest value used in Fig. 13.8, $R = 10^{12}$, corresponds to the total entropy change of $\ln R = 27.6$. If, for simplicity, we translate the heat capacity of $C_0 = 100$ to a number of atoms with the harmonic oscillator value $s = 3N - 6$, the melting entropy per atom is less than 1, which should be compared to the observed bulk values

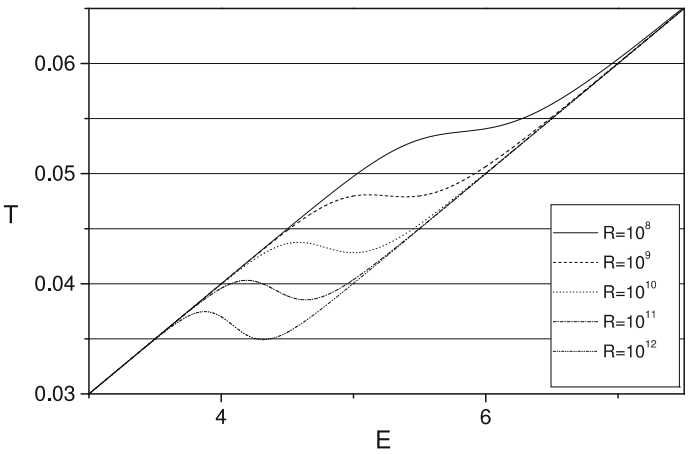


Fig. 13.8 The microcanonical temperature for a model system with the level density given by (13.28) with $C_0 = 100$ and $E_0 = 1$. The different values of R give rise to different depths of the back-bending of the caloric curve. For $R = 10^8$ the curve is without back-bending, but already at $R = 10^9$ does the first very shallow minimum appear. Energy and temperature units are E_0

that vary around unity (see Fig. 13.1). Hence, negative heat capacities are more the rule than the exception.

The fact that R will be sufficiently large even for smaller systems will, everything else equal, ensure that it will also be so for larger systems. The condition for negative microcanonical heat capacity can be stated as

$$\Delta S > \ln(R_{min}) \approx 2\sqrt{C_0 - 2}. \tag{13.44}$$

Because ΔS is roughly equal to N which is also an approximate value for C_0, this inequality tells us that a back-bending caloric curve as the one shown in Fig. 13.9 occurs even more pronounced for larger particles. The phenomenon only disappears at the sizes where phase separation appears.

An example of negative heat capacity based on a specific system is given in Fig. 13.9, where a MD simulation of the caloric curve for the $N = 55$ Lennard-Jones cluster is shown. The temperature and the heat capacity are sampled as described in Chap. 9. The back-bending of the caloric curve is seen very clearly. Because the temperature is the microcanonical version in the simulations (the reciprocal of the derivative of the level density with respect to the energy), the level density can be found by integration. In the region with negative heat capacity the logarithm of the level density, which we recall is basically the entropy, will have a section with positive curvature, called a convex intruder and shown in Fig. 13.10 (see also Exercise 13.7). All other parts of the curve are concave, corresponding to positive heat capacities. The caloric curves for neighboring cluster sizes (not shown) are very similar to that of $N = 55$.

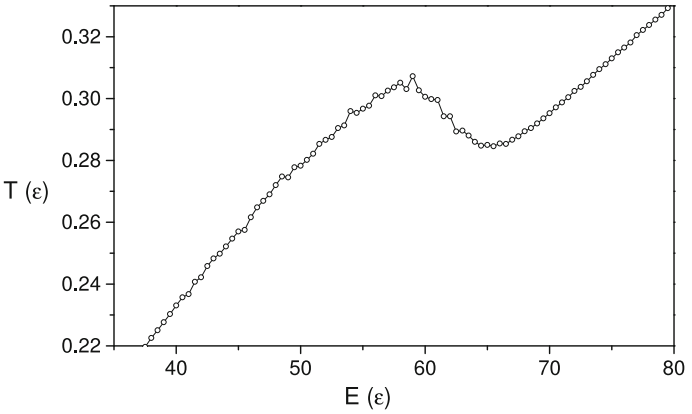


Fig. 13.9 The microcanonical caloric curve for the Lennard-Jones $N = 55$ cluster. Each point is sampled over 10^8 time steps except for a few points at the high energy end where lower statistics was used

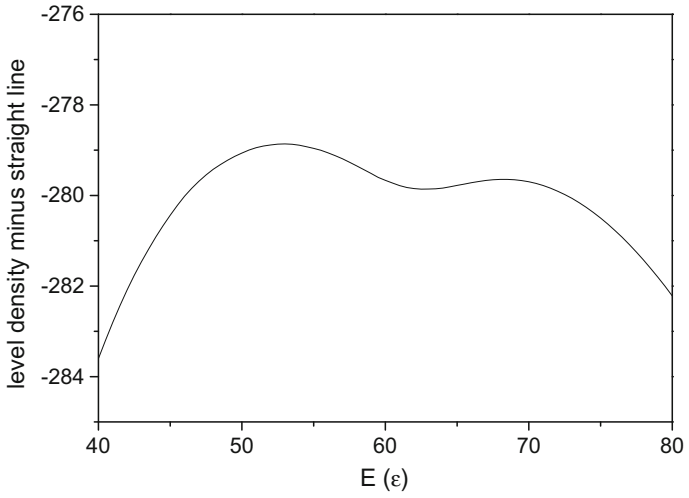


Fig. 13.10 The logarithm of the level density of the $N = 55$ Lennard-Jones cluster around the melting transition energy. A straight line has been subtracted to show the presence of the convex intruder

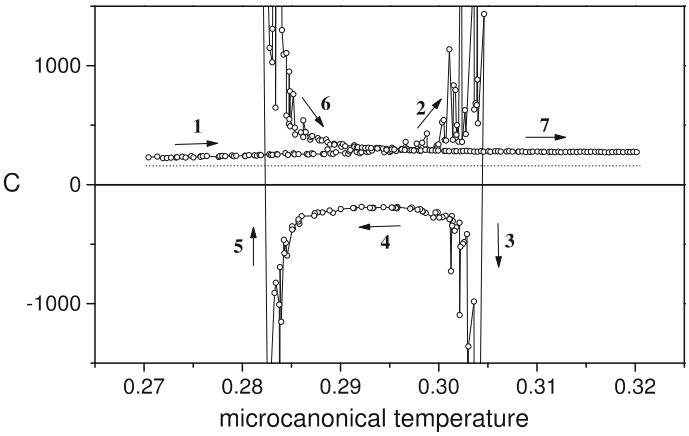


Fig. 13.11 The microcanonical heat capacity for the Lennard-Jones cluster $N = 55$. The numbered arrows give the direction of increasing energy. The dotted line is the harmonic oscillator value

The corresponding heat capacity is shown in Fig. 13.11. The divergences that appear are seen to be located at the temperatures corresponding to the local maximum and minimum in Fig. 13.9, as expected.

13.6 Solid-to-Solid Phase Transitions

Any phase transitions that come associated with a latent heat is called first order and is characterized by a heat capacity peak. Melting-freezing is an intuitively well understood example of this. Another type of transitions is called second order and are associated with correlation lengths that diverge at the transition temperature. Much joy has been derived from that phenomenon, and work on solid-to-solid transitions tend to focus on this. For nanoparticles, in which divergent lengths are excluded by their finite size, they only appear in a very disguised form, if at all. But solid-to-solid phase transitions may also be first order and those transitions will have a chance to occur in nanoparticles.

One possible origin of a first order solid-to-solid phase transition is the existence of a high energy solid modification that has softer vibrational modes than the low energy structure. Such structural phase transitions occur in magnetic substances, for example. The situation is illustrated in Fig. 13.12. The smaller level spacing in the high temperature phase gives it a higher entropy which compensates for the extra energy needed to excite it from the ground state.

There is no compelling argument why a solid-to-solid phase transition should happen, and neither does it for all materials. But if it does, the fundamental picture is, from the point of view of the statistical properties, a variation of the model derived for the liquid-solid transition. Restricting ourselves to the case where the heat capacities of the two phases are identical, the level density can be written as

$$\rho(E) = a_{s,1} E^{C_0 - 1} + \Theta(E - E_0) a_{s,2} (E - E_0)^{C_0 - 1}. \tag{13.45}$$

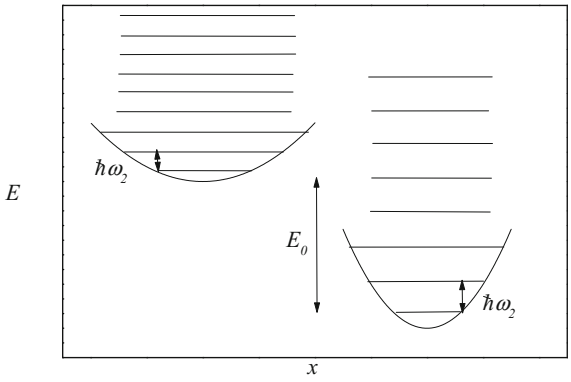


Fig. 13.12 A schematic representation of a state diagram that can give rise to a solid-to-solid first order phase transition. The lower ground state energy has higher vibrational frequencies, indicated by the larger level spacing in the one dimensional harmonic oscillator, compared with the high temperature phase. The reduced level spacing of the high temperature phase is the driving entropic force for the transition

Formally, the only difference from the solid-liquid level density in (13.28) is to replace the coefficient in front of the liquid phase contribution to the the level density by a second solid phase coefficient. For harmonic oscillators and to next to leading order in energy, the two coefficients are

$$a_{s,i} = \frac{1}{(C_0 - 1)! \prod_{j=1}^{C_0} \hbar\omega_j^i}. \tag{13.46}$$

If the geometric average of the (high energy) phase 2 phonons is less than for phase 1, the potential for a phase transition is present.

Whether or not this potential is realized depends on the relative magnitude of the two coefficients, in complete analogy to the discussion of liquid-solid transitions. The only difference here is that the value of R is now determined by the ratio of vibrational frequencies, without the need to introduce an extra parameter in the form of an empirical entropy difference. The value becomes

$$R \equiv \frac{a_2}{a_1} = \mathrm{e}^{N \ln(\overline{\omega}_1/\overline{\omega}_2)}, \tag{13.47}$$

where $\overline{\omega}_i$ is the geometric average of the vibrational frequencies in phase i. This is in fact exactly the same equation as (13.27) if the difference in entropy is assigned to the difference in vibrational frequencies. We can use this result to calculate the finite size particle width of the transition. Expressing the differences in frequencies as $\overline{\omega_1} = (1 + \alpha)\overline{\omega_2}$, and a value of α which is small compared to unity, one gets the analogue of (13.36) for the width of the structural phase transition

$$\delta T_m \approx 8 \frac{T_m}{N\alpha}, \tag{13.48}$$

corresponding to a well defined melting temperature at sizes above

$$N \approx \frac{8}{\alpha} = \frac{8}{\frac{\overline{\omega}_1}{\overline{\omega}_2} - 1}. \tag{13.49}$$

This will be a larger number than the 8 previously suggested for a melting-freezing transition. The critical size where phase transitions appear can not be too small for another reason, viz. that these transitions are related to changes in crystal structures, and a particle must contain a not too small number of atoms before these are established.

We can also calculate the criterion for the existence of negative heat capacity with (13.43), which is not specific to liquid-solid transitions. Assuming for simplicity also here a small relative difference between the two sets of vibrational frequencies gives

$$N \approx \frac{12}{\alpha^2} = \frac{12}{\left(\frac{\overline{\omega}_1}{\overline{\omega}_2} - 1\right)^2}. \tag{13.50}$$

This can be a significant particle size in terms of the number of atoms.

There are cases where α is not that small. A good example is tin, which comes in two varieties (three, but the third can be ignored here), α (grey) and β (white) tin below and above 13.2 °C, respectively. The low temperature allotrope has the highest Debye temperature of the two, 232 K, consistent with the condition for a structural phase transition. The Debye temperature of the other allotrope is 132 K, and the enthalpy of transformation is 2.1 kJ/mol. This gives a calculated transition temperature of

$$T_m = \frac{\Delta H_m}{\ln\left(\frac{\omega_{D,\alpha}}{\omega_{D,\beta}}\right)} = 430 \text{ K}, \tag{13.51}$$

which should be compared with the measured 286 K. The agreement is not striking, but neither is it unreasonable considering the simplifications in the modelling.

Exercises

13.1 The depth of a liquid, premolten surface layer can be estimated from some approximate relations between the quantities that enter (13.10). Because most of the quantities enter logarithmically, we can be pretty crude and still get a reasonable estimate. The difference in surface tensions can be set equal to a typical surface tension, $\Delta\gamma \simeq \gamma$, and this can be related to the heat (enthalpy) of vaporization (see Chap. 8). The heat of vaporization is roughly 20 times the melting temperature, and the melting temperature is on the order of the heat of fusion. Finally, ℓ can be set to the monomer radius, r_1, for the sake of the argument. With these values, find the thickness of the surface layer in units of the monomer radius, for $T = 0.99\, T_m$, $T = 0.999\, T_m$ and $T = 0.9999\, T_m$.

13.2 Use similar approximations as in Exercise 13.1 to find a numerical value for the temperature at which the first molten surface appears.

13.3 Modify (13.12) with the factor that corresponds to the function g in (13.9) and find the melting point. Find the condition the surface tensions etc. must fulfill to create a liquid layer.

13.4 Use the data in Fig. 13.2 and a tabulated value of the melting enthalpy of gold to estimate the liquid-solid surface tension.

13.5 Make an order of magnitude estimate of the maximum value the Lindemann index.

13.6 Estimate R and the melting entropy for the Lennard-Jones cluster $N = 55$ from the simulation data shown in Fig. 13.9. Assume that when the system is at the melting point, the system is half the time in the liquid phase and the other half in the solid phase. Use this to calculate the variance of the energy at this temperature. Estimate

the highest value of the heat capacity in the melting region with (1.70). Compare the result with the number derived from (13.33).

13.7 Use the definition of the microcanonical temperature to show that the logarithm of the level density has a negative curvature when the heat capacity is positive (is concave) and vice versa.

13.8 In the text it is claimed that (13.47, 13.27) are identical if entropy differences are due to the difference in vibrational frequencies alone. Show this. Hint: The use of entropy implies that you can use canonical properties to solve the problem. Next, give the conditions on the difference of frequencies for a phase transition to exist.

Appendix A
Additional Reading

General Statistical Physics

L. D. Landau and E. M. Lifshitz, Statistical Physics, vol. 1. This is a fundamental textbook in statistical physics. It derives and applies the rules to real world systems, i.e. molecules, crystals etc. Given the time it was written, it necessarily misses out on a number of more recent developments. As the other volumes in the series, this is not recommended as the first textbook on the subject.

D. A. McQuarrie, Statistical Physics. A general purpose textbook on the subject. It contains a number of applications of the different ensembles.

R. C. Tolman, The Principles of Statistical Physics. A textbook from the first part of the 20th century. It contains a thorough discussion of classical statistical physics.

D. Gross, Microcanonical Thermodynamics. A volume with a number of examples from the author's own work on the thermodynamics of small systems.

C. Garrod, Statistical Mechanics and Thermodynamics. Another general purpose textbook that covers precisely what the title says at the undergraduate student level.

G. H. Wannier, Statistical Physics. Discusses both the fundamentals of statistical physics and applications with great physical insight.

K. Huang, Statistical Mechanics. Treats thermodynamics, statistical physics and has a special topics part. Not for the equation-shy reader.

R. K. Pathria, Statistical Mechanics. A thorough introduction to a number of subjects, starting from scratch. A number of interesting subjects, such a the Ising model, renormalization group theory, virial coefficients and magnetic systems that are not covered in this work are introduced with the emphasis on the physical content.

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9

Classical Mechanics

H. Goldstein, Classical mechanics. A classical work in more than one sense of the word, this volume has taught generations of students analytical mechanics.

Clusters etc.

H. Haberland, Ed., Clusters of Atoms and Molecules. A collection of chapters written by leading experts in their fields on a wide number subjects related to the physics and chemistry of clusters. It gives a broad overview of the status of the field in the early 1990s, and is not focused on the statistical physics aspects.

R. L. Johnston, Atomic and Molecular Clusters. A more recent textbook which covers a lot of ground, with all the pros and cons resulting from that philosophy.

K. D. Bonin and V. V. Kresin, Electric Dipole Polarizabilities of Atoms, Molecules and Clusters. A small volume devoted to the theory and experiments of the leading long range interaction between two particles of which at least one is neutral. The only drawback is the use of Gaussian units.

H. Vehkamäki, Classical Nucleation Theory in Multicomponent Systems. Perhaps not all you ever wanted to know about nucleation, but it certainly goes a long way, with a strong emphasis on systems containing more than one type of molecule. Approaches the problem from the thermodynamic side.

Fermionic Degrees of Freedom

A. J. Cole, Statistical Models for Nuclear Decay. Gives a number of results useful for the statistical treatment of quantum systems, derived with no more fuzz than necessary.

A. Bohr and B. R. Mottelson, Nuclear Structure. This two-volume work is a standard reference for nuclear physics, and contains a number of results pertaining to finite fermionic systems. The material is presented in a condensed but comprehensive manner and is suitable for a student with a good knowledge of quantum physics.

J. M. Blatt and V. F. Weisskopf, Theoretical Nuclear Physics. This is a classic among textbooks on nuclear physics.

N.W. Ashcroft and N.D. Mermin; Solid state Physics. A standard textbook in its field. The first two chapters give details about simple Fermi gasses beyond the material presented here.

Mathematics and Numerical Methods

M. Abramowitz and I. A. Stegun (Eds.), Handbook of Mathematical Functions. A standard reference, and for a good reason. It contains basically everything you need to know about named mathematical functions. It is essential for anybody who uses classical mathematical analysis. Legal copies are available for free on the internet.

I. S. Gradshteyn & I. M. Ryzhik, Table of integrals, Series, and Products. Another standard volume. With this and the Handbook you will need no other math handbooks.

M. L. Boas, Mathematical Methods in the Physical Sciences. This is a textbook at an introductory level. It contains a large amount of useful techniques and results, explained in a straightforward manner.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in (insert your preferred programming language here). A volume which will give you the ability to calculate numerically basically anything your computing power will allow you to calculate.

P. Stoltse, Simulations. Unfortunately out of print, but contains detailed instructions for numerical simulations, in particular molecular dynamics simulations.

Physical Chemistry

P.W. Atkins and R.S. Friedman, Molecular Quantum Mechanics. Starts from scratch and covers what the title promises. The reader who wants to be an expert spectroscopist will need to supplement it.

Internet Resources

David Wales and collaborators have compiled an extensive list of ground state configurations and energies for clusters of particles interacting with different potentials. They can be downloaded from
http://www-wales.ch.cam.ac.uk/CCD.html

Compilers for different coding languages can be downloaded for free from GNU webpages. Originally written for the UNIX-like GNU operating system, they have been adapted to a number of operating systems. More information at
https://gcc.gnu.org/

The NIST Digital Library of Mathematical Functions gives you the equivalent of Handbook of Mathematical Functions, but made for the internet.
https://dlmf.nist.gov/

Appendix B
Constants of Nature and Conversion Factors

$\hbar = c = k_B = 1$ is a useful system of units but to get in touch with macroscopic life it is necessary to convert to physical world units. Here are some of the most important.

Constants

$1\ \mathrm{u} = 1.660539 \cdot 10^{-27}\ \mathrm{kg}$	Atomic mass unit
$N_A = 6.02214 \cdot 10^{23}\ \mathrm{mol}^{-1}$	Avogadro's number
$k_B = 1.38065 \cdot 10^{-23}\ \mathrm{J/K}$	Boltzmann's constant
$\varepsilon_0 = 8.854187817... \cdot 10^{-12}\ \mathrm{F/m}$	Vacuum permittivity
$m_e = 9.10938 \cdot 10^{-31}\ \mathrm{kg}$	Electron mass
$e = 1.602177 \cdot 10^{-19}\ \mathrm{C}$	Elementary charge
$\mu_0 = 4\pi \cdot 10^{-7}\ \mathrm{N/A^2}$	Vacuum permeability
$\hbar = 1.054571 \cdot 10^{-34}\ \mathrm{J}\,s = 0.6582119\ \mathrm{eVfs}$	Planck's constant
$c = 299792458\ \mathrm{m/s} = 299.792458\ \mathrm{nm/fs}$	Speed of light
$\sigma_{SB} = \frac{\pi^2 k_B^4}{60\hbar^3 c^2} = 5.670 \cdot 10^{-8}\ \mathrm{Wm^{-2}K^{-4}}$ $= \left(\frac{1}{64.80\ \mathrm{K}}\right)^4\ \mathrm{Wm^{-2}} = \left(\frac{1}{129.7\ \mathrm{K}}\right)^4\ \mathrm{eVs^{-1}\AA^{-2}}$	Stefan-Boltzmann's constant
$\mu_B = 9.27401 \cdot 10^{-24}\ \mathrm{J/T} = 0.671713\ \mathrm{K/T}$	Bohr magneton

Energy Conversion

$1\ \mathrm{cm^{-1}} = 1.43877\ \mathrm{K} = 1.239842 \cdot 10^{-4}\ \mathrm{eV} = 1.98645 \cdot 10^{-23}\ \mathrm{J}$
$1\ \mathrm{eV} = 1.602176 \cdot 10^{-19}\ \mathrm{J} = 11604.5\ \mathrm{K} = 8065.544\ \mathrm{cm^{-1}}$
$1\ \mathrm{Hartree} = 27.21138\ \mathrm{eV}$

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9

Derived Constants

$\alpha = \frac{e^2}{4\pi\varepsilon_0}\frac{1}{\hbar c} = 1/137.035999$	Fine structure constant
$\frac{e^2}{4\pi\varepsilon_0} = 14.39965\ \text{eVÅ} \approx \hbar c/137$	Electrostatic constant
$a_0 = \frac{\hbar^2}{m_e \frac{e^2}{4\pi\varepsilon_0}} = \frac{\hbar}{m_e c\alpha} = 0.5291772\ \text{Å}$	Bohr radius
$e^2 a_0^2/\text{Hartree} = 1.648777 \cdot 10^{-41}\ \text{C}^2\text{m}^2\text{J}^{-1}$	Atomic unit of polarizability

Appendix C
Mathematical Help

C.1 The Euler-Maclaurin Summation Formula

In statistical physics one often encounters sums of the type

$$\sum_n f(n), \tag{C.1}$$

where $f(n)$ can be the Boltzmann factor of a certain energy determined by the quantum number n, which is an integer. The summand may also include a degeneracy factor that depends on n. These kinds of sums are not always possible to perform in closed form and it is very useful to have some approximate formula to calculate them. The Euler-Maclaurin formulas provide the answer. They generalize and formalize the approximation

$$\sum_n f(n) \approx \int f(n)\mathrm{d}n, \tag{C.2}$$

which is good for functions that vary slowly with n. One version of the equations states that a sum from a to $b = a + mh$ in steps of h can be approximated as

$$\begin{aligned}\sum_{k=0}^{m} f(a+kh) \approx \frac{1}{h}\int_a^b f(x)\mathrm{d}x + \frac{1}{2}(f(a)+f(b)) \\ + \sum_{p=1}^{n-1} \frac{h^{2p-1}}{(2p)!} B_{2p}\left(f^{(2p-1)}(b) - f^{(2p-1)}(a)\right)\end{aligned} \tag{C.3}$$

where $f^{(2p-1)}$ is the $(2p-1)$'th derivative of f, and B_{2p} are the Bernoulli numbers. The order of approximation, $n-1$, is the user's choice and determines the limits one can set on the error in the formula. The error is

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9

$$\frac{h^{2n}}{(2n)!}B_{2n}\sum_{k=0}^{m} f^{(2n)}(a+kh+\theta h), \tag{C.4}$$

where $0 < \theta < 1$. The first few even Bernoulli numbers are $B_0 = 1,\ B_2 = 1/6,\ B_4 = -1/30,\ B_6 = 1/42,\ B_8 = -1/30$, and $B_{10} = 5/66$. This gives the summation formula

$$\sum_{k=0}^{m} f(a+kh) \approx \frac{1}{h}\int_a^b f(x)\mathrm{d}x + \frac{1}{2}(f(a)+f(b)) + \tag{C.5}$$
$$\frac{h}{12}\left(f^{(1)}(b) - f^{(1)}(a)\right) - \frac{h^3}{720}\left(f^{(3)}(b) - f^{(3)}(a)\right) + \cdots$$

with an error of

$$\frac{h^6}{30240}\sum_{k=0}^{m} f^{(6)}(a+kh+\theta h), \tag{C.6}$$

where, as above, $0 < \theta < 1$. If one wants more terms, the Bernoulli numbers can be calculated with (C.18) and the definition of the ζ function given below. For $n \geq 12$, $\zeta(n) = 1+\epsilon$ where $\epsilon < 3 \cdot 10^{-4}$.

C.2 Some Frequently Occurring Integrals

Integrals that involve a Boltzmann factor and some pre-exponential factor from the degeneracies or phase space factors keep popping up. This section gives you a listing of some of these integrals.

$$\int_0^\infty x^q \mathrm{e}^{-x^m}\mathrm{d}x = \frac{1}{m}\int_0^\infty x^{\frac{q+1}{m}-1}\mathrm{e}^{-x}\mathrm{d}x = \frac{1}{m}\Gamma\left(\frac{q+1}{m}\right), \quad q > -1. \tag{C.7}$$

Special cases are:

$$\int_0^\infty x^p \mathrm{e}^{-x}\mathrm{d}x = \Gamma(p+1) = p! \text{ for } p > -1 \text{ and integer}. \tag{C.8}$$

(This is the definition of the Γ function.) For p half integer:

$$\int_0^\infty x^p \mathrm{e}^{-x}\mathrm{d}x = \Gamma(p+1) = \frac{(2p)!!}{2^{p+1/2}}\pi^{1/2}, \tag{C.9}$$

where the 'double factorial' $(2p)!!$ is defined as $1 \cdot 3 \cdot 5 \cdot \ldots \cdot 2p$.

Gaussian integrals are special cases of (C.7) with $m = 2$. If also q is an even integer, $2n$, they can alternatively be calculated as

$$\int_0^\infty x^{2n}\mathrm{e}^{-x^2}\mathrm{d}x = \left(-\frac{\partial}{\partial\alpha}\right)^n \int_0^\infty \mathrm{e}^{-\alpha x^2}\mathrm{d}x\Big|_{\alpha=1} = \left(-\frac{\partial}{\partial\alpha}\right)^n \frac{1}{2}\sqrt{\pi}\alpha^{-1/2}\Big|_{\alpha=1}, \tag{C.10}$$

which, after substitution are identical to the integrals in (C.9), but may be a more portable equation.

Bosonic degrees of freedom may make it necessary to calculate integrals such as:

$$\int_0^\infty \frac{x^p\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}}\mathrm{d}x = \int_0^\infty \sum_{k=1}^\infty x^p\mathrm{e}^{-kx}\mathrm{d}x = \Gamma(p+1)\zeta(p+1). \tag{C.11}$$

$\zeta(p)$ is Riemann's ζ function,

$$\zeta(p) \equiv \sum_{n=1}^\infty n^{-p}. \tag{C.12}$$

Some special values of the ζ-function are

$$\zeta(3/2) = 2.6123... \tag{C.13}$$

$$\zeta(2) = \frac{\pi^2}{6} = 1.6449... \tag{C.14}$$

$$\zeta(3) = 1.2020 \tag{C.15}$$

$$\zeta(4) = \frac{\pi^4}{90} = 1.0823... \tag{C.16}$$

$$\zeta(5) = 1.0369 \tag{C.17}$$

$$\zeta(2n) = (-1)^{n+1}\frac{(2\pi)^{2n}}{2(2n)!}B_{2n}, \quad (n \text{ integer}) \tag{C.18}$$

B_{2n} are the Bernoulli numbers.

Other integrals of the same general nature are:

$$\int_0^\infty x^p\frac{\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2}\mathrm{d}x = \int_0^\infty x^p\sum_{k=1}^\infty \frac{\mathrm{e}^{-kx}}{1-\mathrm{e}^{-x}}\mathrm{d}x = \int_0^\infty x^p\sum_{k=1}^\infty\sum_{m=0}^\infty \mathrm{e}^{-(k+m)x}\mathrm{d}x. \tag{C.19}$$

There are $m+k$ terms of magnitude $\exp(-(k+m)x)$ in the double sum over k and m, and the integral therefore becomes

$$\int_0^\infty x^p\frac{\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2}\mathrm{d}x = \int_0^\infty x^p\sum_{n=1}^\infty n\mathrm{e}^{-nx}\mathrm{d}x = \sum_{n=1}^\infty n^{-p}\Gamma(p+1) = \zeta(p)\Gamma(p+1). \tag{C.20}$$

C.3 Stirling's Formula and the Γ Function

One often needs to calculate the factorial of a number N, defined as $N! \equiv 1 \cdot 2 \cdot ... \cdot N$, e.g. when handling the indistinguishability of particles. A direct calculation is always possible and should be considered before applying an approximation. One may for example calculate

$$\Delta_1 \ln(N!) \equiv \frac{\ln(N!) - \ln((N-1)!)}{N - (N-1)} = \ln(N), \tag{C.21}$$

or the symmetric derivative

$$\Delta_1 \ln(N!) \equiv \frac{\ln((N+1)!) - \ln((N-1)!)}{(N+1) - (N-1)} = \frac{\ln(N(N+1))}{2}, \tag{C.22}$$

if this is necessary and possible.

In other cases it may be better to use an approximation by Stirling which reads

$$(N-1)! \approx \mathrm{e}^{-N} N^{N-1/2} \sqrt{2\pi} \left[1 + \frac{1}{12N} + \frac{1}{288N^2} - \frac{139}{51840N^3} - \frac{571}{2488320N^4} + \cdots \right]. \tag{C.23}$$

The formula is asymptotic, i.e. it does not converge for finite N. In practise one includes the terms in the bracket until they start to increase.

There is a similar formula for the logarithm:

$$\ln((N-1)!) \approx (N - 1/2)\ln(N) - N + \frac{1}{2}\ln(2\pi) + \frac{1}{12N} - \frac{1}{360N^3} + \frac{1}{1260N^5} - \frac{1}{1680N^7} + \cdots \tag{C.24}$$

There is also an exact formula that represents $N!$ down to rather small arguments. It reads

$$N! = \sqrt{2\pi} N^{N+1/2} \exp\left(-N + \frac{\theta}{12N} \right), \tag{C.25}$$

where $0 < \theta < 1$. All that is required for N is that it is positive, and the formula is therefore also valid for non-integer values of N. The value of θ varies with N. If we restrict our considerations to positive values above 1, a numerical evaluation gives values for θ that are very close to unity. Hence we can write

$$N! \approx \sqrt{2\pi} N^{N+1/2} \exp\left(-N + \frac{1}{12N} \right). \tag{C.26}$$

The generalization of the factorial to non-integer arguments is the Gamma function, defined as

$$\Gamma(x) = \int_0^\infty t^{x-1} \mathrm{e}^{-t} \mathrm{d}t. \tag{C.27}$$

The function obeys the relation $\Gamma(x+1) = x\Gamma(x)$. For positive integer values of x this gives $\Gamma(x+1) = x!$ Some non-integer values are

$$\Gamma(5/2) = \frac{3}{4}\sqrt{\pi}, \tag{C.28}$$

$$\Gamma(3/2) = \frac{\sqrt{\pi}}{2}, \tag{C.29}$$

$$\Gamma(1/2) = \sqrt{\pi}, \tag{C.30}$$

$$\Gamma(1/3) = 2.67893853.., \tag{C.31}$$

$$\Gamma(2/3) = 1.35411793.., \tag{C.32}$$

$$\Gamma(3/4) = 1.22541670.., \tag{C.33}$$

$$\Gamma(1/4) = 3.62560990... \tag{C.34}$$

The function has singularities at the non-positive integers, but is otherwise a well behaved function in the entire complex plane.

C.4 The δ Function and the Step Function

The δ function is defined by the properties

$$\delta(x) = 0, x \neq 0 \tag{C.35}$$

$$\delta(0) = \infty \tag{C.36}$$

$$\int_{-\infty}^{\infty} \delta(x)\mathrm{d}x = 1. \tag{C.37}$$

It has a dimension which is the reciprocal of its argument's. Integrals involving the δ function act to pick out the integrand at the zero of its argument:

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)\mathrm{d}x = f(x_0). \tag{C.38}$$

When the argument of the function is another function with distinct zeros, integrals become

$$\int_{-\infty}^{\infty} \delta(f(x))\mathrm{d}x = \sum_i |f'(x_i)|^{-1}, \tag{C.39}$$

where f' is the derivative and x_i are the zeros of f.

The step function Θ is a dimensionless function defined as the integral of the δ function:

$$\Theta(x) = \int_{-\infty}^{x} \delta(x)\mathrm{d}x = \begin{cases} 1, & x > 0 \\ \frac{1}{2}, & x = 0 \\ 0, & x < 0 \end{cases} \tag{C.40}$$

Index

K. Hansen, *Statistical Physics of Nanoparticles in the Gas Phase*,
Springer Series on Atomic, Optical, and Plasma Physics 73,
https://doi.org/10.1007/978-3-319-90062-9