

Statistical Physics

Course material with JAVA applets

Franz J. Vesely
Computational Physics Group
Institute of Experimental Physics, University of Vienna,
Boltzmannngasse 5, A-1090 Vienna, Austria, Europe

Copyright F. Vesely 1996-2005
franz.vesely@univie.ac.at

This tutorial is regularly updated while I give the 2004/05 course.
The final version will be available towards January 31, 2005.

January 25, 2005

Chapter 1

Why is water wet?



Statistical mechanics is the attempt to explain the properties of matter from first principles.

In the framework of the *phenomenological* theory of energy and matter, in *Thermodynamics*, these properties were measured and interrelated in a logically consistent manner. The resulting theoretical body allowed scientists to write up exact relations between measured properties. However, as long as the molecular or atomic nature of matter was not included in the theory, a complete prediction of the numerical values of these observables was impossible.

To take an example: the entropy difference $S_{1,2}$ between two states of a system was operationally defined as the integral $S_2 - S_1 \equiv \int_1^2 dQ/T$ (assuming a reversible path from 1 to 2), but the numerical value of S could not be predicted even for an ideal gas. Similarly, thermodynamic reasoning shows that the difference between the specific heats C_P and C_V must exactly equal $C_P - C_V = TV\alpha^2/\kappa_T$, where α and κ_T are the expansions coefficient and the isothermal compressibility, respectively; but a specific value for C_P or C_V cannot be predicted.

Let us dream for a moment: what if we could calculate, say, the *entropy* S of a piece of matter on the basis of microscopic considerations? Assuming we could find an explicit expression for $S(E, V)$ – entropy as a function of energy and volume – then this would give us access to all the riches of thermodynamic phenomena. All we would have to do is take the inverse function $E(S, V)$ and apply the thermodynamic relations $T \equiv -(\partial E/\partial S)_V$, $C_V \equiv (\partial E/\partial T)_V$, $P \equiv -(\partial E/\partial V)_S$ etc.

Our project then is to describe the mechanics of an assembly of many ($\approx 10^{24}$) particles with mutual interactions. The only way to do this is by application of statistical methods: a rigorous

analysis of the coupled motions of many particles is simply impossible.

It should be noted, however, that with the emergence of fast computing machines it became possible to perform numerical simulations on suitable model systems. It is sometimes sufficient to simulate systems of several hundred particles only: the properties of such small systems differ by no more than a few percent from those of macroscopic samples. In this manner we may check theoretical predictions referring to simplified – and therefore not entirely realistic – model systems. An important example are “gases” made up of hard spheres. The properties of such gases may be predicted by theory and “measured” in a simulation. The importance of computer simulation for research does not end here. In addition to simple, generic models one may also simulate more realistic and complex systems. In fact, some microscopic properties of such systems are accessible neither to direct experiment nor to theory – but to simulation.

In the context of this course simulations will be used mainly for the visualisation of the statistical-mechanical truths which we derive by mathematical means. Or vice versa: having watched the chaotic buffeting of a few dozen simulated particles we will courageously set out to analyse the essential, regular features hidden in that disorderly motion.

Thus our modest goal will be to identify a microscopic quantity that has all the properties of the thing we call entropy. For a particularly simple model system, the ideal gas, we will even write down an explicit formula for the function $S(E, V)$. Keeping our eyes open as we take a walk through the neat front yard of ideal gas theory, we may well learn something about other, more complex systems such as crystals, liquids, or photon gases.

The grown-up child’s inquiry why the water is wet will have to remain unanswered for some more time. Even if we reformulate it in appropriate terms, asking for a microscopic-statistical explanation for the phenomenon of *wetting*, it exceeds the frame of this introductory treatment. – And so much the better: curiosity, after all, is the well spring of all science.

1.1 A quick resumé of thermodynamics

- **Thermodynamic state — State variables**

We can describe the state of a thermodynamic system uniquely by specifying a small number of macroscopic observables (**state variables**) X_i :
amount of substance (in mol) n ; pressure P ; volume V ; temperature T ; internal energy E ;
etc.

Equations of state of the form $f(X_1, X_2, ..) = 0$ serve to reduce the number of independent state variables:

$PV = nRT$; $E = nC_V T$ (ideal gas); $(P + an^2/V^2)(V - bn) = nRT$ (real gas; Van der Waals); etc.

- **First Law**

JOULE’s experiment (heat of friction: stirring water with a spindle driven by a weight) proved that heat is a form of energy. Mechanical energy (W) and heat (Q) may be converted both ways. Following general custom we assign a positive sign to mechanical work *performed by* the system, while heat is counted positive when it is *fed into* the system. The law of energy conservation then reads, taking into account the internal energy E ,

$$dE = dQ - dW \tag{1.1}$$

Frequently the mechanical energy occurs in the form of “volume work”, i.e. work performed against the ambient pressure when the system expands: $dW = P dV$ (piston lifting weight). It should be noted that E is a *state variable*, while Q and W may depend not only on the state of the system but also on the way it has reached that state; we may stress this by using the notation $dE = \bar{d}Q - \bar{d}W$.¹

Internal energy of an ideal gas: The expansion experiment (also by JOULE: expand air from

¹Quantities such as Q and W are often called *process variables*.

one bottle into an additional, evacuated bottle) demonstrates that $E_{idG} = C_V T$, meaning that the internal energy of an *ideal* gas depends only on the temperature T . (Do not confuse this with JOULE and THOMPSON'S throttling experiment, which relates to *real* gases.)

Since $\Delta W = 0$ we have $\Delta E = \Delta Q$; but as $T_2 = T_1$ (empirical fact) there follows $\Delta Q \propto \Delta T = 0$. Thus $\Delta E = 0$. However, since the volumes of final and initial state differ, the internal energy E cannot depend on V . In other words, for an ideal gas the formula $E = E(V, T)$ reads simply $E = E(T)$.

We know even more about $E(T)$. Feeding energy to the gas in an *isochoric* way, i.e. with $V = \text{const}$, is possible only by heating: $\Delta E = \Delta Q = C_V \Delta T$. Therefore we find, using the empirical fact that $C_V = \text{const}$, that $E(T) = C_V T$.

- **Thermal interaction — Thermal equilibrium — Temperature**

Bringing two bodies in contact such that they may exchange energy, we will in general observe a flow of a certain amount of energy in one direction. As soon as this flow has ebbed off, we have reached **thermal equilibrium** – by definition.

A measurable quantity to characterize this equilibrium state is the **temperature**: $[T_1 = T_2]_{eq}$

- **Second Law**

Not all transformations of energy that are in accordance with the First Law are really possible. The Second Law tells us which transformations are excluded. There are several logically equivalent formulations:

Impossible are those processes in which nothing happens but a complete transformation of ΔQ into ΔW . [Kelvin]

Example: Expansion of a gas against a weighted piston; here we have a complete change of ΔQ into $\int P dV$, but the end state is different from the initial one (larger V).

Impossible are those processes in which nothing happens but a transport of heat ΔQ from a colder to a warmer reservoir. [Clausius]

Example: Heat pump; here we do transport ΔQ , but in order to do so we have to feed in mechanical energy.

- **Reversible processes**

If a process is such that the inversion of the arrow of time $t \rightarrow -t$ will again refer to a possible process, this process is called reversible.

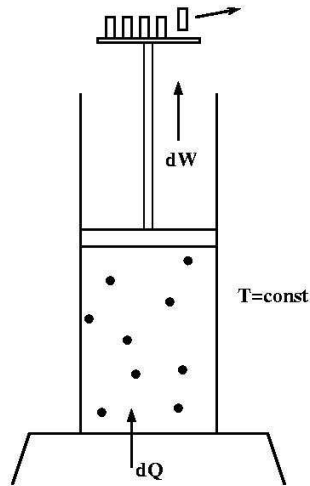
It is a consequence of the Second Law that there are processes which are not reversible.

Example 1: Reversible expansion

Let an ideal gas expand isothermally – i.e. with appropriate input of heat – against a (slowly diminishing!) weight. The internal energy of the gas remains constant, ($E_{idG} = n C_V T$), and we have

$$\Delta Q = \Delta W = \int P(V) dV = nRT \int \frac{1}{V} dV = nRT \ln \frac{V_2}{V_1} \quad (1.2)$$

This process may be reversed: by slowly increasing the weight load on the piston and removing any excess heat to a thermostatted heat bath we may recompress the gas isothermally,



to reach the initial state again.

Example 2: Irreversible expansion

Let an ideal gas expand into vacuum (JOULE's expansion experiment). Result: $T_2 = T_1$, therefore no change of internal energy. In this case no energy has been exerted and stored, but the volume has increased.

Changing the direction of time does not produce a possible process: the shrinking of a gas to a smaller volume without input of mechanical energy does not occur.

• **Entropy**

A mathematically powerful formulation of the Second Law relies on a quantity which is defined as follows:

Let $1 \rightarrow 2$ be a **reversible** path from state 1 to 2. The difference of entropies is then given by

$$\Delta S \equiv S(2) - S(1) = \int_1^2 \frac{dQ}{T} \quad (1.3)$$

S has the following important properties:

- S is a **state variable**, i.e. in any particular physical state of a system the entropy S has a well-defined value, independent of what happened to the system before. This value, however, is defined only up to an additive constant; the *Third Law* removes this ambiguity by stating that $S(T \rightarrow 0) = 0$.
- $\Delta S > \int_1^2 dQ/T$ in all *irreversible* processes.
- When a thermally isolated system is in equilibrium the entropy is at its maximum.
- Thus in a thermally isolated system (the universe being a notable special case) we have $dQ = 0$ and therefore $\Delta S \geq 0$, meaning that the entropy will never decrease spontaneously.

- As a consequence of the First Law $dE = T dS - P dV$ we have

$$\left. \frac{dE}{dS} \right|_V = T \quad (1.4)$$

But this means that we can describe thermal equilibrium in yet another way:
The flow of energy between two bodies in contact will come to an end when

$$\left. \frac{dE}{dS} \right|_1 = \left. \frac{dE}{dS} \right|_2 \quad (1.5)$$

- If two systems are in thermal equilibrium then $S_{tot} = S_1 + S_2$; entropy is additive (extensive).

- **Entropy applied:**

By scrutinizing the entropy balance in a thermodynamic process we may find out whether that process is, according to the Second Law, a possible one. In particular, we can easily decide whether the change of state is reversible or not.

Example 1: Reversible expansion

In the former experiment the entropy of the gas increases by the amount

$$\Delta S_{gas} = \frac{\Delta Q}{T} = \frac{\Delta W}{T} = nR \ln \frac{V_2}{V_1} \quad (1.6)$$

At the same time the heat reservoir delivers the energy ΔQ to the gas; this decreases the reservoir's entropy by

$$\Delta S_{res} = -\frac{\Delta Q}{T} = -nR \ln \frac{V_2}{V_1} \quad (1.7)$$

Thus the entropy of the combined system remains the same, and the reverse process is possible as well. \implies *reversible*.

Example 2: Irreversible expansion:

Let the initial and final states of the gas be the same as above. Since S is a state variable we have as before

$$\Delta S_{gas} = nR \ln \frac{V_2}{V_1} \quad (1.8)$$

This time, however, no energy is fed in, which makes $\Delta S_{res} = 0$ and thus $\Delta S_{tot} > 0$. The reverse process would therefore give rise to a decrease in entropy – which is forbidden: \implies *irreversible*.

- **Macroscopic conditions and thermodynamic potentials**

The thermodynamic state – the “macrostate” – of a gas is uniquely determined by two independent state variables. Any of the two *mechanical* variables – P or V – may be combined with one of the two *thermal* variables – T or S – to describe the state. Of course, the natural choice will be that pair of variables that is controlled, or monitored, by the experimenter. For example, if a sample of gas is contained in a thermostated (T) cylinder with a piston whose *position* (V) is directly observed, the description in terms of (V, T) is indicated. Then again, it may be the *force* on the piston, and not its position, which we measure or control; the natural mechanical variable is then the pressure P . And instead of holding the *temperature* constant we may put a thermal insulation around the system, thus conserving its *thermal energy* or, by $0 = dQ = T dS$, its *entropy*. To describe this experimental situation we will of course choose the pair (P, S) . Let us consider the possible choices:

- * (S, V) : The system's volume and heat content are controlled. This may be done by a thermally insulating the sample (while allowing for controlled addition of heat, $dQ = T dS$) and manipulating the *position* of the piston, i. e. the sample volume V . The internal energy is a (often unknown) function of these variables: $E = E(S, V)$.

Now let the macrostate (S, V) be changed reversibly, meaning that we introduce small changes dS, dV of the state variables. According to the First Law, the change dE of internal energy must then be the net sum of the mechanical and thermal contributions: $dE(S, V) = (\partial E/\partial S)_V dS + (\partial E/\partial V)_S dV$. Comparing this to $dE = T dS - P dV$ we derive the thermodynamic ("Maxwell's") relations $P = -(\partial E/\partial V)_S$ and $T = (\partial E/\partial S)_V$.

Special case $dS = 0$: For *adiabatic* expansion we see that the *internal energy* of the gas changes by $\Delta E = -\int P dV$ - just the amount of mechanical work done by the gas.

The internal energy $E(S, V)$ is our first example of a **Thermodynamic Potential**. Its usefulness stems from the fact that the thermodynamic observables P and T may be represented as *derivatives* of $E(S, V)$. We will now consider three more of these potentials, each pertaining to another choice of state variables.

- * (V, T) : The system's volume and its *temperature* are controlled.

The appropriate thermodynamic potential is then the **Helmholtz Free Energy** defined by $A(V, T) \equiv E - TS$. Using $dA = dE - T dS - S dT = -P dV - S dT$ we may derive the relations $P = -(\partial A/\partial V)_T$ and $S = -(\partial A/\partial T)_V$.

Special case $dT = 0$: Considering a process of *isothermal* expansion we find that now the *free energy* of the gas changes by $\Delta A = -\int P dV$ - again, the amount of mechanical work done by the gas.

This is worth remembering: the mechanical work done by a substance (not just an ideal gas!) that is allowed to expand isothermally and reversibly from volume V_1 to V_2 is equal to the difference between the free energies of the final and initial states, respectively.

- * (T, P) : Heat bath and force-controlled piston.

Gibbs Potential $G(P, T) \equiv A + PV$: $dG = dA + P dV + V dP = -S dT + V dP \implies S = -(\partial G/\partial T)_P$ and $V = (\partial G/\partial P)_T$.

Special case $dT = 0$: The basic experiment described earlier, *isothermal* expansion against a controlled force (= load on the piston) is best discussed in terms of G : the change ΔG equals simply $\int_1^2 V(P; T = \text{const}) dP$, which may be either measured or calculated using an appropriate equation of state $V(P, T)$.

- * (P, S) : As in the first case, the heat content is controlled; but now the *force* acting on the piston (and not the latter's *position*) is the controllable mechanical parameter.

Define the **Enthalpy** $H \equiv E + PV$: $dH = dE + P dV + V dP = T dS + V dP \implies T = (\partial H/\partial S)_P$ and $V = (\partial H/\partial P)_S$.

Special case $dS = 0$: Again, the system may be allowed to expand, but without replacement of the exported mechanical energy by the import of heat (*adiabatic* expansion). The enthalpy difference is then just $\Delta H = \int_1^2 V(P; S = \text{const}) dP$, which may be measured. (In rare cases, such as the ideal gas, there may be an equation of state for $V(P; S = \text{const})$, which allows us to calculate the integral.)

The multitude of thermodynamic potentials may seem bewildering; it should not. As mentioned before, they were actually introduced for convenience of description by the “natural” independent variables suggested by the experiment – or by the machine design: we should not forget that the early thermodynamicists were out to construct thermomechanical machines.

There is an extremely useful and simple mnemonic scheme, due to Maxwell, that allows us to reproduce the basic thermodynamic relations without recourse to longish derivations: *Maxwell’s square*. Here it is:

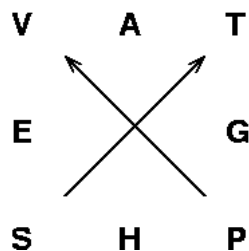


Figure 1.1: Thermodynamics in a square: Maxwell’s mnemonic device

- First of all, it tells us that $A = A(V, T)$, $E = E(S, V)$, etc.
- Second, the relations between the various thermodynamic potentials are barely hidden – discover the rules yourself: $A = E - TS$; $H = E + PV$; $G = A + PV$.
- Next, the eight particularly important *Maxwell relations* for the state variables V, T, P, S may be extracted immediately:

$$V = \left. \frac{\partial H}{\partial P} \right|_S = \left. \frac{\partial G}{\partial P} \right|_T \quad (1.9)$$

$$T = \left. \frac{\partial H}{\partial S} \right|_P = \left. \frac{\partial E}{\partial S} \right|_V \quad (1.10)$$

$$P = - \left. \frac{\partial A}{\partial V} \right|_T = - \left. \frac{\partial E}{\partial V} \right|_S \quad (1.11)$$

$$S = - \left. \frac{\partial A}{\partial T} \right|_V = - \left. \frac{\partial G}{\partial T} \right|_P \quad (1.12)$$

- Using these relations we may easily write down the description of infinitesimal state changes in the various representations given above. To repeat,

$$dE(V, S) = -P dV + T dS \quad (1.13)$$

$$dA(V, T) = -P dV - S dT \quad (1.14)$$

$$dG(P, T) = V dP - S dT \quad (1.15)$$

$$dH(P, S) = V dP + T dS \quad (1.16)$$

Chemical Potential: So far we have assumed that the mass (or particle) content of the considered system was constant. If we allow the number n of moles to vary, another important

parameter enters the scene. Generalizing the internal energy change to allow for a substance loss or increase, we write

$$dE(S, V, n) = -PdV + TdS + \mu dn \quad (1.17)$$

In other words, the *chemical potential* μ quantifies the energy increase of a system upon addition of substance, keeping entropy and volume constant: $\mu = \partial E / \partial n|_{S, V}$. Since S is itself dependent on n another, equally valid relation is usually preferred:

$$\mu = \left. \frac{\partial G}{\partial n} \right|_{T, P} \quad (1.18)$$

The name given to this state variable stems from the fact that substances tend to undergo a chemical reaction whenever the sum of their individual chemical potentials exceeds that of the reaction product. However, the importance of μ is not restricted to chemistry. We will encounter it several times on our way through statistical/thermal physics.

That thing called entropy:

We now set a definite goal for this tutorial. From an analysis of the statistical properties of many-body systems we hope to find a microscopic correlate for the quantity which thermodynamicists introduced under the name of **Entropy**.

Thus we will look out for a function $S(E, V)$ that must have these properties:

Given two systems (e.g. ideal gases) with n_1, V_1, E_1 and $n_2, V_2, E_2 = E - E_1$ which are allowed to exchange heat with the constraint that the total energy E remains constant:

- The regular flow of energy in one direction is to stop as soon as

$$\left. \frac{dS}{dE} \right|_{E_1} = \left. \frac{dS}{dE} \right|_{E_2} \quad (\text{equilibrium})$$

- The final state shall be characterized by

$$S = S_1 + S_2 \quad (\text{additivity})$$

If we succeed in finding such a $S(E, V)$, we can in principle use the inverse function $E(S, V)$ and the relations $(\partial E / \partial V)_S = -P$ etc. to build up thermodynamics from first principles!

1.2 Model systems

Statistical physics must make use of models. To simplify matters we sometimes invoke crudely simplified images of reality, which may seem to render the results of theory useless. However, many fundamental statements are all but independent of the particular model system used in their derivation. Instead, they are consequences of a few basic properties common to all many-body systems, and of the laws of statistics. Thus the following models should be understood primarily as vehicles for **demonstrating** statistical-mechanical predictions.

- **STADIUM BILLIARD**

This model, introduced by Bunimovich and Sinai, will serve us to demonstrate the existence of **chaos** even in quite small systems. Chaos, it must be understood, is a fundamental prerequisite for the application of statistical rules. It is a basic property of chaotic systems that they will acquire each one of a certain set of “states” with equal probability, i.e. with

equal relative frequency. It is this “equal a priori probability” of states which we need to proceed into the heart of Statistical Mechanics.

The stadium billiard is defined as follows. Let a “light ray” or “mass point” move about in a two-dimensional container with reflecting or ideally elastic walls. The boundaries have both flat and semicircular parts, which gives rise to an efficient mixing of the flight directions upon reflection. Such a system is chaotic, meaning that the motional degrees of freedom $\{v_x, v_y\}$ exhibit a uniform probability distribution. Since we have $v_x^2 + v_y^2 = \text{const}$ the “phase points” describing the momentary motion are distributed evenly over the periphery of a circle. The model is easily extended into three dimensions; the velocity then has three components, energy conservation keeps the phase points on the surface of a sphere, and equal a priori probability (or chaos) means that the distribution of points on this “energy surface” is homogeneous – nowhere denser or thinner than elsewhere.

Simulation: Stadium Billiard in 2 Dimensions

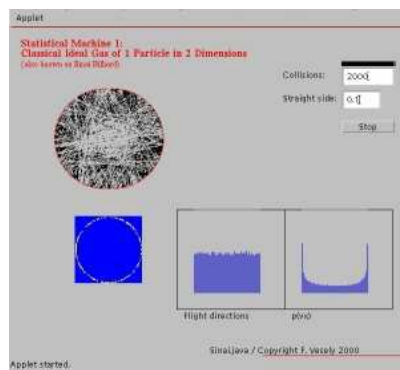


Figure 1.2: Applet STADIUM

- See the trajectory of the particle (ray); note the frequency histograms for flight direction $\phi \equiv \arctan(v_y/v_x)$ and x -velocity v_x .
 - “Chaos” is demonstrated by simultaneously starting a large number of trajectories with nearly identical initial directions: \rightarrow fast emergence of equidistribution on the circle $v_x^2 + v_y^2 = \text{const}$.
- [Code: Stadium]

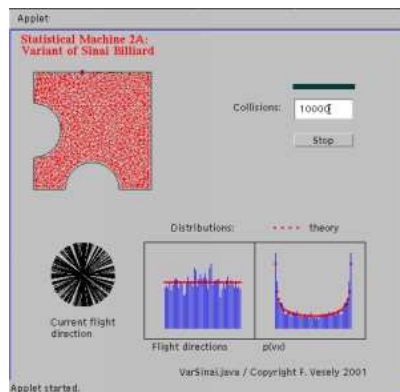


Figure 1.3: Applet VARSINAI

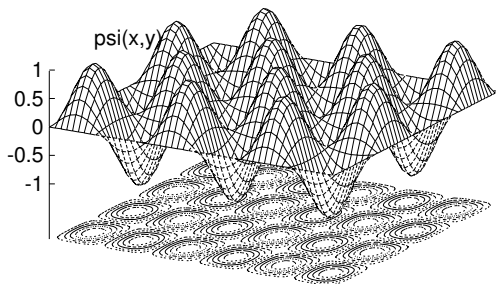


Figure 1.4: Quantum particle in a 2-dimensional box: $n_x = 3$, $n_y = 2$

Simulation: Sinai Billiard

- Ideal one particle gas in a box having randomizers on its walls
 - See the trajectory of the particle (ray); note the frequency histograms for flight direction $\phi \equiv \arctan(v_y/v_x)$ and x -velocity v_x
- [Code: VarSinai]

• CLASSICAL IDEAL GAS

N particles are confined to a volume V . There are no mutual interactions between molecules – except that they may exchange energy and momentum in some unspecified but conservative manner. The theoretical treatment of such a system is particularly simple; nevertheless the results are applicable, with some caution, to gases at low densities. Note that air at normal conditions may be regarded an almost ideal gas.

The momentary state of a classical system of N point particles is completely determined by the specification of all positions $\{\vec{r}_i\}$ and velocities $\{\vec{v}_i\}$. The energy contained in the system is entirely kinetic, and in an isolated system with ideally elastic walls remains constant.

• IDEAL QUANTUM GAS

Again, N particles are enclosed in a volume V . However, the various states of the system are now to be specified not by the positions and velocities but according to the rules of quantum mechanics. Considering first a single particle in a one-dimensional box of length L . The solutions of Schroedinger's equation

$$\mathcal{H}\Psi(x) = E\Psi(x) \quad \text{with } \mathcal{H} \equiv -\frac{\hbar^2}{2m}\nabla^2 \quad (1.19)$$

are in this case $\Psi_n(x) = (1/\sqrt{L})\sin(2\pi nx/L)$, with the energy eigenvalues $E_n = \hbar^2 n^2 / 8mL^2$ ($n = 1, 2, \dots$).

In two dimensions – the box being quadratic with side length L – we have for the energies

$$E_{\vec{n}} = \frac{\hbar^2}{8mL^2}(n_x^2 + n_y^2) \quad (1.20)$$

where $\vec{n} \equiv \{n_x, n_y\}$. A similar expression may be found for three dimensions.

If there are N particles in the box we have, in three dimensions,

$$E_{N,\vec{n}} = \frac{h^2}{8mL^2} \sum_{i=1}^N [n_{ix}^2 + n_{iy}^2 + n_{iz}^2] \quad (1.21)$$

where n_{ix} is the x quantum number of particle i .

Note: In writing the sum 1.21 we only assumed that each of the N particles is in a certain state $\vec{n}_i \equiv (n_{ix}, n_{iy}, n_{iz})$. We have not considered yet if any combination of single particle states \vec{n}_i ($i = 1, \dots, N$) is indeed permitted, or if certain \vec{n}_i, \vec{n}_j might exclude each other (Pauli principle for fermions.)

• HARD SPHERES, HARD DISCS

Again we assume that N such particles are confined to a volume V . However, the finite-size objects may now collide with each other, and at each encounter will exchange energy and momentum according to the laws of elastic collisions. A particle bouncing back from a wall will only invert the respective velocity component.

At very low densities such a model system will of course resemble a classical ideal gas. However, since there is now a – albeit simplified – mechanism for the transfer of momentum and energy the model is a suitable reference system for **kinetic theory** which is concerned with the transport of mechanical quantities. The relevant results will be applicable as first approximations also to moderately dense gases and fluids.

In addition, the HS model has a special significance for the simulation of fluids.

Simulation: Hard discs

– “Gas” of hard discs in a square box with reflecting walls. To achieve chaos even with a single moving particle the walls are decorated with “randomizers”, i.e. reflecting semicircles.

The motion of 1, 2, ... N particles is followed. [Code: Harddisks]

– Displayed: trajectory; frequency histograms for flight direction $\phi \equiv \arctan(v_y/v_x)$ and x -speed v_x .

[Code: Harddisks]

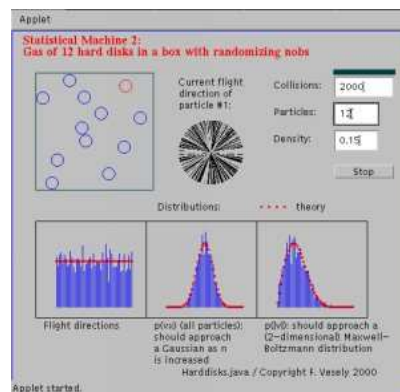


Figure 1.5: Applet HARDDISKS

Simulation: Hard spheres

– “Gas” of N hard spheres in a cubic box with reflecting walls. To achieve chaos even with a single moving particle the walls are decorated with “randomizers”, i.e. reflecting semispheres.

The motion of $1, 2, \dots, N$ particles is followed.

– Displayed: trajectory; frequency histograms for flight direction $\phi \equiv \arctan(v_y/v_x)$ and x -speed v_x .

[Code: Hspheres]

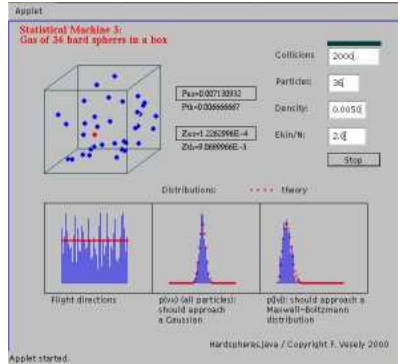


Figure 1.6: Applet HSPHERES

⇒ **Simulation 1.4: N hard discs in a 2D box with periodic boundaries.** [Code: Hdiskspbc, UNDER CONSTRUCTION]

• LENNARD-JONES MOLECULES

This model fluid is defined by the interaction potential (see figure)

$$U(r) = 4\epsilon[(r/\sigma)^{-12} - (r/\sigma)^{-6}] \quad (1.22)$$

where ϵ (potential well depth) and σ (contact distance) are substance specific parameters.

In place of the hard collisions we have now a continuous repulsive interaction at small distances; in addition there is a weak attractive force at intermediate pair distances. The model is fairly realistic; the interaction between two rare gas atoms is well approximated by equ. 1.22.

The LJ model achieved great importance in the Sixties and Seventies, when the microscopic structure and dynamics of simple fluids was an all-important topic. The interplay of experiment, theory and simulation proved immensely fruitful for laying the foundations of modern liquid state physics.

In simulation one often uses the so-called “periodic boundary conditions” instead of reflecting vessel walls: a particle leaving the (quadratic, cubic, ...) cell on the right is then fed in with identical velocity from the left boundary, etc. This guarantees that particle number, energy and total momentum are conserved, and that each particle is at all times surrounded by other particles instead of having a wall nearby. The situation of a molecule well within a macroscopic sample is better approximated in this way.

Simulation: Lennard-Jones fluid

N Lennard-Jones particles in a 2d box with periodic boundary conditions.

[Code: LJones]

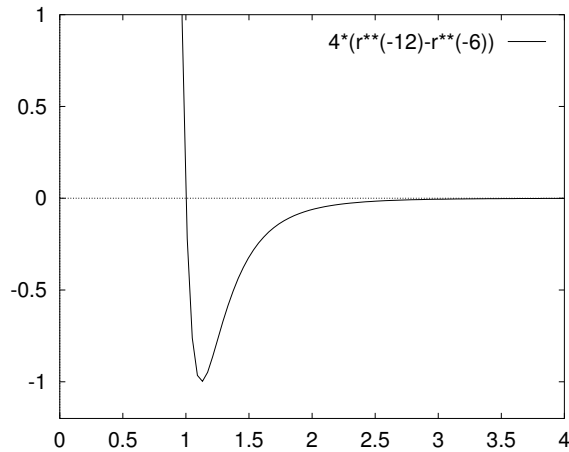


Figure 1.7: Lennard-Jones potential with $\epsilon = 1$ (well depth) and $\sigma = 1$ (contact distance where $U(r) = 0$). Many real fluids consisting of atoms or almost isotropic molecules are well described by the LJ potential.

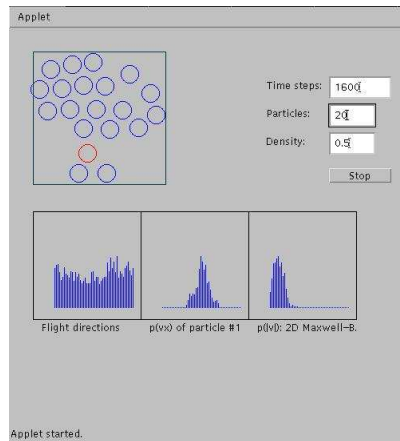


Figure 1.8: Applet LJONES

- **HARMONIC CRYSTAL**

The basic model for a solid is a regular configuration of atoms or ions that are bound to their nearest neighbors by a suitably modelled pair potential. Whatever the functional form of this potential, it may be approximated, for small excursions of any one particle from its equilibrium position, by a harmonic potential. Let l denote the equilibrium (minimal potential) distance between two neighbouring particles; the equilibrium position of atom j in a one-dimensional lattice is then given by $X_j = jl$. Defining the *displacements* of the atoms from their lattice points by x_j we have for the energy of the lattice

$$E(\vec{x}, \dot{\vec{x}}) = \frac{f}{2} \sum_{j=2}^N (x_j - x_{j-1})^2 + \frac{m}{2} \sum_{j=1}^N \dot{x}_j^2 \quad (1.23)$$

where f is a force constant and m is the atomic mass.

The generalization of 1.23 to 2 and 3 dimensions is straightforward.

The further treatment of this model is simplified by the approximate assumption that each particle is moving independently from all others in its own oscillator potential: $E(\vec{x}, \dot{\vec{x}}) \approx f \sum_{j=1}^N x_j^2 + (m/2) \sum_{j=1}^N \dot{x}_j^2$. In going to 2 and 3 dimensions one introduces the further simplifying assumption that this private oscillator potential is isotropically “smeared out”: $E_{pot}(\{x, y, z\}) \approx f \sum_{j=1}^N (x_j^2 + y_j^2 + z_j^2)$. The model thus defined is known as the “Einstein model” of solids.

- **MODELS FOR COMPLEX MOLECULES**

Most real substances consist of more complex units than isotropic atoms or Lennard-Jones type particles. There may be several interaction centers per particle, containing electrical charges, dipoles or multipoles, and the units may be joined by rigid or flexible bonds. Some of these models are still amenable to a theoretical treatment, but more often than not the methods of numerical simulation – Monte Carlo or molecular dynamics – must be invoked.

- **SPIN LATTICES**

Magnetically or electrically polarizable solids are often described by models in which “spins” with the discrete permitted values $\sigma_i = \pm 1$ are located at the vertices of a lattice.

If the spins have no mutual interaction the discrete states of such a system are easy to enumerate. This is why such models are often used to demonstrate of statistical-mechanical – or actually, combinatorial – relations (Reif: Berkeley Lectures). The energy of the system is then defined by $E \equiv -H \sum_i \sigma_i$, where H is an external field.

Of more physical significance are those models in which the spins interact with each other, For example, the parallel alignment of two neighboring spins may be energetically favored over the antiparallel configuration (**Ising** model). Monte Carlo simulation experiments on systems of this type have contributed much to our understanding of the properties of ferromagnetic and -electric substances.

1.3 Fundamentals of Statistics

By *statistics* we denote the investigation of regularities in apparently non-deterministic processes. An important basic quantity in this context is the “relative frequency” of an “event”. Let us consider a repeatable experiment – say, the throwing of a die – which in each instance leads to one of several possible results e – say, $e \equiv \text{number of points} = 4$. Now repeat this experiment n times under equal conditions and register the number of cases in which the specific result e occurs; call this number $f_n(e)$. The relative frequency of e is then defined as $r(e) \equiv f_n(e)/n$.

Following R. von Mises we denote as the “probability ” of an event e the *expected value* of the relative frequency in the limit of infinitely many experiments:

$$\mathcal{P}(e) \equiv \lim_{n \rightarrow \infty} \frac{f_n(e)}{n} \tag{1.24}$$

EXAMPLE: Game die; 100-1000 trials; $e \equiv \{\text{no. of points} = 6\}$, or $e \equiv \{\text{no. of points} \leq 3\}$.

Now, this definition does not seem very helpful. It implies that we have already done some experiments to determine the relative frequency, and it tells us no more than that we should expect more or less the same relative frequencies when we go on repeating the trials. What we want, however, is a recipe for the **prediction** of $\mathcal{P}(e)$.

To obtain such a recipe we have to reduce the event e to so-called “elementary events” ϵ that obey the postulate of **equal a priori probability**. Since the probability of any particular one among K possible elementary events is just $\mathcal{P}(\epsilon_i) = 1/K$, we may then derive the probability of a compound event by applying the rules

$$\mathcal{P}(e = \epsilon_i \text{ or } \epsilon_j) = \mathcal{P}(\epsilon_i) + \mathcal{P}(\epsilon_j) \quad (1.25)$$

$$\mathcal{P}(e = \epsilon_i \text{ and } \epsilon_j) = \mathcal{P}(\epsilon_i) \cdot \mathcal{P}(\epsilon_j) \quad (1.26)$$

Thus the predictive calculation of probabilities reduces to the **counting** of the possible elementary events that make up the event in question.

EXAMPLE: The result $\epsilon_6 \equiv \{\text{no. of points} = 6\}$ is one among 6 mutually exclusive elementary events with equal a priori probabilities ($= 1/6$). The compound event $e \equiv \{\text{no. of points} \leq 3\}$ consists of the elementary events $\epsilon_i = \{1, 2 \text{ or } 3\}$; its probability is thus $\mathcal{P}(1 \vee 2 \vee 3) = 1/6 + 1/6 + 1/6 = 1/2$.

How might this apply to statistical mechanics? – Let us assume that we have N equivalent mechanical systems with possible states s_1, s_2, \dots, s_K . A relevant question is then: what is the probability of a situation in which

$$e \equiv \{k_1 \text{ systems are in state } s_1, k_2 \text{ systems in state } s_2, \text{ etc.}\} \quad (1.27)$$

EXAMPLE: $N = 60$ dice are thrown (or *one* die 60 times!). What is the probability that 10 dice each have numbers of points 1, 2, ... 6? What, in contrast, is the probability that *all* dice show a “one”?

The same example, but with more obviously physical content:

Let $N = 60$ gas atoms be contained in a volume V , which we imagine to be divided into 6 equal partial volumes. What is the probability that at any given time we find $k_i = 10$ particles in each subvolume? And how probable is the particle distribution $(60, 0, 0, 0, 0, 0)$? (Answer: see below under the heading “multinomial distribution”.)

We can generally assume that both the number N of systems and the number K of accessible states are very large – in the so-called “thermodynamic limit” they are actually taken to approach infinity. This gives rise to certain mathematical simplifications.

Before advancing into the field of physical applications we will review the fundamental concepts and truths of statistics and probability theory, focussing on events that take place in number space, either \mathcal{R} (real numbers) or \mathcal{N} (natural numbers).

DISTRIBUTION FUNCTION

Let x be a real random variate in the region (a, b) . The distribution function

$$P(x_0) \equiv \mathcal{P}\{x < x_0\} \quad (1.28)$$

is defined as the probability that some x is smaller than the given value x_0 . The function $P(x)$ is monotonically increasing and has $P(a) = 0$ and $P(b) = 1$. The distribution function is dimensionless: $[P(x)] = 1$.

The most simple example is the equidistribution for which

$$P(x_0) = \frac{x_0 - a}{b - a} \quad (1.29)$$

Another important example, with $a = -\infty$, $b = \infty$, is the normal distribution

$$P(x_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_0} dx e^{-x^2/2} \quad (1.30)$$

and its generalization, the Gaussian distribution

$$P(x_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x_0} dx e^{-(x-\langle x \rangle)^2/2\sigma^2} \quad (1.31)$$

where the parameters $\langle x \rangle$ and σ^2 define an ensemble of functions.

DISTRIBUTION DENSITY

The distribution or probability density $p(x)$ is defined by

$$p(x_0) dx \equiv \mathcal{P}\{x \in [x_0, x_0 + dx]\} \equiv dP(x_0) \quad (1.32)$$

In other words, $p(x)$ is just the differential quotient of the distribution function:

$$p(x) = \frac{dP(x)}{dx}, \quad \text{i.e. } P(x_0) = \int_a^{x_0} p(x) dx \quad (1.33)$$

$p(x)$ has a dimension; it is the reciprocal of the dimension of its argument x :

$$[p(x)] = \frac{1}{[x]} \quad (1.34)$$

For the equidistribution we have

$$p(x) = 1/(b - a), \quad (1.35)$$

and for the normal distribution

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (1.36)$$

If x is limited to discrete values x_α with a step $\Delta x_\alpha \equiv x_{\alpha+1} - x_\alpha$ one often writes

$$p_\alpha \equiv p(x_\alpha) \Delta x_\alpha \quad (1.37)$$

for the probability of the event $x = x_\alpha$. This p_α is by definition dimensionless, although it is related to the distribution density $p(x)$ for continuous arguments. The definition 1.37 includes the special case that x is restricted to integer values k ; in that case $\Delta x_\alpha = 1$.

MOMENTS OF A DENSITY

By this we denote the quantities

$$\langle x^n \rangle \equiv \int_a^b x^n p(x) dx \quad \text{or, in the discrete case,} \quad \equiv \sum_\alpha p_\alpha x_\alpha^n \quad (1.38)$$

The **first** moment $\langle x \rangle$ is also called the **expectation value** or **mean value** of the distribution density $p(x)$, and the **second** moment $\langle x^2 \rangle$ is related to the **variance** and the **standard deviation**: variance $\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$ (standard deviation $\sigma =$ square root of variance).

EXAMPLES:

- .) For an equidistribution $\epsilon[0, 1]$ we have $\langle x \rangle = 1/2$, $\langle x^2 \rangle = 1/3$ and $\sigma^2 = 1/12$.
- .) For the normal distribution we find $\langle x \rangle = 0$ and $\langle x^2 \rangle = \sigma^2 = 1$.

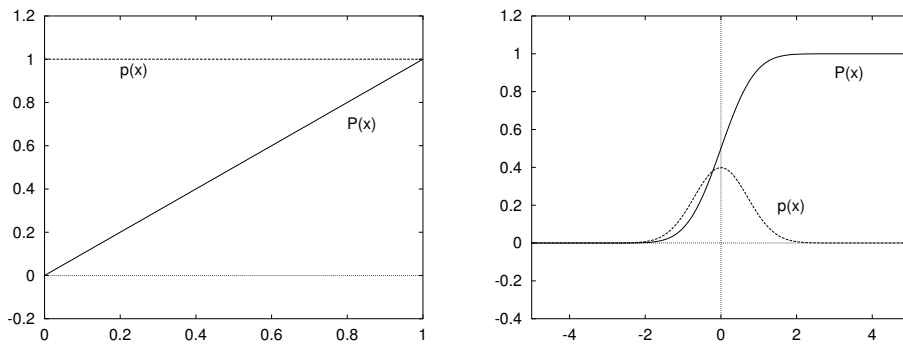


Figure 1.9: Equidistribution and normal distribution functions P and densities p

SOME IMPORTANT DISTRIBUTIONS

- **Equidistribution:** Its great significance stems from the fact that this distribution is central both to statistical mechanics and to practical numerics. In the theory of statistical-mechanical systems, one of the fundamental assumptions is that all states of a system that have the same energy are equally probable (axiom of equal *a priori* probability). And in numerical computing the generation of homogeneously distributed pseudo-random numbers is relatively easy; to obtain differently distributed random variates one usually “processes” such primary equidistributed numbers.

- **Gauss distribution:** This distribution pops up everywhere in the quantifying sciences. The reason for its ubiquity is the “central value theorem”: *Every* random variate that can be expressed as a *sum of arbitrarily distributed* random variates will in the limit of many summation terms be *Gauss distributed*. For example, when we have a complex measuring procedure in which a number of individual errors (or uncertainties) add up to a total error, then this error will be nearly Gauss distributed, regardless of how the individual contributions may be distributed. In addition, several other physically relevant distributions, such as the binomial and multinomial densities (see below), approach the Gauss distribution under certain – quite common – circumstances.

- **Binomial distribution:** This **discrete** distribution describes the probability that in n independent trials an event that has a single trial probability p will occur exactly k times:

$$\begin{aligned}
 p_k^n &\equiv \mathcal{P}\{k \text{ times } e \text{ in } n \text{ trials}\} \\
 &= \binom{n}{k} p^k (1-p)^{n-k}
 \end{aligned}
 \tag{1.39}$$

For the first two moments of the binomial distribution we have $\langle k \rangle = np$ (not necessarily integer) and $\sigma^2 = np(1-p)$ (i.e. $\langle k^2 \rangle - \langle k \rangle^2$).

APPLICATION: Fluctuation processes in statistical systems are often described in terms of the binomial distribution. For example, consider a particle freely roaming a volume V . The probability to find it at some given time in a certain partial volume

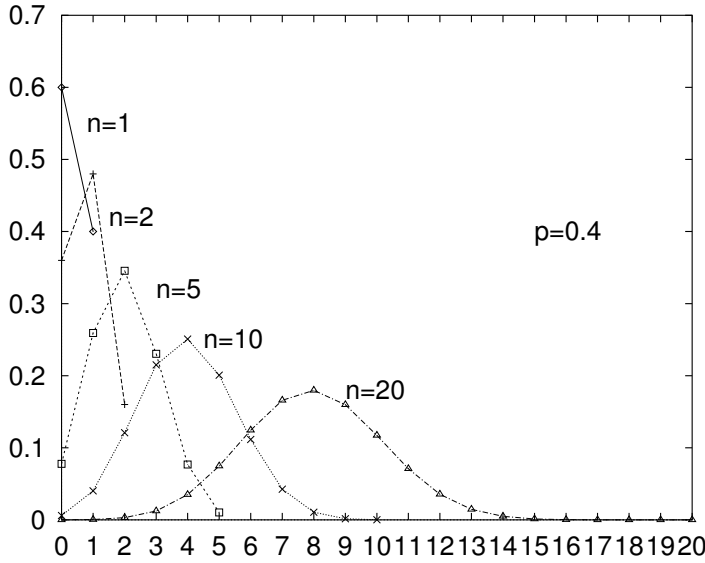


Figure 1.10: Binomial distribution density p_k^n

V_1 is $p(V_1) = V_1/V$. Considering now N independent particles in V , the probability of finding just N_1 of them in V_1 is given by

$$p_{N_1}^N = \binom{N}{N_1} (V_1/V)^{N_1} (1 - V_1/V)^{N-N_1} \quad (1.40)$$

The average number of particles in V_1 and its standard deviation are

$$\langle N_1 \rangle = NV_1/V \quad \text{and} \quad \sigma \equiv \sqrt{(\Delta N_1)^2} = \sqrt{N(V_1/V)(1 - V_1/V)}. \quad (1.41)$$

Note that for $Np \approx 1$ we have for the variance $\sigma^2 \approx \sigma \approx \langle N_1 \rangle \approx 1$, meaning that the population fluctuations in V_1 are then of the same order of magnitude (namely, 1) as the mean number of particles itself.

For large n such that $np \gg 1$ the **binomial distribution** approaches a **Gauss distribution** with mean np and variance $\sigma^2 = npq$ (theorem of Moivre-Laplace):

$$p_k^n \Rightarrow p_G(k) = \frac{1}{\sqrt{2\pi npq}} \exp[-(k - np)^2/2npq] \quad (1.42)$$

with $q \equiv 1 - p$.

If $n \rightarrow \infty$ and $p \rightarrow 0$ such that their product $np \equiv \lambda$ remains finite, the density 1.39 approaches

$$p_n(k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (1.43)$$

which goes by the name of **Poisson distribution**.

An important element in the success story of statistical mechanics is the fact that with increasing n the **sharpness** of the distribution 1.39 or 1.42 becomes very large. The **relative width** of the maximum, i. e. $\sigma/\langle k \rangle$, decreases as $1/\sqrt{n}$. For $n = 10^4$ the width of the peak is no more than 1% of $\langle k \rangle$, and for “molar” orders of particle numbers $n \approx 10^{24}$ the relative width $\sigma/\langle k \rangle$ is already $\approx 10^{-12}$. Thus the density approaches a “delta distribution”. This, however, renders the calculation of averages particularly simple:

$$\langle f(k) \rangle = \sum_k p_k^n f(k) \rightarrow f(\langle k \rangle) \quad (1.44)$$

or

$$\langle f(k) \rangle \approx \int dk \delta(k - \langle k \rangle) f(k) = f(\langle k \rangle) \quad (1.45)$$

• **Multinomial distribution:** This is a generalization of the binomial distribution to more than 2 possible results of a single trial. Let e_1, e_2, \dots, e_K be the (mutually exclusive) possible results of an experiment; their probabilities in a single trial are p_1, p_2, \dots, p_K , with $\sum_i^K p_i = 1$. Now do the experiment n times; then

$$p_n(k_1, k_2, \dots, k_K) = \frac{n!}{k_1! k_2! \dots k_K!} p_1^{k_1} p_2^{k_2} \dots p_K^{k_K} \quad (\text{with } k_1 + k_2 + \dots + k_K = n) \quad (1.46)$$

is the probability to have the event e_1 just k_1 times, e_2 accordingly k_2 times, etc.

We get an idea of the significance of this distribution in statistical physics if we interpret the K possible events as “states” that may be taken on by the n particles of a system (or, in another context, by the n systems in an ensemble of many-particle systems). The above formula then tells us the probability to find k_1 among the n particles in state e_1 , etc.

EXAMPLE: A die is cast 60 times. The probability to find each number of points just 10 times is

$$p_{60}(10, 10, 10, 10, 10, 10) = \frac{60!}{(10!)^6} \left(\frac{1}{6}\right)^{60} = 7.457 \cdot 10^{-5}. \quad (1.47)$$

To compare, the probabilities of two other cases: $p_{60}(10, 10, 10, 10, 9, 11) = 6.778 \cdot 10^{-5}$, $p_{60}(15, 15, 15, 5, 5, 5) = 4.406 \cdot 10^{-8}$. Finally, for the quite improbable case $(60, 0, 0, 0, 0, 0)$ we have $p_{60} = 2.046 \cdot 10^{-47}$

Due to its large number of variables (K) we cannot give a graph of the multinomial distribution. However, it is easy to derive the following two important properties:

Approach to a multivariate Gauss distribution: just as the binomial distribution approaches, for large n , a Gauss distribution, the multinomial density approaches an appropriately generalized – “multivariate” – Gauss distribution.

Increasing sharpness: if n and $k_1 \dots k_K$ become very large (multiparticle systems; or ensembles of $n \rightarrow \infty$ elements), the function $p_n(k_1, k_2, \dots, k_K) \equiv p_n(\vec{k})$ has an **extremely sharp maximum** for a certain partitioning $\vec{k}^* \equiv \{k_1^*, k_2^*, \dots, k_K^*\}$, namely $\{k_i^* = np_i; i = 1, \dots, K\}$. This particular partitioning of the particles to the various possible states is then “almost always” realized, and all other allotments (or distributions) occur very rarely and may safely be neglected.

This is the basis of the **method of the most probable distribution** which is used with great success in several areas of statistical physics.[2.2]

STIRLING’S FORMULA

For large values of m the evaluation of the factorial $m!$ is difficult. A handy approximation is **Stirling’s formula**

$$m! \approx \sqrt{2\pi m} (m/e)^m \quad (1.48)$$

EXAMPLE: $m = 69$ (Near most pocket calculators’ limit): $69! = 1.7112 \cdot 10^{98}$;
 $\sqrt{2\pi \cdot 69} (69/e)^{69} = 1.7092 \cdot 10^{98}$.

The same name **Stirling’s formula** is often used for the **logarithm** of the factorial:

$$\ln m! \approx m(\ln m - 1) + \ln \sqrt{2\pi m} \approx m(\ln m - 1) \quad (1.49)$$

(The term $\ln \sqrt{2\pi m}$ may usually be neglected in comparison to $m(\ln m - 1)$.)

EXAMPLE 1: $\ln 69! = 226.1905$; $69(\ln 69 - 1) + \ln \sqrt{2\pi \cdot 69} = 223.1533 + 3.0360 = 226.1893$.

EXAMPLE 2: The die is cast again, but now there are 120 trials. When asked to produce 120! most pocket calculators will cancel their cooperation. So we apply Stirling's approximation:

The probability of throwing each number of points just 20 times is

$$p_{120}(20, \dots, 20) = \frac{120!}{(20!)^6} \left(\frac{1}{6}\right)^{120} \approx \sqrt{240\pi} \left[\left(\frac{120}{6e}\right)^{20} \frac{1}{20!} \right]^6 = 1.350 \cdot 10^{-5}, \quad (1.50)$$

and the probability of the partitioning (20, 20, 20, 20, 19, 21) is $1.285 \cdot 10^{-5}$.

STATISTICAL (IN)DEPENDENCE

Two random variates x_1, x_2 are statistically mutually independent (uncorrelated) if the distribution density of the **compound probability** (i.e. the probability for the joint occurrence of x_1 and x_2) equals the product of the individual densities:

$$p(x_1, x_2) = p(x_1)p(x_2). \quad (1.51)$$

EXAMPLE: In a fluid or gas the distribution density for a single component of the particle velocity is given by (Maxwell-Boltzmann)

$$p(v_\alpha) = \sqrt{\frac{m}{2\pi kT}} \exp\left\{-\frac{mv_\alpha^2}{2kT}\right\}, \quad \alpha = x, y, z \quad (1.52)$$

The degrees of freedom $\alpha = x, y, z$ are statistically independent; therefore the compound probability is given by

$$p(\vec{v}) \equiv p(v_x, v_y, v_z) = p(v_x)p(v_y)p(v_z) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left\{-\frac{m}{2kT}(v_x^2 + v_y^2 + v_z^2)\right\} \quad (1.53)$$

By **conditional distribution density** we denote the quantity

$$p(x_2|x_1) \equiv \frac{p(x_1, x_2)}{p(x_1)} \quad (1.54)$$

(For uncorrelated x_1, x_2 we have $p(x_2|x_1) = p(x_2)$).

The density of a **marginal distribution** describes the density of one of the variables *regardless of the specific value of the other one*, meaning that we integrate the joint density over all possible values of the second variable:

$$p(x_2) \equiv \int_{a_1}^{b_1} p(x_1, x_2) dx_1 \quad (1.55)$$

TRANSFORMATION OF DISTRIBUTION DENSITIES

From 1.32 we can immediately conclude how the density $p(x)$ will transform if we substitute the variable x . Given some bijective mapping $y = f(x)$; $x = f^{-1}(y)$ the conservation of probability requires

$$|dP(y)| = |dP(x)| \quad (1.56)$$

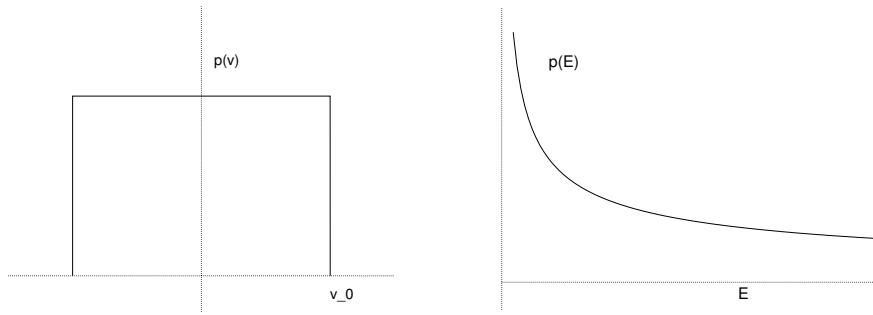


Figure 1.11: Transformation of the distribution density (see Example 1)

(The absolute value appears because we have not required $f(x)$ to be an increasing function.) This leads to

$$|p(y) dy| = |p(x) dx| \quad (1.57)$$

or

$$p(y) = p(x) \left| \frac{dx}{dy} \right| = p[f^{-1}(y)] \left| \frac{df^{-1}(y)}{dy} \right| \quad (1.58)$$

Incidentally, this relation is true for any kind of density, such as mass or spectral densities, and not only for distribution densities.

EXAMPLE 1: A particle of mass m moving in one dimension is assumed to have any velocity in the range $\pm v_0$ with equal probability; so we have $p(v) = 1/2v_0$. The distribution density for the kinetic energy is then given by (see Figure 1.11)

$$p(E) = 2p(v) \left| \frac{dv}{dE} \right| = \frac{1}{2\sqrt{E_0}} \frac{1}{\sqrt{E}} \quad (1.59)$$

in the limits $0..E_0$, where $E_0 = mv_0^2/2$. (The factor 2 in front of $p(v)$ comes from the ambiguity of the mapping $v \leftrightarrow E$).

EXAMPLE 2: An object is found with equal probability at any point along a circular periphery; so we have $p(\phi) = 1/2\pi$ for $\phi \in [0, 2\pi]$. Introducing cartesian coordinates $x = r \cos \phi$, $y = r \sin \phi$ we find for the distribution density of the coordinate x , with $x \in [\pm r]$, that

$$p(x) = p(\phi) \left| \frac{d\phi}{dx} \right| = \frac{1}{\pi} \frac{1}{\sqrt{r^2 - x^2}} \quad (1.60)$$

(see Figure 1.12)

Problems equivalent to these examples:

- A homogeneously blackened glass cylinder – or a semitransparent drinking straw – held sideways against a light source: absorption as a function of the distance from the axis?
- Distribution of the x -velocity of a particle that can move randomly in two dimension, keeping its kinetic energy constant.

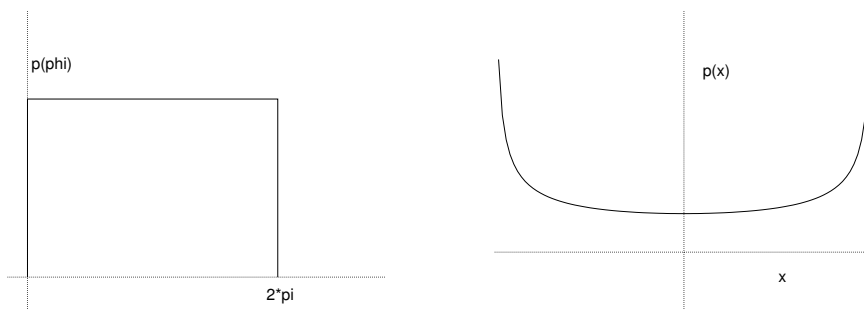


Figure 1.12: Transformation of the distribution density (see Example 2)

⇒ **Simulation 1.6: Stadium Billiard.** Distribution of the velocity component v_x . [Code: Stadium]

c) Distribution of the velocity of any of two particles arbitrarily moving in one dimension, keeping only the sum of their kinetic energies constant.

For the joint probability density of **several** variables the transformation formula is a direct generalization of 1.58, viz.

$$p(\vec{y}) = p(\vec{x}) \left| \frac{d\vec{x}}{d\vec{y}} \right| \quad (1.61)$$

Here we write $|d\vec{x}/d\vec{y}|$ for the **functional determinant** (or Jacobian) of the mapping $\vec{x} = \vec{x}(\vec{y})$,

$$|d(x_1, x_2, \dots)/d(y_1, y_2, \dots)| = \begin{vmatrix} dx_1/dy_1 & dx_1/dy_2 & \dots \\ dx_2/dy_1 & dx_2/dy_2 & \dots \\ \vdots & & \ddots \end{vmatrix} \quad (1.62)$$

EXAMPLE 3: Again, let $\vec{v} \equiv \{v_x, v_y, v_z\}$, and $p(\vec{v})$ as in equ. 1.53. Now we write $\vec{w} \equiv \{v, \phi, \theta\}$, with

$$v_x = v \sin \theta \cos \phi, \quad v_y = v \sin \theta \sin \phi, \quad v_z = v \cos \theta. \quad (1.63)$$

The Jacobian of the mapping $\vec{v} = \vec{v}(\vec{w})$ is

$$|d(v_x, v_y, v_z)/d(v, \phi, \theta)| = \begin{vmatrix} \sin \theta \cos \phi & -v \sin \theta \sin \phi & v \cos \theta \cos \phi \\ \sin \theta \sin \phi & v \sin \theta \cos \phi & v \cos \theta \sin \phi \\ \cos \theta & 0 & -v \sin \theta \end{vmatrix} = -v^2 \sin \theta \quad (1.64)$$

Therefore we have for the density of the **modulus** of the particle velocity

$$p(v) \equiv \int_0^{2\pi} d\phi \int_{-\pi}^{\pi} d\theta v^2 \sin \theta \left(\frac{m}{2\pi kT} \right)^{3/2} \exp\left\{-\frac{m}{2kT}v^2\right\} \quad (1.65)$$

$$= 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 \exp\left\{-\frac{m}{2kT}v^2\right\} \quad (1.66)$$

1.4 Problems for Chapter 1

EXERCISES:

1.1 Entropy and irreversibility: (a) Two heat reservoirs are at temperatures T_1 and $T_2 > T_1$. They are connected by a metal rod that conducts the heat Q per unit time from 2 to 1. Do an entropy balance to show that this process is irreversible.

(b) Using the specific values $T_1 = 300K$, $T_2 = 320K$ and $Q = 10J$, calculate the entropy increase per second.

1.2 Reversible, isothermal expansion: Compute the entropy balance for the experiment described (or shown) in the lecture, using estimated values of the necessary experimental parameters.

Discuss the role of the ideal gas assumption: is it necessary / unnecessary / convenient ?

1.3 Maxwell square:

Draw up Maxwell's square and use it to complete the following equations:

- (a) $G = A \pm \dots$; (b) $A = E \pm \dots$;
- (c) $P = \dots = \dots$ (in terms of derivatives of A and E);
- (d) $dG = V dP - \dots$

1.4 Random directions in 2D: Using your favourite random number generator, sample a number $K \approx 50 - 100$ of angles ϕ_i ($i = 1, \dots, K$) equidistributed in $[0, 2\pi]$. Interpret $\cos \phi_i$ and $\sin \phi_i$ as the $v_{x,y}$ -components of a randomly oriented velocity vector \vec{v}_i with $|\vec{v}_i| = 1$.

(a) Draw a histogram of the empirical probability (i.e. event frequency) of $p(v_x)$. Compare the shape of your $p(v_x)$ to the bottom right histogram in Applet *Stadium* (see 1.2).

(b) Normalize the histogram such that a sum over all bins equals one. What is the value of $p(0.5)$?

1.5 Binomial distribution: Suggest an experiment with given p and n . Perform the experiment N times and draw an empirical frequency diagram; compare with 1.39.

1.6 Density fluctuations in air:

(a) Calculate the mean number of molecules (not discerning between N_2 and O_2) that are to be found at a pressure of $10 Pa$ in a cube with a side length of the wave length of light ($\approx 6 \cdot 10^{-7} m$). What is the standard deviation of the particle number, both absolutely and relative to the mean particle number? (Air is to be treated as an ideal gas at normal temperature $T_0 = 273.15 K$.)

(b) Compute the value of the probability density of the event $N(\Delta V) = k \approx \langle N \rangle$, i. e. the probability of finding an integer number k next to the mean number of particles in the sample volume? (Hint: Don't attempt to evaluate factorials of large numbers, such as appear in the binomial distribution p_k^n ; rather, use that distribution which resembles p_k^n when n becomes large.) What is the probability of finding only 95 percent of the mean particle number in the sample volume?

1.7 Multinomial distribution: A volume V is divided into $m = 5$ equal-sized cells. The $N = 1000$ particles of an ideal gas may be allotted randomly to the cells.

a) What is the probability of finding in a snapshot of the system the partitioning $\{N_1, N_2, \dots, N_5\}$? Explain the formula.

b) Demonstrate *numerically* that the partitioning with the greatest probability is given by $N_i = N/m = 200$. For example, compare the situations $(201, 199, \dots)$, $(202, 199, 199, \dots)$, $(202, 198, \dots)$, $(205, 195, \dots)$, and $(204, 199, 199, 199, 199)$ to the most probable one.

c) (1 bonus point) Prove *analytically* that $N_i = 200$ is the most probable distribution. Hint: minimize the function $f(\vec{k}) \equiv \log p_n(\vec{k})$ under the condition $\sum_i k_i = n$.

1.8 Transformation of a distribution density: Repeat the calculation of Example 3 for the

two-dimensional case, i.e. $\vec{v} \equiv \{v_x, v_y\}$ and $\vec{w} \equiv \{v, \phi\}$. Draw the distribution density $p_{2D}(v)$.

TEST YOUR UNDERSTANDING OF CHAPTER 1:

1.1 Thermodynamic concepts:

- What does the entropy balance tell us about the reversibility/irreversibility of a process? Demonstrate, using a specific example.
- Describe the process of thermal interaction between two bodies. When will the energy flow stop?
- Which thermodynamic potential is suited for the description of isothermal-isochoric systems?

1.2 Model systems:

- Describe 2-3 model systems of statistical mechanics.
- What quantities are needed to completely describe the momentary state of a classical ideal gas of N particles?
- What quantities are needed for a complete specification of the state of a quantum ideal gas?

1.3 Statistical concepts:

- Explain the concepts “distribution function” and “distribution density”; give two examples.
- What are the moments of a distribution? Give a physically relevant example.

1.4 Equal a priori probability: Explain the concept and its significance for statistical mechanics.

Chapter 2

Elements of Kinetic Theory



Figure 2.1: Ludwig Boltzmann

In a dilute gas the molecules are in free, linear flight most of the time; just occasionally their flight will be interrupted by a collision with a single other particle. The result of such an event, which in general may be treated as a classical elastic collision, is a change in the speeds and directions of both partners. So it should be possible to derive, from a **statistical** treatment of such binary collisions, predictions on the properties of gases. This idea was the starting point from which Ludwig Boltzmann proceeded to develop his “Kinetic Theory of Gases”.

First he defined a **distribution density** f such that $f(\vec{r}, \vec{v}; t) d\vec{r} d\vec{v}$ denotes the number of particles which at time t situated at \vec{r} and have a velocity \vec{v} . The 6-dimensional space of the variables $\{\vec{r}, \vec{v}\}$ has later been called “ μ space”.¹

The evolution in time of the function $f(\vec{r}, \vec{v}; t)$ is then described by **Boltzmann’s Transport Equation**. We will shortly sketch the derivation and shape of this important formula.

A particularly important result of Boltzmann’s equation is its **stationary solution**, i.e. the function that solves the equation in the long-time limit $t \rightarrow \infty$. It turns out that this **equilibrium**

¹The term is probably due to Paul Ehrenfest who in his classical book on Statistical Mechanics refers to μ -space, or “molecular” space as opposed to Γ , or “gas” space (see Chapter 3 below).

distribution in μ space may be derived without explicitly solving the transport equation itself. The strategy used to do so, called the **method of the most probable distribution**, will be encountered again in other parts of Statistical Mechanics.

2.1 Boltzmann's Transport Equation

With his "Kinetic Theory of Gases" Boltzmann undertook to explain the properties of dilute gases by analysing the elementary collision processes between pairs of molecules.

The evolution of the distribution density in μ space, $f(\vec{r}, \vec{v}; t)$, is described by **Boltzmann's transport equation**. A thorough treatment of this beautiful achievement is beyond the scope of our discussion. But we may sketch the basic ideas used in its derivation.

- If there were no collisions at all, the swarm of particles in μ space would flow according to

$$f\left(\vec{r} + \vec{v} dt, \vec{v} + \frac{\vec{K}}{m} dt; t + dt\right) = f(\vec{r}, \vec{v}; t) \quad (2.1)$$

where \vec{K} denotes an eventual external force acting on particles at point (\vec{r}, \vec{v}) . The time derivative of f is therefore, in the collisionless case,

$$\left(\frac{\partial}{\partial t} + \vec{v} \cdot \nabla_{\vec{r}} + \frac{\vec{K}}{m} \cdot \nabla_{\vec{v}}\right) f(\vec{r}, \vec{v}; t) = 0 \quad (2.2)$$

where

$$\vec{v} \cdot \nabla_{\vec{r}} f \equiv v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z} \quad (2.3)$$

and

$$\frac{\vec{K}}{m} \cdot \nabla_{\vec{v}} f \equiv \frac{1}{m} \left(K_x \frac{\partial f}{\partial v_x} + K_y \frac{\partial f}{\partial v_y} + K_z \frac{\partial f}{\partial v_z} \right) \quad (2.4)$$

To gather the meaning of equation 2.2 for free flow, consider the collisionless, free flow of gas particles through a thin pipe: there is no force (i. e. no change of velocities), and μ -space has only two dimensions, x and v_x (see Figure 2.2).

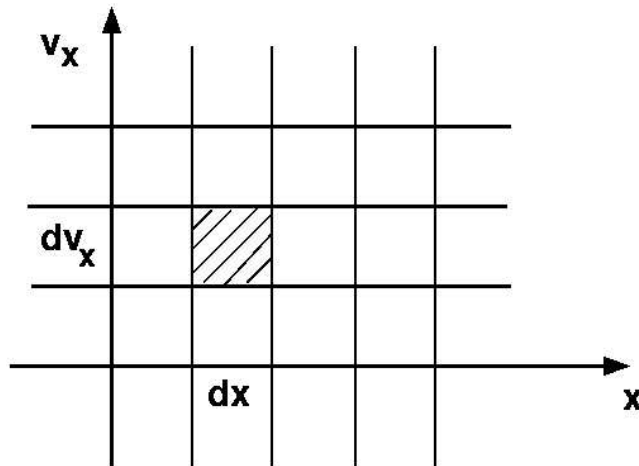


Figure 2.2: A very simple μ -space

At time t a differential “volume element” at (x, v_x) contains, on the average, $f(x, v_x)dx dv_x$ particles. The temporal change of $f(x, v_x)$ is then given by

$$\frac{\partial f(x, v_x)}{\partial t} = -v_x \frac{\partial f(x, v_x)}{\partial x} \quad (2.5)$$

To see this, count the particles entering during the time span dt from the left (assuming $v_x > 0$, $n_{in} = f(x - v_x dt, v_x)dx dv_x$ and those leaving towards the right, $n_{out} = f(x, v_x)dx dv_x$. The local change per unit time is then

$$\frac{\partial f(x, v_x)}{\partial t} = \frac{n_{in} - n_{out}}{dt dx dv_x} \quad (2.6)$$

$$= \frac{f(x - v_x dt, v_x) - f(x, v_x)}{dt} \quad (2.7)$$

$$= \frac{f(x, v_x) - (v_x dt) (\partial f / \partial x) - f(x, v_x)}{dt} \quad (2.8)$$

$$= \frac{f(x, v_x) - (v_x dt) (\partial f / \partial x) - f(x, v_x)}{dt} \quad (2.9)$$

$$= -v_x \frac{\partial f(x, v_x)}{\partial x} \quad (2.10)$$

$$= -v_x \frac{\partial f(x, v_x)}{\partial x} \quad (2.11)$$

The relation $\frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} = 0$ is then easily generalized to the case of a non-vanishing force,

$$\frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} + \frac{K_x}{m} \frac{\partial f}{\partial v_x} = 0 \quad (2.12)$$

(this would engender an additional vertical flow in the figure), and to six instead of two dimensions (see equ. 2.2).

All this is for collisionless flow only.

- In order to account for collisions a term $(\partial f / \partial t)_{coll} dt$ is added on the right hand side:

$$\left(\frac{\partial}{\partial t} + \vec{v} \cdot \nabla_{\vec{r}} + \frac{\vec{K}}{m} \cdot \nabla_{\vec{v}} \right) f(\vec{r}, \vec{v}; t) = \left(\frac{\partial f}{\partial t} \right)_{coll} \quad (2.13)$$

The essential step then is to find an explicit expression for $(\partial f / \partial t)_{coll}$. Boltzmann solved this problem under the simplifying assumptions that

- only binary collisions need be considered (dilute gas);
- the influence of container walls may be neglected;
- the influence of the external force \vec{K} (if any) on the rate of collisions is negligible;
- velocity and position of a molecule are uncorrelated (assumption of **molecular chaos**).

The effect of the binary collisions is expressed in terms of a “differential scattering cross section” $\sigma(\Omega)$ which describes the probability density for a certain change of velocities,

$$\{\vec{v}_1, \vec{v}_2\} \rightarrow \{\vec{v}'_1, \vec{v}'_2\}. \quad (2.14)$$

(Ω thus denotes the relative orientation of the vectors $(\vec{v}'_2 - \vec{v}'_1)$ and $(\vec{v}_2 - \vec{v}_1)$). The function $\sigma(\Omega)$ depends on the intermolecular potential and may be either calculated or measured.

Under all these assumptions, and by a linear expansion of the left hand side of equ. 2.1 with respect to time, the Boltzmann equation takes on the following form:

$$\left(\frac{\partial}{\partial t} + \vec{v}_1 \cdot \nabla_{\vec{r}} + \frac{\vec{K}}{m} \cdot \nabla_{\vec{v}_1} \right) f_1 = \int d\Omega \int d\vec{v}_2 \sigma(\Omega) |\vec{v}_1 - \vec{v}_2| \left(f'_1 f'_2 - f_1 f_2 \right) \quad (2.15)$$

where $f_1 \equiv f(\vec{r}, \vec{v}_1; t)$, $f'_1 \equiv f(\vec{r}, \vec{v}'_1; t)$ etc. This integrodifferential equation describes, under the given assumptions, the spatio-temporal behaviour of a dilute gas. Given some initial density $f(\vec{r}, \vec{v}; t = 0)$ in μ -space the solution function $f(\vec{r}, \vec{v}; t)$ tells us how this density changes over time. Since f has up to six arguments it is difficult to visualize; but there are certain moments of f which represent measurable averages such as the local particle density in 3D space, whose temporal change can thus be computed.

Chapman and Enskog developed a general procedure for the approximate solution of Boltzmann's equation. For certain simple model systems such as hard spheres their method produces predictions for $f(\vec{r}, \vec{v}; t)$ (or its moments) which may be tested in computer simulations. Another more modern approach to the numerical solution of the transport equation is the "Lattice Boltzmann" method in which the continuous variables \vec{r} and \vec{v} are restricted to a set of discrete values; the time change of these values is then described by a modified transport equation which lends itself to fast computation.

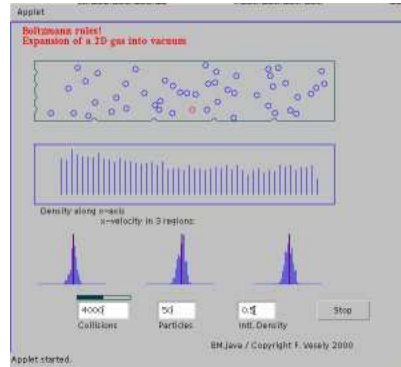
The initial distribution density $f(\vec{r}, \vec{v}; 0)$ may be of arbitrary shape. To consider a simple example, we may have all molecules assembled in the left half of a container – think of a removable shutter – and at time $t = 0$ make the rest of the volume accessible to the gas particles:

$$f(\vec{r}, \vec{v}, 0) = A\Theta(x_0 - x)f_0(\vec{v}) \quad (2.16)$$

where $f_0(\vec{v})$ is the (Maxwell-Boltzmann) distribution density of particle velocities, and $\Theta(x_0 - x)$ denotes the Heaviside function. The subsequent expansion of the gas into the entire accessible volume, and thus the approach to the stationary final state (= equilibrium state) in which the particles are evenly distributed over the volume may be seen in the solution $f(\vec{r}, \vec{v}; t)$ of Boltzmann's equation. Thus the greatest importance of this equation is its ability to describe also **non-equilibrium processes**.

Simulation: The power of Boltzmann's equation

- Irreversible Expansion of a hard disc gas
 - Presentation of the distribution densities in r-space and in v-space
- [Code: BM]



The **Equilibrium distribution** $f_0(\vec{r}, \vec{v})$ is that solution of Boltzmann's equation which is **stationary**, meaning that

$$\frac{\partial f(\vec{r}, \vec{v}; t)}{\partial t} = 0 \quad (2.17)$$

It is also the limiting distribution for long times, $t \rightarrow \infty$.

It may be shown that this equilibrium distribution is given by

$$f_0(\vec{r}, \vec{v}) = \rho(\vec{r}) \left[\frac{m}{2\pi kT(\vec{r})} \right]^{3/2} \exp\{-m [\vec{v} - \vec{v}_0(\vec{r})]^2 / 2kT(\vec{r})\} \quad (2.18)$$

where $\rho(\vec{r})$ and $T(\vec{r})$ are the local density and temperature, respectively.

If there are no external forces such as gravity or electrostatic interactions we have $\rho(\vec{r}) = \rho_0 = N/V$. In case the temperature is also independent of position, and if the gas as a whole is not moving ($\vec{v}_0 = 0$), then $f(\vec{r}, \vec{v}) = \rho_0 f_0(\vec{v})$, with

$$f_0(\vec{v}) = \left[\frac{m}{2\pi kT} \right]^{3/2} e^{-mv^2/2kT}, \quad (2.19)$$

This is the famous **Boltzmann distribution**; it may be derived also in different ways, without requiring the explicit solution of the transport equation. \implies See next section, equ.2.30.

2.2 The Maxwell-Boltzmann distribution

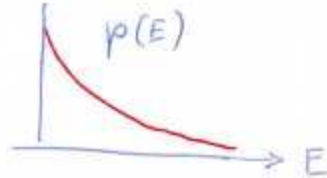
We want to apply statistical procedures to the swarm of points in Boltzmann's μ space. To do this we first divide that space in 6-dimensional cells of size $(\Delta x)^3(\Delta v)^3$, labelling them by j ($= 1, \dots, m$). There is a characteristic energy $E_j \equiv E(\vec{r}_j, \vec{v}_j)$ pertaining to each such cell. For instance, in the ideal gas case this energy is simply $E_j = mv_j^2/2$, where \vec{v}_j is the velocity of the particles in cell j .

Now we distribute the N particles over the m cells, such that n_j particles are allotted to cell no. j . In a closed system with total energy E the population numbers n_j must fulfill the condition $\sum_j n_j E_j = E$. The other condition is, of course, the conservation of the number of particles, $\sum_j n_j = N$. Apart from these two requirements the allotment of particles to cells is completely random.

We may understand this prescription as the rule of a game of fortune, and with the aid of a computer we may actually play that game!

Playing Boltzmann's game

[Code: LBRoulette]



Instead of playing the game we may calculate its outcome by probability theory. For good statistics we require that $N \gg m \gg 1$. A specific m -tuple of population numbers $\vec{n} \equiv \{n_j; j = 1, \dots, m\}$ will here be called a **partitioning**. (If you prefer to follow the literature, you may refer to it as a *distribution*.) Each partitioning may be performed in many different ways, since the labels of the particles may be permuted without changing the population numbers in the cells. This means that many specific **allotments** pertain to a single **partitioning**. Assuming that the allotments are elementary events of equal probability, we simply count the number of possible allotments to calculate the probability of the respective partitioning.

The number of possible permutations of particle labels for a given partition $\vec{n} \equiv \{n_j; j = 1, \dots, m\}$ is

$$P(\vec{n}) = \frac{N!}{n_1! n_2! \dots n_m!} \quad (2.20)$$

(In combinatorics this is called *permutations of m elements – i. e. cell numbers – with repetition*).

Since each allotment is equally probable, the most probable **partitioning** is the one allowing for the largest number of **allotments**. In many physically relevant cases the probability of that optimal partitioning is very much larger than that of any other, meaning that we can restrict the further discussion to this one partitioning. (See the previous discussion of the multinomial distribution.)

Thus we want to determine that specific partitioning \vec{n}^* which renders the expression 2.20 a maximum, given the additional constraints

$$\sum_{j=1}^m n_j - N = 0 \quad \text{and} \quad \sum_{j=1}^m n_j E_j - E = 0 \quad (2.21)$$

Since the logarithm is a monotonically increasing function we may scout for the maximum of $\ln P$ instead of P – this is mathematically much easier. The proven method for maximizing a function of many variables, allowing for additional constraints, is the **variational method with Lagrange multipliers** (see Problem 2.1). The variational equation

$$\delta \ln P - \delta \left(\alpha \sum_j n_j + \beta \sum_j E_j n_j \right) = 0 \quad (2.22)$$

with the undetermined multipliers α and β leads us, using the Stirling approximation for $\ln P$, to

$$-\ln n_j^* - \alpha - \beta E_j = 0 \quad (2.23)$$

Thus the optimal partitioning n_j^* is given by

$$n_j^* = e^{-\alpha - \beta E_j} ; \quad j = 1, \dots, m \quad (2.24)$$

Consequently

$$f(\vec{r}_j, \vec{v}_j) \propto n_j^* = A \exp\{-\beta E(\vec{r}_j, \vec{v}_j)\} \quad (2.25)$$

In particular, we find for a dilute gas, which in the absence of external forces will be homogeneous with respect to \vec{r} ,

$$f(\vec{v}_j) = B \exp\{-\beta(mv_j^2/2)\} \quad (2.26)$$

Using the normalizing condition $\int f(\vec{v}) d\vec{v} = 1$ or

$$B \int 4\pi v^2 e^{-\beta mv^2/2} dv = 1 \quad (2.27)$$

we find $B = (\beta m/2\pi)^{3/2}$ and therefore

$$f(\vec{v}) = \left[\frac{\beta m}{2\pi} \right]^{3/2} e^{-\beta mv^2/2} \quad (2.28)$$

Now we take a closer look at the quantity β which we introduced at first just for mathematical convenience. The mean kinetic energy of a particle is given by

$$\langle E \rangle \equiv \int d\vec{v} (mv^2/2) f(\vec{v}) = \frac{3}{2\beta} \quad (2.29)$$

But we will learn in Section 2.3 that the average kinetic energy of a molecule is related to the macroscopic observable quantity T (temperature) according to $\langle E \rangle = 3kT/2$; therefore we have $\beta \equiv 1/kT$. Thus we may write the distribution density of the velocity in the customary format

$$f(\vec{v}) = \left[\frac{m}{2\pi kT} \right]^{3/2} e^{-mv^2/2kT} \quad (2.30)$$

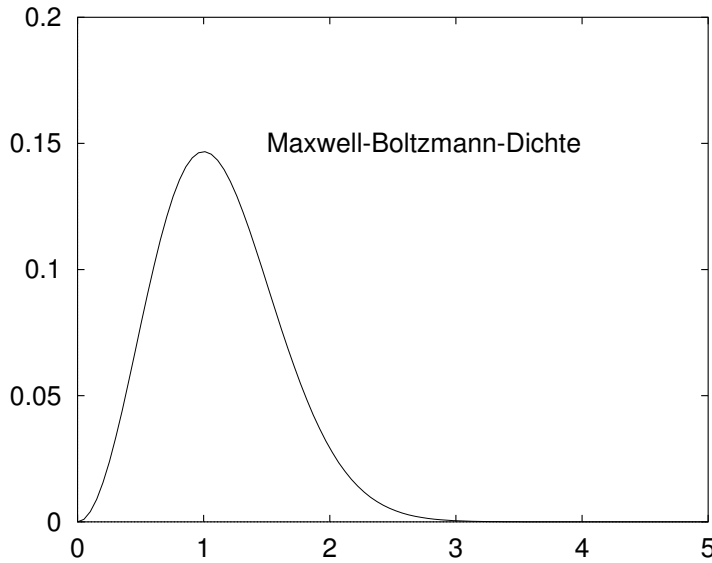


Figure 2.3: Maxwell-Boltzmann distribution

This density in velocity space is commonly called **Maxwell-Boltzmann distribution density**. The same name is also used for a slightly different object, namely the distribution density of the **modulus** of the particle velocity (the “speed”) which may easily be derived as (see equ. 1.66).

$$f(|\vec{v}|) = 4\pi v^2 f(\vec{v}) = 4\pi \left[\frac{m}{2\pi kT} \right]^{3/2} v^2 e^{-mv^2/2kT} \quad (2.31)$$

So we have determined the population numbers n_j^* of the cells in μ space by maximizing the number of possible allotments. It is possible to demonstrate that the partitioning we have found is not just the most probable but **by far** the most probable one. In other words, any noticeable deviation from this distribution of particle velocities is extremely improbable (see above: multinomial distribution.) This makes for the great practical importance of the MB distribution: it is simply **the** distribution of velocities in a many particle system which we may assume to hold, neglecting all other possible but improbable distributions.

As we can see from the figure, $f(|\vec{v}|)$ is a skewed distribution; its **maximum** is located at

$$\tilde{v} = \sqrt{\frac{2kT}{m}} \quad (2.32)$$

This **most probable** speed is not the same as the **mean** speed,

$$\langle v \rangle \equiv \int_0^\infty dv v f(|\vec{v}|) = \sqrt{\frac{8}{\pi}} \sqrt{\frac{kT}{m}} \quad (2.33)$$

or the **root mean squared velocity** or *r.m.s. velocity*),

$$\bar{v}_{rms} \equiv \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3kT}{m}} \quad (2.34)$$

EXAMPLE: The mass of an H_2 molecule is $m = 3.346 \cdot 10^{-27} kg$; at room temperature (appr. 300 K) we have $kT = 1.380 \cdot 10^{-23} \cdot 300 = 4.141 \cdot 10^{-21} J$; therefore the most probable speed of such a molecule under equilibrium conditions is $\tilde{v} = 1.573 \cdot 10^3 m/s$.

2.3 Thermodynamics of dilute gases

The pressure of a gas in a container is produced by the incessant drumming of the gas molecules upon the walls. At each such wall collision – say, against the right wall of a cubic box – the respective momentum of the molecule $p_x \equiv mv_x$ is reversed. The momentum transferred to the wall is thus $\Delta p_x = 2mv_x$. The force acting on the unit area of the wall is then just the time average of the momentum transfer:

$$P \equiv \frac{\langle K \rangle_t}{F} \equiv \left\langle \frac{1}{Ft} \sum_{k=1}^{M(t)} \Delta p_x^{(k)} \right\rangle \equiv \left\langle \frac{1}{Ft} \sum_{k=1}^{M(t)} 2mv_x^{(k)} \right\rangle \quad (2.35)$$

where $M(t)$ is the number of wall impacts within the time t .

To obtain a theoretical prediction for the value of the pressure we argue as follows:

The number of particles having x -velocity v_x and impinging on the right wall per unit time is obviously proportional to v_x . The momentum transfer from such a particle to the wall is $2mv_x$. Thus we find

$$P = \frac{N}{V} \int_{v_x > 0} d\vec{v} 2mv_x^2 f(\vec{v}) \quad (2.36)$$

Inserting the Boltzmann density for $f(\vec{v})$ and performing the integrations we have

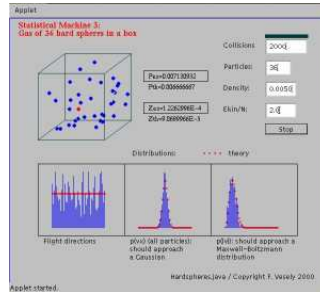
$$P = \frac{2}{3} \frac{N \langle E \rangle}{V} \quad (2.37)$$

Simulation: Thermodynamics of dilute gases

3-dimensional system of hard spheres in a cubic box. The pressure is determined

- from the collisions of the particles with the container walls (P_{ex})
- from equation 2.37 (P_{th}).

[Code: Hspheres]



Where do we stand now? Just by statistical reasoning we have actually arrived at a prediction for the pressure – a thermodynamic quantity!

Having thus traversed the gap to macroscopic physics we take a few steps further. From thermodynamics we know that the pressure of a dilute gas at density $\rho \equiv N/V$ and temperature T is $P = \rho kT$. Comparing this to the above formula we find for the mean energy of a molecule

$$\langle E \rangle = \frac{3kT}{2}, \quad (2.38)$$

and the parameter β (which was introduced in connection with the equilibrium density in velocity space) turns out to be just the inverse of kT .

The further explication of the dilute gas thermodynamics is easy. Taking the formula for the internal energy, $E \equiv N \langle E \rangle = 3NkT/2$, and using the First Law

$$dQ = dE + PdV \quad (2.39)$$

we immediately find that amount of energy that is needed to raise the temperature of a gas by 1 degree – a.k.a. the heat capacity C_V :

$$C_V \equiv \left. \frac{dQ}{dT} \right|_V = \frac{dE}{dT} = \frac{3}{2}Nk \quad (2.40)$$

2.4 Transport processes

A system is in “equilibrium” if its properties do not change spontaneously over time. (In case there are external fields – such as gravity – the material properties may vary in space, otherwise they will be independent also of position: the system is not only in equilibrium but also homogeneous.)

In particular, in equilibrium the local **energy density**, i.e. the energy contained in a volume element dV divided by that dV , as well as the **momentum density** and the **particle** or **mass density** remain constant. These densities, which refer to the three conserved quantities of mechanics, play a special role in what follows.

Again considering the dilute gas, and assuming there are no external fields, the densities are

$$\rho_E \equiv \frac{E}{V} = \frac{1}{V} \sum_{i=1}^N \frac{mv_i^2}{2}, \quad \rho_{\vec{p}} \equiv \frac{\vec{p}}{V} = \frac{1}{V} \sum_{i=1}^N m\vec{v}_i, \quad \text{and} \quad \rho_m \equiv \frac{Nm}{V} \quad (2.41)$$

By looking more closely we would see that these local densities are in fact not quite constant; rather, they will fluctuate about their mean. In other words, there will be a spontaneous waxing and waning of local **gradients** of those densities. Also, the experimenter may intervene to create a gradient. For instance, one might induce in a horizontal layer of a gas or liquid a certain x -velocity $u_x \neq 0$, and thereby a momentum p_x . The difference between the momenta in adjacent layers then defines a gradient of the momentum density. If we left the system to itself this gradient would decrease and eventually vanish. The property governing the speed of this decrease is called **viscosity**.

If a physical quantity – such as mass, or momentum – is conserved, any local change can only be achieved by a flow into or out of the space region under consideration; we are then speaking of **transport processes** whose speeds are governed by the respective **transport coefficients** – such as the viscosity η , the heat conductivity λ (for the energy transport) and the diffusion constant D (for mass transport).

In real experiments aimed at determining these transport coefficients the respective gradient is artificially maintained. Thus there will be a continuing flow of momentum, energy or matter in the (reverse) direction of the gradient – which is obviously a non-equilibrium situation. However, by a careful setup we can keep at least these flows and the local densities and gradients constant in time. This is called a **stationary non-equilibrium situation**.

DEFINITION OF TRANSPORT COEFFICIENTS

In order to define the transport coefficients η , λ and D we consider the basic experimental setup.

- **Viscosity:** To measure the coefficient of momentum transport η we generate a **laminary flow** by placing a gas or fluid layer between two horizontal plates and moving the upper plate with constant velocity u_0 to the right. In this manner we superimpose a systematic x -velocity onto the random thermal motion of the molecules.

The magnitude of the thermal speed is of the order $10^3 m/s$; by adding a “shear velocity” $u_x(z)$ of some centimeters per second the local equilibrium is not considerably disturbed. Thus we may assume that we have still a Maxwell-Boltzmann distribution of velocities at any point in the fluid, with the same value of $\langle E \rangle$ (bzw. kT) everywhere.

Yet by imposing a velocity gradient we have slightly perturbed the equilibrium; a certain amount of x -momentum will flow against the gradient – in our case downwards – so as to reestablish equilibrium. The amount of momentum flowing down through a unit of area

per unit of time is called **flow density** \vec{j}_p . In a first (linear) approximation this flow will be proportional to the imposed velocity gradient, and the coefficient of proportionality is, by definition, the transport coefficient η :

$$j_p = -\eta \frac{du_x(z)}{dz} \quad (2.42)$$

This defining equation for η is often called **Newton's law** of viscous flow. The parameter η is known as viscosity.

- **Thermal conductivity:** To generate a gradient of energy density we may place the substance between two heat reservoirs with different temperatures. If we use the same simple geometry as for the viscosity then the temperature (or energy) gradient will have only the one non-vanishing component dT/dz which accordingly produces a flow of energy in the counter direction. The coefficient λ of thermal conduction is defined by (**Fourier's law**),

$$j_E = -\lambda \frac{dT(z)}{dz} \quad (2.43)$$

- **Diffusion constant:** Even in a homogeneous gas (or a fluid, or a solid, for that matter) the individual molecules will alter their positions in the course of time. This process is known as **self diffusion**. It is possible to “tag” some of the particles – for instance, by giving them a radioactive nucleus. The density of the marked species – call it 1 – may then once more have a gradient that will be balanced by a particle flow (**Fick's law**):

$$j_1 = -D \frac{d\rho_1(z)}{dz} \quad (2.44)$$

MEAN FREE PATH

We will presently attempt to calculate the above defined transport coefficients from the microscopic dynamics of the molecules. In a simple kinetic treatment the elementary process determining the various flows is the free flight of a gas particle between two collisions with other molecules. Thus the first step should be an estimation of the average path length covered by a particle between two such collisions.

Let $P(x)$ be the – as yet unknown – probability that a particle meets another one **before** it has passed a distance x . Then $1 - P(x)$ is the probability for the molecule to fly a distance x without undergoing a collision. But the probability of an encounter within an infinitesimal path element dx is given by $p = \rho\sigma^2\pi dx$, where σ is the diameter of the particles. (p equals the fraction of the “target area” covered by other particles.) Thus the differential probability for a collision between x and $x + dx$ is

$$dP(x) = [1 - P(x)]\rho\sigma^2\pi dx \quad (2.45)$$

which upon integration yields

$$P(x) = 1 - e^{-\rho\sigma^2\pi x} \quad (2.46)$$

The **probability density** $p(x) \equiv dP(x)/dx$ for a collision within that interval is

$$p(x) = \rho\sigma^2\pi e^{-\rho\sigma^2\pi x} \quad (2.47)$$

The **mean free path** is then the first moment of this density:

$$l \equiv \langle x \rangle \equiv \int_0^\infty dx x p(x) = \frac{1}{\rho\sigma^2\pi} \quad (2.48)$$

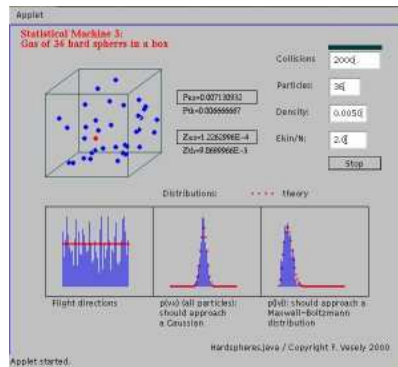
EXAMPLE: The diameter of a H_2 molecule is $\sigma \approx 10^{-10}m$ ($= 1 \text{ \AA}$ ngström). According to Loschmidt (and the later but more often quoted Avogadro) a volume of $22.4 \cdot 10^{-3}m^3$ contains, at standard conditions, $6.022 \cdot 10^{23}$ particles; the particle density is thus $\rho = 2.69 \cdot 10^{25}m^{-3}$, and the mean free path is $l = 1.18 \cdot 10^{-6}m$.

Simulation: Collision rate in a dilute gas

3-dimensional system of hard spheres in a cubic box:

– Compare the theoretical collision rate Z_{th} with the empirical one Z_{ex}

[Code: Hspheres]



ESTIMATING THE TRANSPORT COEFFICIENTS

In the case of a dilute gas we can actually find expressions for the transport coefficients. Assuming once more the basic geometric setting used in most experiments, we suppose that the gradient of the respective conserved quantity has only a z component, leading to a flow that is also directed along the z axis.

- **Viscosity:** In the course of the **random** (thermal) motion of the particles the small **systematic** term $u_x(z)$ is carried along. On the average, an equal number of particles will cross the interface between two horizontal layers from below and above. However, since the upper layer has a larger u_x (assuming $du_x/dz > 0$, it will gradually slow down, while the lower layer will take up speed. The flow of momentum due to this mechanism can be estimated as follows:

The mean number of particles crossing the unit of area from below and from above, respectively, is $\rho\langle v \rangle/6$. Each of the molecules carries that systematic $u_x(z)$ which pertains to the layer of its previous collision, i.e. $u_x(z \pm l)$. The momentum flow density is therefore

$$j_{\vec{p}} = \frac{1}{6}\rho\langle v \rangle [mu(z - l) - mu(z + l)] \approx \frac{1}{6}\rho\langle v \rangle \left[-2m \frac{du}{dz} l \right] \tag{2.49}$$

Comparing this to the defining equation for the viscosity η we find

$$\eta = \frac{1}{3}\rho l \langle v \rangle m \tag{2.50}$$

It is a somewhat surprising consequence of this formula that, since $l \propto 1/\rho$, the viscosity *eta* is apparently independent of density! Maxwell, who was the first to derive this result, had to convince himself of its validity by accurate experiments on dilute gases.

- **Thermal conductivity:** A similar consideration as in the case of momentum transport yields, with $j_E = -\lambda dT/dz$,

$$\lambda = \frac{1}{3}\rho l \langle v \rangle m c_V \quad (2.51)$$

where $c_V \equiv C_V/M$ is the specific heat (C_V .. molar heat capacity; M ..molar mass).

- **Diffusion:** Again we assume a gradient of the relevant density – in this case, the particle density of some species 1 – along the z direction. Writing $j_1 = -D d\rho_1/dz$ and using the same reasoning as before we find

$$D = \frac{1}{3}l \langle v \rangle \quad (2.52)$$

2.5 Just in case ...

... someone has forgotten how to find the extremum of a many-variable function under additional constraints. To remind you: the method of undetermined Lagrange multipliers waits to be applied here.

Let $f(x, y) = x^2 + y^2$ be the given function. Of course, this is the equation of a paraboloid with its tip – or minimum – at the origin. However, let $g(x, y) \equiv x - y - 1 = 0$ be the constraint equation, meaning that we don't search for the global minimum but for the minimum along the line $y = x - 1$. There are two ways to go about it. The simple but inconvenient way is to substitute $y = x - 1$ in $f(x, y)$, thus rendering f a function of x only. Equating the derivative df/dx to zero we find the locus of the conditional minimum, $x = 1/2$ and $y = -1/2$. The process of substitution is, in general, tedious.

A more elegant method is this: defining a (undetermined) Lagrange multiplier α , find the minimum of the function $f(x, y) - \alpha g(x, y)$ according to

$$\frac{\partial f}{\partial x} - \alpha \frac{\partial g}{\partial x} = 0 \quad (2.53)$$

$$\frac{\partial f}{\partial y} - \alpha \frac{\partial g}{\partial y} = 0 \quad (2.54)$$

Eliminating α we find the solution without substituting anything. In our case

$$2x - \alpha = 0 \quad (2.55)$$

$$2y + \alpha = 0 \quad (2.56)$$

$\implies x + y = 0$, and from $g(x, y) = 0$: $(x, y) = (1/2, -1/2)$.

2.6 Problems for Chapter 2

EXERCISES:

2.1 Variational method of Lagrange: Determine the minimum of the function $f(x, y) = x^2 + 2y^2$ under the condition (i.e. along the curve) $g(x, y) = x + y - 6 = 0$, by two different methods:

a) by inserting $y = 6 - x$ in $f(x, y)$ and differentiating by x ;

b) using a Lagrange multiplier α and evaluating Lagrange's equations $\partial f/\partial x - \alpha(\partial g/\partial x) = 0$ and $\partial f/\partial y - \alpha(\partial g/\partial y) = 0$. (Shorthand notation: $\delta f - \alpha \delta g = 0$; δf ... "Variation of f ").

For your better understanding sketch the functions $f(x, y)$ and $g(x, y) = 0$.

2.2 Method of the most probable distribution: Having understood the principle of Lagrange variation, reproduce the derivation of the Boltzmann distribution of energies (see text). Compare your result with the bottom right histogram in Applet *LBRoulette*.

2.3 Moments of the Maxwell-Boltzmann distribution: Verify the expressions for the most probable velocity, the average and the r.m.s. velocity.

2.4 Pressure in a dilute gas: Verify the formula $P = (2/3)(N\langle E\rangle/V)$.

2.5 Transport properties: For nitrogen under standard conditions, estimate the mean free path and the transport coefficients η , λ and D . ($\sigma \approx 4 \cdot 10^{-10}m$).

TEST YOUR UNDERSTANDING OF CHAPTER 2:

2.1 Maxwell-Boltzmann distribution: What is the formula for the distribution density $f_0(\vec{v})$ of the molecular velocities in equilibrium; what is the respective formula for the speeds (absolute values of the velocity), $f_0(|\vec{v}|)$?

2.2 Pressure in an ideal gas: Derive the pressure equation of state for the ideal gas from simple kinetic theory.

2.3 Transport coefficients:

- Write down the defining equation for one of the three transport coefficients.
- What is the mean free path in a gas of spheres with diameter σ ?
- How does the viscosity of a dilute gas depend on the density?

Chapter 3

Phase space

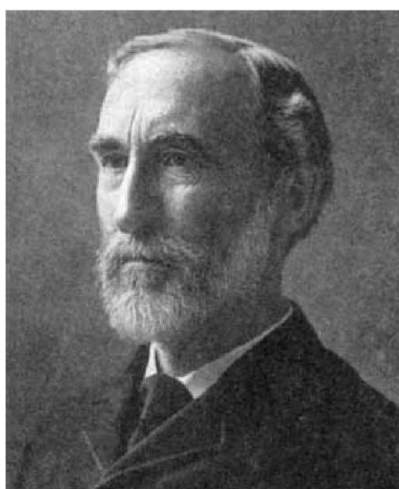


Figure 3.1: Josiah Willard Gibbs

The Kinetic Theory explained in Chapter 2 is applicable only to gases and – with certain add-ons – to fluids. A more general and extremely powerful method for analyzing statistical-mechanical systems is based on a geometrical exploration of the high-dimensional phase space spanned by the complete set of microvariables.

3.1 Microscopic variables

The **microscopic state** of a model system is uniquely determined by the specification of the complete set of **microscopic variables**. The number of such variables is of the same order as the number of particles. In contrast, the **macroscopic state** is specified by a small number of measurable quantities such as net mass, energy, or volume. Generally a huge number of different microstates are compatible with one single macrostate; and it is a primary task of statistical mechanics to find out just how many microstates there are for a given set of macroscopic conditions.

Interpreting the microscopic variables as coordinates in a high-dimensional space we may represent a particular microstate as a **point** or **vector** in that space. The space itself is called **Gibbs phase space**; the state vector is often symbolized as $\vec{\Gamma}$.

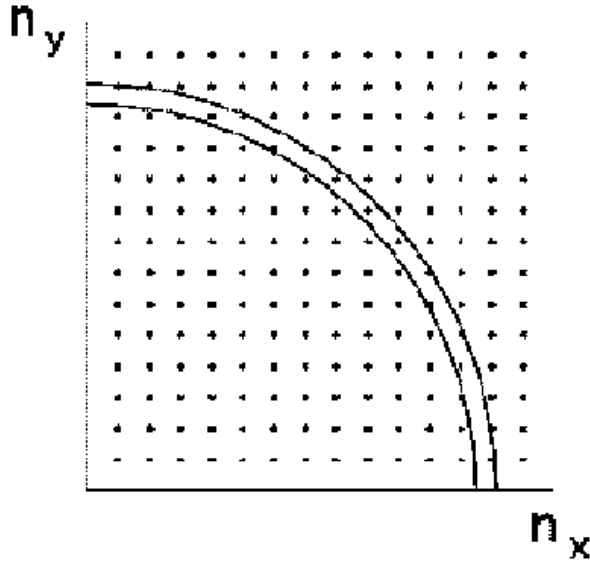


Figure 3.2: Phase space of the 2D quantum gas

Consider a simple, classical many-particle system – say, an ideal gas, or a fluid made up of hard spheres or Lennard-Jones molecules. The state vector is then defined by all position and velocity coordinates:

$$\vec{\Gamma} \equiv \{\vec{r}_1, \dots, \vec{r}_N; \vec{v}_1, \dots, \vec{v}_N\} \quad \vec{r}_i \in V; v_{i,\alpha} \in (\pm\infty) \quad (3.1)$$

The number of degrees of freedom (d.o.f.) is $n = 6N$ (or, in 2 dimensions, $4N$); the number of velocity d.o.f. is $3N$ (or $2N$). It is often allowed – and always advantageous – to treat the subspaces $\vec{\Gamma}_r \equiv \{\vec{r}_i\}$ (position space) and $\vec{\Gamma}_v \equiv \{\vec{v}_i\}$ (velocity space) separately.

The representation of a microstate of the entire system by a single point in $6N$ -dimensional Gibbs space must not be mixed up with the treatment of Chapter 2, introduced by Boltzmann. The μ -space defined there had only 6 dimensions $\{\vec{r}, \vec{v}\}$, and each **particle** in the system had its own representative point. Thus the microstate of a system of N particles corresponded to a swarm of N points $\{\vec{r}_i, \vec{v}_i\}$.

In the case of an ideal quantum gas it suffices to specify all quantum numbers to define a state:

$$\vec{\Gamma} \equiv \{(n_{ix}, n_{iy}, n_{iz}); i = 1, \dots, N\} \quad (3.2)$$

Model systems made up of N spins are specified by

$$\vec{\Gamma} \equiv \{\sigma_i; i = 1, \dots, N\} \quad (3.3)$$

with $\sigma_i = \pm 1$.

ENERGY SURFACE

We consider a collection of M systems, all having the same energy E , the same number of particles N , and – in the case of a gas or fluid – the same volume V . All microstates $\vec{\Gamma}$ compatible with these macroscopic conditions are then equally probable, that is, they have the same relative

frequency within the ensemble of systems. The size M of the ensemble is assumed to be very large – indeed, to approach infinity.

The assumption that all microstates that are compatible with the condition $E = E_0$ have the same probability is one of the solid foundations Statistical Mechanics is built upon. It is called the “postulate of equal a priori probability”. For a mathematically exact formulation of this axiom we use the **phase space density** which is supposed to have the property

$$\rho(\vec{r}, \vec{v}) = \begin{cases} \rho_0 & \text{for } E(\vec{r}, \vec{v}) = E_0 \\ 0 & \text{elsewhere} \end{cases} \quad (3.4)$$

The condition $E = \text{const}$ defines an $(n-1)$ -dimensional “surface” within the n -dimensional phase space of the system. The set of all points upon that “energy surface” is named **microcanonical ensemble**. To take an example, in the $3N$ -dimensional velocity space of a classical ideal gas the equation

$$E_{kin} \equiv \frac{m}{2} \sum_i v_i^2 = \text{const} = E_0 \quad (3.5)$$

defines the $(3N-1)$ -dimensional surface of a $3N$ -sphere of radius $r = \sqrt{2E_0/m}$.

A typical question to be answered by applying statistical methods to such an ensemble is this: what is the average of the squared particle velocity $\langle v_i^2 \rangle$ over all – a priori equally probable – states on this surface?

ERGODICITY

Instead of considering an ensemble of systems let us now watch just one single system as it evolves in time according to the laws of mechanics. In a closed system the total energy will remain constant; the microstates visited by the system must therefore lie on the $(N-1)$ -dimensional energy surface defining the microcanonical ensemble.

The **ergodic hypothesis** states that in the course of such a “natural evolution” of the system any permitted microstate will be reached (or closely approximated) with the same relative frequency.

This hypothesis cannot be proven in general; in fact, it does not always hold. However, for many relevant systems such as gases or fluids under normal conditions it is quite true. In such cases the time t_0 needed for a sufficiently thorough perambulation of the energy surface is in the range of $10^{-9} - 10^{-6}$ seconds, i.e. safely below the typical observation time in an experiment. Among those systems which we may characterize as “barely ergodic” or non-ergodic we have supercooled liquids and glasses. In such systems the state vector Γ remains trapped for long times in a limited region of the energy surface; it may then take seconds, days, or even centuries before other parts of the microcanonical surface are reached.

The ergodic hypothesis, if true, has an important practical consequence: for the calculation of mean values over the microstates on the energy surface it does not matter if we take the average over states randomly picked from a microcanonical ensemble, or over the successive states of one single, isolated system. This corollary of the ergodic hypothesis is often succinctly stated as

$$\text{ensemble average} = \text{time average}$$

The assumption of ergodicity enables us to support our theoretical arguments by “molecular dynamics” computer experiments. These are deterministic simulations reproducing the temporal evolution of a single isolated N -particle system. We will later touch upon another kind of computer simulation, in which the state space is perambulated in a *stochastic* manner; it bears the suggestive name “Monte Carlo simulation”.

ENERGY SHELL

In place of the strict condition $E = E_0$ we will generally require the weaker condition $E_0 - \Delta E \leq E \leq E_0$ to hold. In other words, the permitted states of the system are to be restricted to a

Model	Microvariables ($i = 1, \dots, N$)	Dimension of phase space	Energy
Classical ideal gas	\vec{r}_i, \vec{v}_i	$6N$	$E_{kin} \equiv \frac{m}{2} \sum v_i^2$
Hard spheres in a box	\vec{r}_i, \vec{v}_i	$6N$	$E_{kin} \equiv \frac{m}{2} \sum v_i^2$
Hard discs in a box	\vec{r}_i, \vec{v}_i	$4N$	$E_{kin} \equiv \frac{m}{2} \sum v_i^2$
LJ molecules, periodic boundary conditions	\vec{r}_i, \vec{v}_i	$6N$ (or. $6N - 3$)	$E_{kin} + E_{pot} \equiv \frac{m}{2} \sum v_i^2 + \sum_{i,j>i} U_{LJ}(r_{ij})$
Linear or nonlinear molecules	$\vec{r}_i, \vec{v}_i, \vec{e}_i, \vec{\omega}_i$	$10N$ or. $12N$	$E_{kin} + E_{pot} \equiv \frac{m}{2} \sum v_i^2 + \frac{I}{2} \sum \omega_i^2 + E_{pot}$
Harmonic crystal	$\vec{q}_i, \dot{\vec{q}}_i$	$6N$	$E_{kin} + E_{pot} \equiv \frac{m}{2} \sum \dot{q}_i^2 + \frac{f}{2} \sum \vec{q}_i - \vec{q}_{i-1} ^2$
Ideal quantum gas	\vec{n}_i	$3N$	$\frac{\hbar^2}{8mL^2} \sum_i \vec{n}_i ^2$
Spins, non-interacting	σ_i	N	$-H \sum_i \sigma_i$

Table 3.1: State spaces of some model systems

thin “shell” at $E \approx E_0$. This more pragmatic requirement – which is made to keep the mathematics simpler – agrees well with the experimental fact that the exchange of energy between a system and its environment may be kept small, but can never be completely suppressed.

Thus we assume for the density in phase space that

$$\rho(\vec{r}, \vec{v}) = \begin{cases} \rho_0 & \text{for } E(\vec{r}, \vec{v}) \in [E_0, \Delta E] \\ 0 & \text{elsewhere} \end{cases} \quad (3.6)$$

For the classical ideal gas this requirement reads¹

$$E_0 - \Delta E \leq E_{kin} \equiv (m/2) \sum_i v_i^2 \leq E_0 \quad (3.7)$$

For the ideal quantum gas we require

$$E_0 - \Delta E \leq E_{N, \vec{n}} \equiv \frac{\hbar^2}{8mL^2} \sum_i \vec{n}_i^2 \leq E_0 \quad (3.8)$$

In a system of N non-interacting spins the energy shell is defined by $-H \sum_i \sigma_i \in [E_0, \Delta E]$.

Table 3.1 presents an overview of the state spaces and energy functions for various model systems.

PHASE SPACE VOLUME AND THERMODYNAMIC PROBABILITY

Let us now recall the fundamental assumption that all *microstates* having the same energy are equally probable. Thus they represent *elementary events* as defined in Section 1.3.

It follows then that a *macrostate* that allows for more *microstates* than others will occur more often – and thus will be more probable.

¹In an MD simulation not only the total energy but also the total momentum \vec{P} is conserved; this should be kept in mind when comparing simulation results to theoretical statements. Actually, the four conditions $E \in (E_0 - \Delta E, E_0)$ and $\vec{P} = \text{const}$, would define a $3N - 3$ -dimensional spherical shell in velocity space. However, for sufficiently numerous particles ($N \geq 100$) this detail may be neglected, and we can base the argumentation just on the condition $E \approx E_0$ (with a $3N$ -dimensional shell).

As an example, we may ask for the probability to find all molecules of an ideal gas in the left half of the vessel, with no other restrictions on position or speed. All such microstates are located in a small part of the total permitted phase space shell $[E, \Delta E]$; they make up a *sub-ensemble* of the complete microcanonical ensemble. Now, the probability of the macrostate “all particles in the left half of the container” is obviously equal to the size ratio of the subensemble and the total ensemble. We will have to compare the volume (in phase space) of the partial shell pertaining to the subensemble to the volume of the total shell. (We will see later that the ratio of the two volumes is $(1/2)^N$ and may thus be neglected – in accordance with experience and intuition.)

3.2 From Hyperspheres to Entropy

Some model systems have the convenient property that their energy may be written as a sum over the squares of their microvariables. The most popular example is the classical ideal gas with $E \equiv E_{kin} = (m/2) \sum v_i^2$. The conditions $E \leq E_0$ or $E \approx E_0$ then describe a high-dimensional sphere or spherical shell, respectively. Thus it will be worth the while to make ourselves acquainted with the geometrical properties of such bodies. It must be stressed, however, that the restriction to n -spheres is only a matter of mathematical convenience. It will turn out eventually that those properties of n -dimensional phase space which are of importance in Statistical Mechanics are actually quite robust, and not reserved to n -spheres. For instance, the n -rhomboid which pertains to the condition $\sum_i \sigma_i \approx const$ for simple spin systems could be (and often is) used instead.

VOLUME AND SURFACE OF HIGHDIMENSIONAL SPHERES

In the following discussion it will be convenient to use, in addition to the sphere radius r_0 , a variable that represents the **square of the radius**: $z_0 \equiv r_0^2$. In the phase space of interactionless many-particle systems, this quantity is related to an **energy**, as may be seen from eqs. 3.7, 3.8 and Table 3.1); and in an isolated system it is just the energy which is given as a basic parameter.

Attention:

It should be kept in mind that the energy – and thus the square of the sphere radius – is an *extensive* quantity, meaning that it will increase as the number of degrees of freedom (or particles): $z_0 \propto n$. It would be unphysical and misleading to keep z_0 at some constant value and at the same time raise the number of dimensions.

VOLUME

The volume of a n -dimensional sphere with radius $r_0 = \sqrt{z_0}$ is

$$V_n(z_0) = C_n r_0^n = C_n z_0^{n/2} \quad \text{with} \quad C_n = \frac{\pi^{n/2}}{(n/2)!} \quad (3.9)$$

where $(1/2)! = \sqrt{\pi}/2$ and $(x+1)! = x!(x+1)$. The following recursion is useful:

$$C_{n+2} = \frac{2\pi}{n+2} C_n \quad (3.10)$$

EXAMPLES: $C_1 = 2$, $C_2 = \pi$, $C_3 = \pi^{3/2}/(3/2)! = 4\pi/3$, $C_4 = \pi^2/2$, $C_5 = 8\pi^2/15$, $C_6 = \pi^3/6$, ... $C_{12} = \pi^6/720$, ...

For large n we have, using Stirling's approximation for $(n/2)!$,

$$C_n \approx \left[\frac{2\pi e}{n} \right]^{n/2} \frac{1}{\sqrt{\pi n}} \quad \text{or} \quad \ln C_n \approx \frac{n}{2} \ln \frac{2\pi e}{n} \approx \frac{n}{2} (\ln \pi + 1) - \frac{n}{2} \ln \frac{n}{2} \quad (3.11)$$

As soon as the Stirling approximation holds, i.e. for $n \geq 100$, the hypersphere volume may be written

$$V_n(z_0) \approx \frac{1}{\sqrt{\pi n}} \left(\frac{2\pi e}{n} \right)^{n/2} z_0^{n/2} \quad \text{or} \quad \ln V_n(z_0) = \frac{n}{2} \ln \left(\frac{2\pi e}{n} z_0 \right) - \ln \sqrt{\pi n} \quad (3.12)$$

Going to very large $n (\geq 500)$ we find for the logarithm of the volume (the formula for V proper is then difficult to handle due to the large exponents)

$$\ln V_n(z_0) \approx \frac{n}{2} \ln \left(\frac{2\pi e}{n} z_0 \right) \quad (3.13)$$

EXAMPLE: Let $n = 1000$ and $z_0 = 1000$. For the log volume we find $\ln V_{1000}(1000) \equiv 500 \ln(2\pi e) - \ln \sqrt{1000\pi} = 882.76$

SURFACE AREA

For the surface area of a n -dimensional sphere we have

$$O_n(z_0) = \frac{dV_n(z_0)}{dr_0} = nC_n z_0^{(n-1)/2} \quad (3.14)$$

EXAMPLES:

$$\begin{aligned} V_1(z_0) &= 2r_0, & O_1(z_0) &= 2 \\ V_2(z_0) &= \pi z_0, & O_2(z_0) &= 2\pi r_0 \\ V_3(z_0) &= (4\pi/3)r_0^3, & O_3(z_0) &= 4\pi z_0 \\ V_4(z_0) &= (\pi^2/2)z_0^2, & O_4(z_0) &= 2\pi^2 r_0^3 \end{aligned}$$

A very useful representation of the surface is this:

$$O_n(r_0) = \int_{-r_0}^{r_0} dr_1 \frac{r_0}{r_2} O_{n-1}(r_2) \quad (3.15)$$

with $r_2 \equiv \sqrt{r_0^2 - r_1^2}$. This formula provides an important insight; it shows that the “mass” of a spherical shell is distributed along one sphere axis (r_1) as follows:

$$p_n(r_1) = \frac{r_0}{r_2} \frac{O_{n-1}(r_2)}{O_n(r_0)} = \frac{(n-1)C_{n-1}}{nC_n} \frac{r_2^{n-3}}{r_0^{n-2}} \quad (3.16)$$

EXAMPLES: Let $r_0 = 1$; then we find (see Fig. 3.3)

$$\begin{aligned} p_2(r_1) &= \frac{1}{\pi} (1 - r_1^2)^{-1/2} \\ p_3(r_1) &= \frac{1}{2} \quad (\text{constant!}) \\ p_4(r_1) &= \frac{2}{\pi} (1 - r_1^2)^{1/2} && \text{(The astute reader notes that for once we} \\ p_5(r_1) &= \frac{3}{4} (1 - r_1^2) \\ &\dots \\ p_{12}(r_1) &= \frac{256}{63\pi} (1 - r_1^2)^{9/2} \end{aligned}$$

have kept $r_0 = 1$, regardless of the value of n ; this is permitted here because we are dealing with a normalized density p_n .)

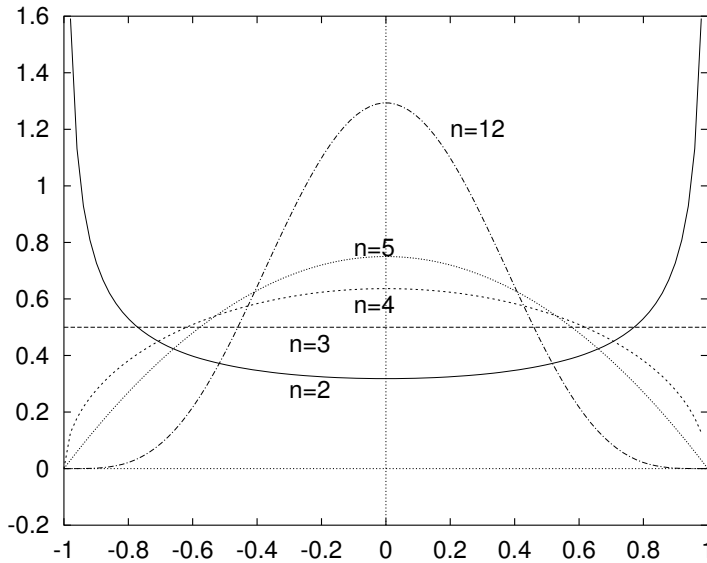


Figure 3.3: Mass distribution $p(r_1)$ of a $(n - 1)$ -dimensional spherical surface along one axis. With increasing dimension the mass concentrates more and more at the center of the axis. If we interpret the surface as the locus of all phase space points with given total energy $E_{kin} = (m/2) \sum v_i^2$, then $p(r_1) \equiv p(v_i)$ is just the distribution density for any single velocity component v_i .

APPLICATION: Assume that a system has n degrees of freedom of translatory motion. (Example: n particles moving on a line, or $n/3$ particles in 3 dimensions.) Let the sum of squares of all velocities (energy!) be given, but apart from that let any particular combination of the values v_1, v_2, \dots be equally probable. All “phase space points” $\vec{v} \equiv \{v_1 \dots v_n\}$ are then homogeneously distributed on the spherical surface $O_n(|v|^2)$, and a single velocity v_1 occurs with probability density 3.16.

As we increase the number of dimensions, the character of this density function changes dramatically at first (see Fig. 3.3). If just *two* particles on a line (or the two d.o.f. of a pin ball) share the total energy $mv^2/2$, then the velocity of one of them is most probably near the possible maximal value while the other has only a small speed. In contrast, for many dimensions (or particles) the maximum of the probability density $p_n(v_1)$ is near zero. The case $n = 3$, meaning 3 particles on a line or one particle in 3 dimensions, is special: all possible values of v_1 occur with equal probability.

Approach to the **Maxwell-Boltzmann distribution:** For very large n we have

$$p_n(v) \approx \frac{1}{\sqrt{2\pi\langle v^2 \rangle}} \exp\{-v^2/2\langle v^2 \rangle\} \quad (3.17)$$

with $\langle v^2 \rangle = 2E_0/nm$.

The Maxwell-Boltzmann distribution may thus be derived solely from the postulate of equal a priori probability and the geometric properties of high-dimensional spheres.

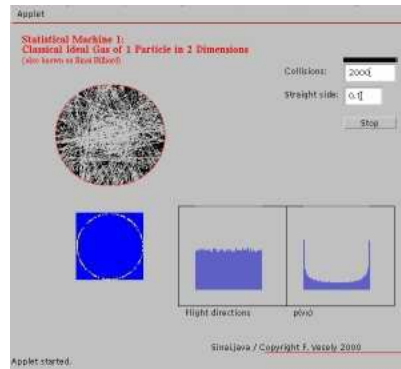


Figure 3.4: Bunimovich’s stadium billiard. Distribution of flight directions, $p(\phi)$ and of a single velocity component, $p(v_x)$

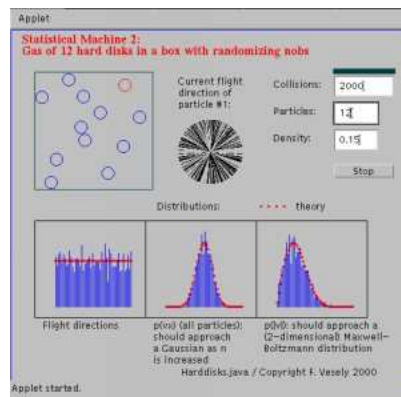


Figure 3.5: Simulation: $N = 1$ to 3 hard disks in a 2D box with reflecting walls. Distribution of flight directions, $p(\phi)$, and of a single velocity component, $p(v_{1x})$. Comparison with the theoretical distribution [Code: Harddisks]

TWO AMAZING PROPERTIES OF HYPERSPHERES

We have seen that spheres in high-dimensional spaces exhibit quite unusual geometrical properties. In the context of Statistical Physics the following facts are of particular relevance:

- Practically the entire volume of a hypersphere is assembled in a thin shell immediately below the surface
- The volume of a hypersphere is – at least on a logarithmic scale – almost identical to the volume of the largest inscribed hyper-cylinder

We will consider these two statements in turn.

VOLUME OF A SPHERICAL SHELL

The ratio of the volume $\Delta V_n(r_0 - \Delta r, r_0)$ of a thin “skin” near the surface and the total volume of the sphere is

$$\frac{\Delta V_n}{V_n} = \frac{r_0^n - (r_0 - \Delta r)^n}{r_0^n} = 1 - \left(1 - \frac{\Delta r}{r_0}\right)^n \longrightarrow 1 - \exp\{-n(\Delta r/r_0)\} \quad (3.18)$$

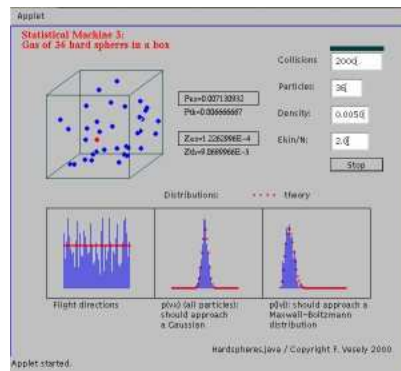


Figure 3.6: Simulation: One hard sphere in a box with reflecting walls. Distribution of one velocity component; demonstration of the special case $n = 3$ (where $p(v_{1x}) = \text{const}$). [Code: Hspheres]

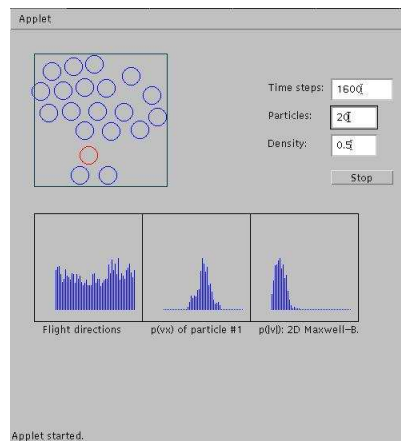


Figure 3.7: Simulation: N Lennard-Jones particles in a 2D box with periodic boundary conditions. Distribution of a velocity component, $p(v_{1x})$. [Code: LJJones]

or, using the quantities z_0 and Δz :

$$\frac{\Delta V_n}{V_n} \longrightarrow 1 - \exp\left[-\frac{n \Delta z}{2 z_0}\right] \quad (3.19)$$

For $n \rightarrow \infty$ this ratio approaches 1, **regardless of how thin the shell may be!**

At very high dimensions the entire volume of a sphere is concentrated immediately below the surface:

$$\Delta V_n(z_0, \Delta z) \longrightarrow V_n(z_0), \quad n \gg 1 \quad (3.20)$$

EXAMPLE: $n = 1000, \Delta r = r_0/100 \rightarrow \Delta V/V = 1 - 4.3 \cdot 10^{-5}$

HYPERSPHERES AND HYPERCYLINDERS

The sphere volume $V_n(z_0)$ may alternatively be written as

$$\begin{aligned}
V_n(z_0) &= \int_0^{r_0} dr_1 O_{n_1}(z_1) V_{n-n_1}(z_2) \\
&\equiv \int_0^{z_0} \frac{dV_{n_1}(z_1)}{dz_1} dz_1 V_{n-n_1}(z_0 - z_1) \\
&\equiv \int_0^{z_0} dV_{n_1}(z_1) V_{n-n_1}(z_0 - z_1)
\end{aligned} \tag{3.21}$$

with $r_1 \equiv \sqrt{z_1}$.

EXAMPLE:

$$V_6(z_0) = \int_0^{r_0} dr_1 O_3(z_1) V_3(z_0 - z_1) = \int_0^{z_0} dr_1 4\pi r_1^2 \frac{4\pi}{3} (z_0 - z_1)^{3/2} = \frac{\pi^3}{6} z_0^3 \tag{3.22}$$

Partitioning the integration interval $[0, z_0]$ in equ. 3.21 into small intervals of size Δz we may write

$$V_n(z_0) \approx \sum_{k=1}^K \Delta V_{n_1}(z_k) V_{n-n_1}(z_0 - z_k) \tag{3.23}$$

where $\Delta V_{n_1}(z_k) \equiv V_{n_1}(k\Delta z) - V_{n_1}((k-1)\Delta z)$.

Similarly, for the volume of the spherical shell we have

$$\Delta V_n(z_0) = \int_0^{z_0} dV_{n_1}(z_1) \Delta V_{n-n_1}(z_0 - z_1) \approx \sum_{k=1}^K \Delta V_{n_1}(z_k) \Delta V_{n-n_1}(z_0 - z_k) \tag{3.24}$$

Remembering equ. 3.20 we may, for high enough dimensions n , n_1 and $(n - n_1)$, always write V in place of ΔV . Therefore,

$$V_n(z_0) \approx \sum_{k=1}^K V_{n_1}(z_k) V_{n-n_1}(z_0 - z_k) \tag{3.25}$$

Now, the terms in these sums are strongly varying. For high dimensions there is always one single term that dominates the sum; all other terms may safely be neglected. To find this term we evaluate

$$\frac{d}{dz} [V_{n_1}(z) V_{n-n_1}(z_0 - z)] = 0 \tag{3.26}$$

From

$$\frac{d [z^{n_1/2} (z_0 - z)^{(n-n_1)/2}]}{dz} = 0 \tag{3.27}$$

we find for the argument z^* that maximizes the integrand 3.21 or the summation term in 3.24 or 3.25,

$$z^* = \frac{n_1}{n} z_0 \tag{3.28}$$

We have thus found a very surprising property of high-dimensional spheres which we may summarize by the following calculation rule:

- Divide the given dimension n in n_1 and $n - n_1$ such that also $n_1 \gg 1$ and $n - n_1 \gg 1$

- Find the maximum of the product $f(z) \equiv V_{n_1}(z) \cdot V_{n-n_1}(z_0 - z)$ (or the maximum of the logarithm $\ln f(z)$) with respect to z ; this maximum is located at

$$z^* = \frac{n_1}{n} z_0 \quad (3.29)$$

- With this, the following holds:

$$\ln V_n(z_0) \approx \ln V_{n_1}(z^*) + \ln V_{n-n_1}(z_0 - z^*) \quad (3.30)$$

and similarly,

$$\ln \Delta V_n(z_0) \approx \ln \Delta V_{n_1}(z^*) + \ln \Delta V_{n-n_1}(z_0 - z^*) \quad (3.31)$$

(Remark: in a numerical verification of these relation it must be remembered that the maximum of the function at the r.h.s. of 3.31 or 3.30 is very sharp indeed; it may be overlooked when the interval $[0, z_0]$ is scanned in regular steps. It is better to localize z^* first, then compare with slightly smaller and larger z .)

The geometric interpretation of the relation 3.30 is amazing:

On a logarithmic scale the volume of an n -sphere equals the product of the volumes of two spheres in the subspaces n_1 and $n - n_1$. But this product may be understood as the volume of a hypercylinder that is inscribed in the n -sphere and which has the “base area” in n_1 space and the “height” in $(n - n_1)$ space.

EXAMPLE: $n = 1000$, $n_1 = 400$, $z_0 = 1000$: the maximum of the quantity $\ln f(z) \equiv \ln V_{400}(z) + \ln V_{600}(1000 - z)$ is located at $z^* = 400$, and we have

$$\begin{aligned} 500 \ln(2\pi e) - \ln \sqrt{1000\pi} &\approx 200 \ln(2\pi e) - \ln \sqrt{400\pi} + 300 \ln(2\pi e) - \ln \sqrt{600\pi} \\ 1418.94 - 4.03 &\approx 567.58 - 3.57 + 851.36 - 3.77 \end{aligned} \quad (3.32)$$

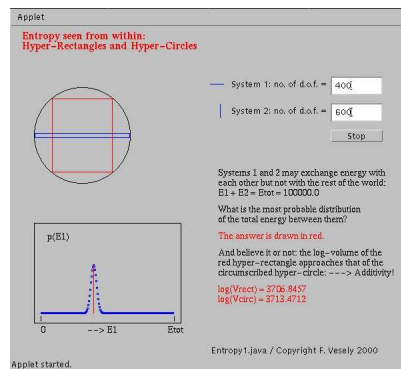


Figure 3.8: Simulation: Hyperspheres and -cylinders. In a given hypersphere of dimension n we inscribe hypercylinders whose “base areas” and “heights” have n_1 and $n - n_1$ dimensions, respectively. The hypercylinder with the maximum volume is identified: its log volume is almost equal to that of the circumscribed sphere. [Code: Entropy1]

DISCRETISATION OF PHASE SPACE; ENTROPY

Returning from geometry to physics, we will from now on denote phase space volumes by Γ , reserving the symbol V for the volume of a system in real space. Thus, $\Gamma_{\vec{v}}$ is the $3N$ -dimensional velocity phase space volume of a N -particle system, $\Gamma_{\vec{r}}$ refers to the positional subspace volume which usually will also have $3N$ dimensions, and Γ denotes the full phase space (or volume in phase space) with $6N$ dimensions.

Depending on which model system we are considering, the microvariables are either continuous (classical manybody system) or discrete (quantum models, spin lattices.) In order to be able to “enumerate” the states in phase space it is opportune to introduce a **raster** even in the case of continuous microvariables. Thus we imagine the phase space of a N -particle system to be divided into **cells** of size ²

$$g \equiv \Delta x \Delta v \quad (3.33)$$

Now let us consider a classical gas, or fluid, with particle number N and volume V . In the $3N$ -dimensional velocity subspace of the $6N$ -dimensional phase space the condition $E_{kin} \in [E, \Delta E]$ again defines a spherical shell whose volume is essentially equal to the volume of the enclosed sphere. The number of Δv cells in (or below) this shell is

$$\Sigma_{\vec{v}}(E, \Delta E) \equiv \frac{\Delta \Gamma_{\vec{v}}}{(\Delta v)^{3N}} \approx \frac{\Gamma_{\vec{v}}}{\Delta v^{3N}} = C_{3N} \left(\frac{2E}{m} \right)^{3N/2} / (\Delta v)^{3N} \quad (3.34)$$

In the case of the **ideal** gas the contribution of the position subspace may be included in a simple manner; we write

$$\Sigma_{\vec{r}} = \frac{V^N}{\Delta x^{3N}} \quad (3.35)$$

Thus the total number of cells in $6N$ space is

$$\begin{aligned} \Sigma'(N, V, E) = \Sigma_{\vec{r}} \Sigma_{\vec{v}} &= C_{3N} \left[\frac{V}{g^3} \left(\frac{2E}{m} \right)^{3/2} \right]^N \\ &= \frac{1}{\sqrt{3N\pi}} \left(\frac{2\pi e}{3N} \right)^{3N/2} \left[\frac{V}{g^3} \left(\frac{2E}{m} \right)^{3/2} \right]^N \end{aligned} \quad (3.36)$$

Now we have to prematurely introduce a result of quantum statistics. One characteristic property of quantum objects is their **indistinguishability**. It is evident that the number of distinct microstates will depend on whether or not we take into account this quantum property. A detailed analysis which must be postponed for now leads up to the simple result that we have to divide the above quantity $\Sigma'(N, V, E)$ by $N!$ to find

$$\Sigma(N, V, E) \equiv \Sigma'(N, V, E)/N! \quad (3.37)$$

which is now indeed proportional to the total number of physically distinct microstates. This rule is known as the **rule of correct Boltzmann enumeration**.

It is characteristic of the physical “instinct” of J. W. Gibbs that he found just this rule for the correct calculation of Σ although quantum mechanics was not yet known to him. He proposed the $1/N!$ rule in an ad hoc manner to solve a certain theoretical problem, the so-called *Gibbs Paradox*.

The quantity $\Sigma(N, V, E)$ is a measure of the available phase space volume, given in units g^{3N} . The *logarithm* of $\Sigma(N, V, E)$ has great physical significance: it is – up to a prefactor k – identical to the *entropy* $S(N, V, E)$ that was introduced in

²A particularly powerful formulation of classical mechanics, known as Hamilton’s formalism, makes use of the variables \vec{q}, \vec{p} (position and momentum) in place of \vec{r}, \vec{v} (position and velocity.) In this notation the phase space cells have the size $h \equiv \Delta q \Delta p$.

thermodynamics. At present, this identification is no more than a hypothesis; the following chapter will show how reasonable the equality

$$S(N, V, E) = k \ln \Sigma(N, V, E) \quad (3.38)$$

really is.

In the case of a classical ideal gas we find, using $\ln N! \approx N \ln N$ and neglecting $\ln \sqrt{3N\pi}$ in equ. 3.36,

$$S(N, V, E) \equiv k \ln \Sigma(N, V, E) = Nk \ln \left[\frac{V}{N} \left(\frac{4\pi Ee}{3Nmg^2} \right)^{3/2} \right] \quad (3.39)$$

This is the famous **Sackur-Tetrode equation**, named for the authors **Otto Sackur** and **Hugo Tetrode**.

The numerical value of Σ , and therefore of S , is obviously dependent on the chosen grid size g . This disquieting fact may be mitigated by the following considerations:

- As long as we are only **comparing** phase space volumes, or entropies, the unit is of no concern
- There is in fact a smallest physically meaningful gridsize which may well serve as the natural unit of Σ ; it is given by the quantum mechanical uncertainty: $g_{min} = h/m$

EXAMPLE: $N = 36$, $m = 2$, $E = N = 36$. The average energy per particle is then $= 1$, and the mean squared velocity is $\langle v^2 \rangle = 1$. Assuming a cubic box with $V = L^3 = 1$ and a rather coarse grid with $\Delta x = \Delta v = 0.1$ we find

$$S(N, V, E)/k = 36 \ln \left[\frac{1}{36} \left(\frac{4 \cdot \pi \cdot 36 \cdot e}{3 \cdot 36 \cdot 2 \cdot 10^{-4}} \right)^{3/2} \right] = 462.27 \quad (3.40)$$

Just for curiosity, let us determine $\ln(\Delta\Sigma)$, where $\Delta\Sigma$ is the phase space volume between $E = 35.5$ and 36.0 :

$$\Delta\Sigma = \Sigma(36.0) - \Sigma(35.5) = \Sigma(36.0) \left[1 - \left(\frac{35.5}{36} \right)^{54} \right] = \Sigma(36.0) \cdot 0.53 \quad (3.41)$$

and thus

$$S(N, V, \Delta E)/k = 462.27 - 0.63 = 461.64 \quad (3.42)$$

We see that even for such a small system the entropy value is quite insensitive to using the spherical shell volume in place of the sphere volume!

3.3 Problems for Chapter 3

EXERCISES:

3.1 Geometry of n -spheres I: Calculate the volumes and surfaces of spheres with $r = 1.0$ in 3,6,12 and 100 dimensions. In evaluating 100! use the **Stirling approximation** $n! \approx \sqrt{2\pi n} (n/e)^n$.

3.2 Geometry of n -spheres II: For the various n -spheres of the foregoing example, compute

the fraction of the total volume contained in a shell between $r = 0.95$ and $r = 1.0$. Comment the result.

3.3 Approximate formula for $\ln V_n(z_0)$: Choose your own example to verify the validity of the approximation 3.30. Put the meaning of this equation in words; demonstrate the geometrical meaning for a three-dimensional sphere – even if the approximation is not yet valid in this case. Make a few runs with Applet *Entropy1* and comment on your findings.

3.4 Sackur-Tetrode equation: Calculate the entropy of an ideal gas of noble gas atoms; choose your own density, energy per particle, and particle mass (keeping in mind that the gas is to be near ideal); use the minimal grid size in phase space, $g = g_{min} \equiv h/m$.

TEST YOUR UNDERSTANDING OF CHAPTER 3:

3.1 Geometry of phase space: Explain the concepts **phase space**, **energy surface**, **energy shell**.

3.2 Entropy and geometry: What is the relation between the entropy of a system and the geometry of the phase space?

3.3 Geometry of n -spheres: Name a property of high-dimensional spheres that simplifies the entropy of an ideal gas. (Hint: shell volume?)

Chapter 4

Statistical Thermodynamics

4.1 Microcanonical ensemble

We recall the definition of this ensemble – it is that set of microstates which for given N, V have an energy in the interval $[E, \Delta E]$. The number of such microstates is proportional to the phase space volume they inhabit. And we found some reason to suspect that this volume – its logarithm, rather – may be identified as that property which the thermodynamicists have dubbed **entropy** and denoted by $S(N, V, E)$. This can hold only if S has the two essential properties of entropy:

1. If the system is divided into two subsystems that may freely exchange energy, then the equilibrium state is the one in which the available energy $E = E_1 + E_2$ is distributed such that

$$\left. \frac{dS(N_1, V_1, E_1)}{dE_1} \right|_{E_1=E_1^*} \stackrel{?}{=} \left. \frac{dS(N_2, V_2, E_2)}{dE_2} \right|_{E_2=E-E_1^*} \quad (4.1)$$

The quantity $T \equiv [dS(N, V, E)/dE]^{-1}$ is commonly called **temperature**.

2. In equilibrium the entropy S is additive:

$$S(N_1, V_1, E_1) + S(N_2, V_2, E_2) \stackrel{?}{=} S(N_1 + N_2, V_1 + V_2, E_1 + E_2) \quad (4.2)$$

THERMAL INTERACTION

Now consider two systems (N_1, V_1, E_1) and (N_2, V_2, E_2) in **thermal contact**, meaning that they can exchange energy but keeping the total energy $E \equiv E_1 + E_2$ constant; their “private” volumes and particles remain separated. We may determine the phase space volume of the combined system.

Ad 1. The optimal partial energy E_1^* fulfills $\Gamma_2(E_2) \partial\Gamma_1(E_1)/\partial E_1 = \Gamma_1(E_1) \partial\Gamma_2(E_2)/\partial E_2$, or

$$\left. \frac{\partial}{\partial E_1} \ln \Gamma_1(E_1) \right|_{E_1=E_1^*} = \left. \frac{\partial}{\partial E_2} \ln \Gamma_2(E_2) \right|_{E_2=E-E_1^*} \quad (4.3)$$

This, however, is nothing else but the well-known thermodynamic equation

$$\left. \frac{\partial S_1(E_1)}{\partial E_1} \right|_{E_1=E_1^*} = \left. \frac{\partial S_2(E_2)}{\partial E_2} \right|_{E_2=E-E_1^*} \quad (4.4)$$

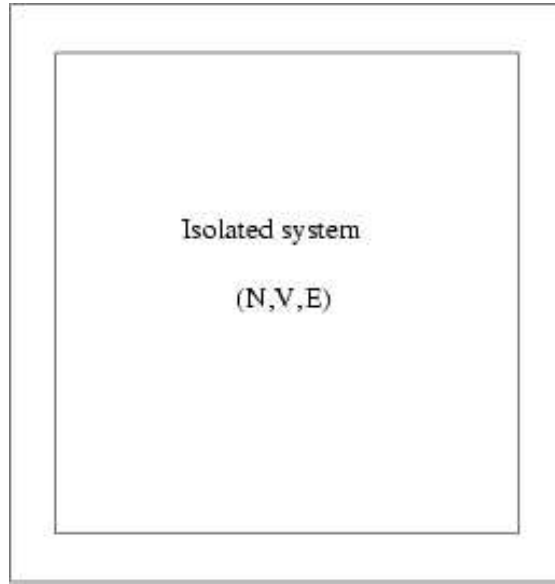


Figure 4.1: Isolated system: \rightarrow microcanonical ensemble

or, with $\partial S/\partial E \equiv 1/T$,

$$T_1 = T_2. \quad (4.5)$$

Ad 2. Let the total energy be divided up according to $E \equiv E_1 + E_2$. Then we have, by equ. 3.30,

$$\Gamma_{1+2}(E) = \Gamma_1(E_1^*) \Gamma_2(E - E_1^*) \quad (4.6)$$

and thus

$$\begin{aligned} S_{1+2} &= k \ln \frac{\Gamma_{1+2}(E)}{g^{3N}} = k \ln \left[\frac{\Gamma_1(E_1^*)}{g^{3N_1}} \frac{\Gamma_2(E - E_1^*)}{g^{3N_2}} \right] \\ &= k \ln \frac{\Gamma_1(E_1^*)}{g^{3N_1}} + k \ln \frac{\Gamma_2(E - E_1^*)}{g^{3N_2}} \\ &= S_1(E_1^*) + S_2(E - E_1^*) \end{aligned} \quad (4.7)$$

where E_1^* is that partial energy of system 1 which maximizes the product $\Gamma_1(E_1)\Gamma_2(E - E_1)$.

In other words, at thermal contact between two systems isolated from the outside world there will be a regular flow of energy until the quantity $\partial S/\partial E$ ($\equiv 1/T$) is equal in both systems. Since the combined system has then the largest extension in phase space, this will be the most probable distribution of energies upon the two systems. There may be fluctuations around the optimal energy distribution, but due to the extreme sharpness of the maximum of $\Gamma_1(E_1)\Gamma_2(E - E_1)$ these deviations remain very small.

It should be noted that these conclusions, although of eminent physical significance, may be derived quite simply from the geometrical properties of high-dimensional spheres.

Example: Consider two systems with $N_1 = N_2 = 36$ and initial energies $E_1 = 36$, $E_2 = 72$. Now bring the systems in thermal contact. The maximum value of the product $\Gamma_1(E_1)\Gamma_2(E - E_1)$ occurs at $E_1^* = E_2^* = 54$, and the respective phase space volume is

$$\Gamma_1(E_1^*)\Gamma_2(108 - E_1^*) = \left[\frac{1}{36} (\pi e 10^4)^{3/2} \right]^{72} \quad (4.8)$$

How does another partitioning of the total energy – say, (46, 62) instead of (54, 54) – compare to the optimal one, in terms of phase space volume and thus probability?

$$\frac{\Gamma_1(46)\Gamma_2(62)}{\Gamma_1(54)\Gamma_2(54)} = \left(\frac{46 \cdot 62}{54 \cdot 54}\right)^{54} = 0.30 \quad (4.9)$$

We can see that the energy fluctuations in these small systems are relatively large: $\delta E_1/E_1^* \approx 8/54 = 0.15$. However, for larger particle numbers $\delta E_1/E_1^*$ decreases as $1/\sqrt{N}$: $(515 \cdot 565/540 \cdot 540)^{540} = 0.31$; thus $\delta E_1/E_1^* \approx 25/540 = 0.05$.

THERMODYNAMICS IN THE MICROCANONICAL ENSEMBLE

Let the system under consideration be in mechanical or thermal contact with other systems. The macroscopic conditions (V, E) may then undergo changes, but we assume that this happens in a **quasistatic** way, meaning that the changes are slow enough to permit the system always to effectively perambulate the microensemble pertaining to the momentary macroconditions. To take the N -particle gas as an example, we require that its energy and volume change so slowly that the system may visit all regions of the the phase space shell $(E(t), V(t))$ before it moves to a new shell. Under these conditions the imported or exported differential energy dE is related to the differential volume change according to

$$dS = \left(\frac{\partial S}{\partial E}\right)_V dE + \left(\frac{\partial S}{\partial V}\right)_E dV \quad (4.10)$$

Defining (in addition to $T \equiv (\partial S/\partial E)_V^{-1}$) the pressure by

$$P \equiv T \left(\frac{\partial S}{\partial V}\right)_E \quad (4.11)$$

then equ. 4.10 is identical to the thermodynamic relation

$$dS = \frac{1}{T}(dE + P dV) \quad \text{or} \quad dE = T dS - P dV \quad \text{(First Law)} \quad (4.12)$$

Example: Classical ideal gas with

$$S(N, V, E) = Nk \ln \left[\frac{V}{N} \left(\frac{4\pi E e}{3Nmg^2} \right)^{3/2} \right] \quad (4.13)$$

Solving this equation for E we find for the **internal energy** $U(S, V) \equiv E$ the explicit formula

$$U(S, V) = \left(\frac{N}{V}\right)^{2/3} \frac{3Nmg^2}{4\pi} e^{2S/3Nk-1} \quad (4.14)$$

From thermodynamics we know that $T = (\partial U/\partial S)_V$; therefore

$$T = \frac{2}{3} \frac{U}{Nk} \quad (4.15)$$

From this we conclude, in agreement with experiment, that the **specific heat** of the ideal gas is

$$C_V = \frac{3}{2} Nk. \quad (4.16)$$

The **pressure** may be found from $P = -(\partial U/\partial V)_S$:

$$P = \frac{2}{3} \frac{U}{V} = \frac{NkT}{V} \quad (4.17)$$

SIMULATION IN THE MICROCANONICAL ENSEMBLE: MOLECULAR DYNAMICS

We found it a simple matter to derive the entropy S of an ideal gas as a function of N , V and E ; and once we had $S(N, V, E)$ the subsequent derivation of thermodynamics was easy. To keep things so simple we had to do some cheating, in that we assumed no interaction between the particles. Statistical mechanics is powerful enough to yield solutions even if there are such interactions. A discussion of the pertinent methods – virial expansion, integral equation theories etc. – is beyond the scope of this tutorial. However, there is a more pragmatic method of investigating the thermodynamic properties of an arbitrary model system: computer simulation. The classical equations of motion of mass points interacting via a physically plausible pair potential $u(r)$ such as the one introduced by Lennard-Jones read

$$\frac{d\vec{v}_i}{dt} = -\frac{1}{m} \sum_{j \neq i} \frac{\partial u(r_{ij})}{\partial \vec{r}_i} \quad (4.18)$$

If at some time t the microstate $\{\vec{r}_i, \vec{v}_i; i = 1, \dots, N\}$ is given we can solve these equations of motion for a short time step Δt by numerical approximation; the new positions and velocities at time $t + \Delta t$ are used as starting values for the next integration step and so forth. This procedure is known as molecular dynamics simulation.

Since we assume no external forces but only forces between the particles, the total energy of the system remains constant: the trajectory in phase space is confined to the energy surface $E = \text{const.}$ If the system is chaotic it will visit all states on this hypersurface with the same frequency. An average over the trajectory is therefore equivalent to an average over the microcanonical ensemble. For example, the internal energy may be calculated according to $U_i \equiv \langle E_{kin} \rangle + \langle E_{pot} \rangle$, where E_{kin} may be determined at any time from the particle velocities, and $E_{pot} \equiv (1/2) \sum_{i,j} u(r_{ij})$ from the positions. By the same token the temperature may be computed via $3NkT/2 \equiv \langle E_{kin} \rangle$, while the pressure is the average of the so-called “virial”; that is the quantity $W \equiv (1/2) \sum_i \sum_{j \neq i} \vec{K}_{ij} \cdot \vec{r}_{ij}$. In particular we have $P = NkT/V + \langle W \rangle / 3V$.

In the case of hard spheres the particle trajectories are computed in a different manner. For given \vec{r}_i, \vec{v}_i the time span t_0 to the next collision between any two particles in the system is determined. Calling these prospective collision partners i_0 and j_0 we first move all spheres along their specific flight directions by $\Delta \vec{r}_i = \vec{v}_i t_0$ and then simulate the collision (i_0, j_0) , computing the new directions and speeds of the two partners according to the laws of elastic collisions. Now we have gone full circle and can do the next t_0 and i_0, j_0 .

Further details of the MD method may be found in [VESELY 94] or [ALLEN 90]

4.2 Canonical ensemble

We once more put two systems in thermal contact with each other. One of the systems is supposed to have many more degrees of freedom than the other:

$$n > n - n_1 \gg n_1 \gg 1 \quad (4.19)$$

The larger system, with $n_2 \equiv n - n_1$ d.o.f., is called “heat bath”. The energy $E_2 = E - E_1$ contained in the heat bath is “almost always” much greater than the energy of the smaller system; the heat reservoir’s entropy may therefore be expanded around $S_2(n_2, E)$:

$$S_2(E - E_1) \approx S_2(E) - E_1 \left. \frac{\partial S_2(E')}{\partial E'} \right|_{E'=E} = S_2(E) - \frac{E_1}{T} \quad (4.20)$$

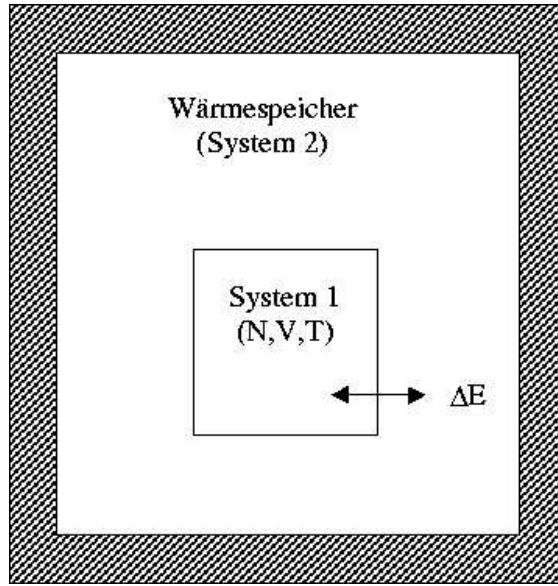


Figure 4.2: System in contact with an energy reservoir: \rightarrow canonical ensemble

where T is the temperature of the heat bath. The number of phase space cells occupied by the larger system is thus

$$\Sigma_2(E - E_1) = e^{S_2/k} e^{-E_1/kT} \quad (4.21)$$

But the larger $\Sigma_2(E - E_1)$, the larger the probability to find system 1 in a microstate with energy E_1 . We may express this in terms of the “canonical phase space density” $\rho_1(\Gamma_1)$:

$$\rho_1[\Gamma_1(E_1)] = \rho_0 e^{-E_1/kT} \quad (4.22)$$

where ρ_0 is a normalizing factor, and

$$\rho_1(\Gamma_1) \equiv \rho[\vec{r}_1, \vec{v}_1; E(\vec{r}_1, \vec{v}_1) = E_1] \quad (4.23)$$

is the density of microstates in that region of phase space of system 1 that belongs to energy E_1 .

For a better understanding of equ. 4.22 we recall that in the microcanonical ensemble only those states of system 1 were considered for which the energy E_1 was in the interval $[E_1, E_1 - \Delta E]$. In the canonical ensemble **all** energy values are permitted, but the density of state points varies strongly, as $\exp[-E_1/kT]$.

Equation 4.22 may not be understood to say that the **most probable** energy of the smaller system be equal to zero. While the **density** of states in the phase space of system 1 indeed drops sharply with increasing E_1 , the **volume** $\Gamma_1(E_1)$ pertaining to E_1 is strongly increasing as $E_1^{3N/2}$. The product of these two factors, i.e. the **statistical weight** of the respective phase space region, then exhibits a maximum at an energy $E_1 \neq 0$.

As a – by now familiar – illustration of this let us recall the Maxwell-Boltzmann distribution: $p(|\vec{v}|)$ is just the probability density for the (kinetic) energy of a subsystem consisting of only one particle, while the heat bath is made up of the $N - 1$ other molecules.

In the more general case, i.e. for a large number of particles, the peak of the energy distribution is so sharp that the most probable energy is all but identical to the mean energy:

$$\frac{\Delta_E}{E} \propto \frac{1}{\sqrt{N}} \quad (4.24)$$

Thus we have found that even non-isolated systems which may exchange energy have actually most of the time a certain energy from which they will deviate only slightly. But this means that

we may calculate averages of physical quantities either in the microcanonical or in the canonical ensemble, according to mathematical convenience. This principle is known as “equivalence of ensembles”.

We have derived the properties of the canonical ensemble using a Taylor expansion of the entropy. The derivation originally given by Gibbs is different. J. W. Gibbs generalized Boltzmann’s “method of the most probable distribution” to an ensemble of microscopically identical systems which are in thermal contact with each other.

Gibbs considered M equal systems ($i = 1, \dots, M$), each containing N particles. The sum of the M energies was constrained to sum up to a given value, $\mathcal{E} = \sum_i E_i$, with an unhindered interchange of energies between the systems. Under these simple assumptions he determined the probability of finding a system in the neighbourhood of a microstate $\vec{\Gamma}$ having an energy in the interval $[E, \Delta E]$. With increasing energy E this probability density drops as $\rho(\Gamma(E)) \propto \exp[-E/kT]$. Since the **volume** of the energy shell rises sharply with energy we again find that most systems will have an energy around $\langle E \rangle = \mathcal{E}/M$. Thus the important **equivalence of canonical and microcanonical ensembles** may alternatively be proven in this manner.

THERMODYNAMICS IN THE CANONICAL ENSEMBLE

The quantity

$$Q(N, V, T) \equiv \frac{1}{N!g^{3N}} \int d\vec{r} d\vec{v} e^{-E(\vec{r}, \vec{v})/kT} \quad (4.25)$$

is called **canonical partition function**. First of all, it is a normalizing factor in the calculation of averages over the canonical ensemble. For example, the internal energy may be written as

$$U \equiv \langle E \rangle = \frac{1}{Q(N, V, T)} \frac{1}{N!g^{3N}} \int d\vec{r} d\vec{v} E(\vec{r}, \vec{v}) e^{-E(\vec{r}, \vec{v})/kT} \quad (4.26)$$

But the great practical importance of the partition function stems from its close relation to Helmholtz’ free energy $A(N, V, T)$, which itself is a central object of thermodynamics. The relation between the two is

$$Q(N, V, T) = e^{-\beta A(N, V, T)} \quad (4.27)$$

where $\beta \equiv 1/kT$. To prove this important fact we differentiate the identity

$$\frac{1}{N!g^{3N}} \int d\vec{r} d\vec{v} e^{\beta[A(N, V, T) - E(\vec{r}, \vec{v})]} = 1 \quad (4.28)$$

by β , obtaining

$$A(N, V, T) - \langle E \rangle + \beta \left(\frac{\partial A}{\partial \beta} \right)_V = 0 \quad (4.29)$$

or

$$A(N, V, T) - U(N, V, T) - T \left(\frac{\partial A}{\partial T} \right)_V = 0 \quad (4.30)$$

But this is, with $S \equiv -(\partial A/\partial T)_V$, identical to the basic thermodynamic relation $A = U - TS$.

All other thermodynamic quantities may now be distilled from $A(N, V, T)$. For instance, the pressure is given by

$$P = - \left(\frac{\partial A}{\partial V} \right)_T, \quad (4.31)$$

Similarly, entropy and Gibbs’ free energy are calculated from

$$S = - \left(\frac{\partial A}{\partial T} \right)_V \quad \text{and} \quad G = A + PV \quad (4.32)$$

Example 1: For the **classical ideal gas** we have

$$Q(N, V, T) = \frac{m^{3N}}{N!h^{3N}} \int d\vec{r}d\vec{v} \exp\left[-\frac{m}{2kT} \sum |\vec{v}_i|^2\right] = \frac{1}{N!h^{3N}} q^N \quad (4.33)$$

with

$$q \equiv \int d\vec{r}d\vec{v} \exp\left[-\frac{m}{2kT} |\vec{v}|^2\right] = \left(\frac{2\pi kT}{m}\right)^{3/2} V \quad (4.34)$$

This leads to

$$Q(N, V, T) = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2}\right)^{3N/2} V^N \quad (4.35)$$

and using $A \equiv -kT \ln Q$ and 4.31 we find

$$P = \frac{NkT}{V} \quad (4.36)$$

which is the well-known equation of state of the ideal gas. Similarly we find from $S = -(\partial A/\partial T)_V$ the entropy in keeping with equ. 3.39.

Example 2: The free energy of one mole of an ideal gas at standard conditions, assuming a molecular mass of $m = 39.95m_H$ (i. e. Argon), is

$$A = -kT \ln Q = -4.25 \cdot 10^7 J \quad (4.37)$$

We have now succeeded to derive the thermodynamics of an ideal gas solely from a geometrical analysis of the phase space of N classical point masses. It must be stressed that similar relations for thermodynamical observables may also be derived for other model systems with their respective phase spaces.

In hindsight it is possible to apply the concept of a “partition function” also to the microcanonical ensemble. After all, the quantity $\Sigma(N, V, E)$ was also a measure of the total accessible phase space volume. In other words, we might as well call it the “microcanonical partition function”. And we recall that its logarithm – the entropy – served as a starting point to unfold statistical thermodynamics.

EQUIPARTITION THEOREM

Without proof we note the following important theorem:

If the Hamiltonian of a system contains some position or velocity coordinate *in quadratic form*, the respective degree of freedom will have the mean energy $kT/2$.

Example 1: The Hamiltonian of the classical ideal gas is

$$H(\vec{v}) = \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{mv_{i\alpha}^2}{2} \quad (4.38)$$

Each of the $3N$ translational d.o.f. appears in the guise $v_{i\alpha}^2$. The equipartition theorem then tells us that for each velocity component

$$\frac{m}{2} \langle v_{i\alpha}^2 \rangle = \frac{kT}{2} \quad \text{or} \quad \frac{m}{2} \langle v_i^2 \rangle = \frac{3kT}{2} \quad (4.39)$$

(As there is no interactional or external potential, the positional d.o.f. contain no energy.)

Example 2: Every classical fluid, such as the Lennard-Jones liquid, has the Hamiltonian

$$H(\vec{r}, \vec{v}) = H_{pot}(\vec{r}) + \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{mv_{i\alpha}^2}{2} \quad (4.40)$$

Therefore each of the $3N$ **translatory** d.o.f. contains, on the average, an energy $kT/2$. (The interaction potential is not quadratic in the positions; therefore the equipartition law does not apply to $r_{i\alpha}$.)

Example 3: A system of N one-dimensional harmonic oscillators is characterized by the Hamiltonian

$$H(\vec{x}, \dot{\vec{x}}) = H_{pot} + H_{kin} = \frac{f}{2} \sum_{i=1}^N x_i^2 + \frac{m}{2} \sum_{i=1}^N \dot{x}_i^2 \quad (4.41)$$

Therefore,

$$\frac{f}{2} \langle x_i^2 \rangle = \frac{kT}{2} \quad \text{and} \quad \frac{m}{2} \langle \dot{x}_i^2 \rangle = \frac{kT}{2}. \quad (4.42)$$

Thus the total energy of the N oscillators is $E = NkT$. The generalization of this to three dimensions is trivial; we find $E_{id.cryst.} = 3NkT$. For the specific heat we have consequently $C_V \equiv (\partial E / \partial T)_V = 3Nk$. This prediction is in good agreement with experimental results for crystalline solids at moderately high temperatures. For low temperatures – and depending on the specific substance this may well mean room temperature – the classical description breaks down, and we have to apply quantum rules to predict C_V and other thermodynamic observables (see Chapter 5.)

Example 4: A fluid of N “dumbbell molecules”, each having 3 translatory and 2 rotatory d.o.f., has the Hamiltonian

$$H(\vec{r}, \vec{e}, \dot{\vec{r}}, \dot{\vec{e}}) = H_{pot}(\vec{r}, \vec{e}) + \sum_{i,\alpha} \frac{mv_{i\alpha}^2}{2} + \sum_{i,\beta} \frac{I\omega_{i\beta}^2}{2} \quad (4.43)$$

where \vec{e} are the orientation vectors of the linear particles, ω are their angular velocities, and I denotes the molecular moment of inertia.

Thus we predict

$$\frac{m}{2} \langle v_i^2 \rangle = \frac{3kT}{2} \quad \text{and} \quad \frac{I}{2} \langle \omega_i^2 \rangle = \frac{2kT}{2}. \quad (4.44)$$

which is again a good estimate as long as a classical description may be expected to apply.

CHEMICAL POTENTIAL

Let us recall the experimental setup of Fig. 4.2. Allowing two systems to exchange energy leads to an equilibration of their temperatures T (see Section 4.1). The energies E_1, E_2 of the two subsystems will then only fluctuate – usually just slightly – around their average values.

Let us now assume that the systems can also exchange particles. In such a situation we will again find some initial equilibration after which the particle numbers N_1, N_2 will only slightly fluctuate around their mean values. The flow of particles from one subsystem to the other will come to an end as soon as the free energy $A = A_1 + A_2$ of the combined system tends to a

constant value. This happens when $(\partial A_1/\partial N)_V = (\partial A_2/\partial N)_V$. The quantity that determines the equilibrium is therefore $\mu \equiv (\partial A/\partial N)_V$; the subsystems will trade particles until $\mu_1 = \mu_2$. (Compare this definition of the **chemical potential** with that of temperature, $(\partial S/\partial E)_V^{-1}$.)

Example: The chemical potential of Argon ($m = 39.95m_H$) at normal conditions is

$$\mu \equiv \frac{dA}{dN} = -kT \frac{d \ln Q}{dN} = -kT \ln \left[\frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \right] = -7.26 \cdot 10^{-20} J \quad (4.45)$$

SIMULATION IN THE CANONICAL ENSEMBLE: MONTE CARLO CALCULATION

We may once more find an appropriate numerical rule for the correct “browsing” of all states with given (N, V, T) (in place of (N, V, E)). Since it is now the Boltzmann factor that gives the statistical weight of a microstate $\vec{\Gamma} \equiv \{\vec{r}_i, \vec{v}_i\}$, the average value of some quantity $a(\vec{\Gamma})$ is given by

$$\langle a \rangle_{N,V,T} = \int a(\vec{\Gamma}) \exp[-E(\vec{\Gamma})/kT] d\vec{\Gamma} / \int \exp[-E(\vec{\Gamma})/kT] d\vec{\Gamma} \quad (4.46)$$

If a depends only on the particle positions and not on velocities – and this is true for important quantities such potential energy, virial, etc. – then we may even confine the weighted integral to the $3N$ -dimensional configurational subspace $\vec{\Gamma}_c \equiv \{\vec{r}_i\}$ of full phase space:

$$\langle a \rangle_{N,V,T} = \int a(\vec{\Gamma}_c) \exp[-E_{pot}(\vec{\Gamma}_c)/kT] d\vec{\Gamma}_c / \int \exp[-E_{pot}(\vec{\Gamma}_c)/kT] d\vec{\Gamma}_c \quad (4.47)$$

In order to find an estimate for $\langle a \rangle$ we formally replace the integrals by sums and construct a long sequence of randomly sampled states $\vec{\Gamma}_c(m); m = 1, \dots, M$ with the requirement that the relative frequency of microstates in the sequence be proportional to their Boltzmann factors. In other words, configurations with high potential energy should occur less frequently than states with small E_{pot} . The customary method to produce such a biased random sequence (a so-called “Markov chain”) is called “Metropolis technique”, after its inventor Nicholas Metropolis.

Now, since the Boltzmann factor is already contained in the frequency of states in the sequence we can compute the Boltzmann-weighted average simply according to

$$\langle a \rangle = \frac{1}{M} \sum_{m=1}^M a(\vec{\Gamma}_c(m)) \quad (4.48)$$

An extensive description of the Monte Carlo method may be found in [Vesely 1978] or [Vesely 2001].

There is also a special version of the molecular dynamics simulation method that may be used to perambulate the canonical distribution. We recall that in MD simulations normally the total system energy $E_{tot} = E_{kin} + E_{pot}$ is held constant, which is roughly equivalent to sampling the microcanonical ensemble. However, by introducing a kind of numerical thermostat we may at each time step adjust all particle velocities so as to keep E_{kin} either constant (isokinetic MD simulation) oder near a mean value such that $\langle E_{kin} \rangle \propto T$ (isothermal MD).

4.3 Grand canonical ensemble

Once again we put a small system (1) in contact with a large one (2). However, this time we do not only permit the exchange of energy but also the crossing over of particles from one subsystem

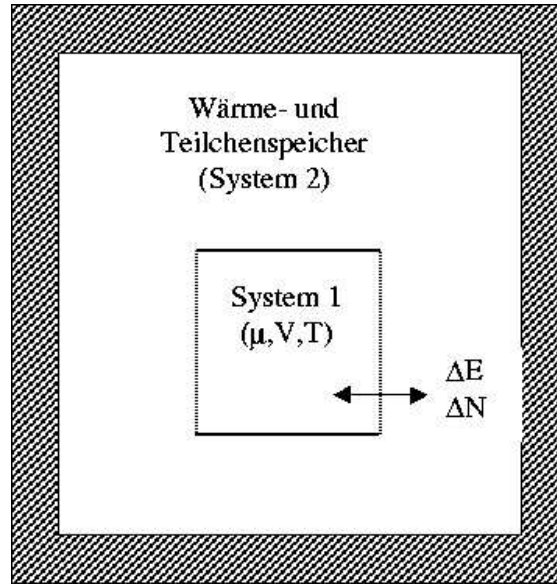


Figure 4.3: System in contact with an energy and particle reservoir: \rightarrow grand canonical ensemble to the other.

And as before we can write down the probability density in the phase space of the smaller system; it depends now both on the number of particles N_1 and on $\{\vec{r}_i, \vec{v}_i; i = 1, \dots, N_1\}$, as follows:

$$p(\vec{r}, \vec{v}; N_1) \propto e^{\mu N_1/kT} e^{-E(\vec{r}, \vec{v})/kT} \quad (4.49)$$

Summing this density over all possible values of N_1 and integrating – at each N_1 – over all $\{\vec{r}_i, \vec{v}_i; i = 1, \dots, N_1\}$ we obtain the **grand partition function**

$$Z(\mu, V_1, T) \equiv \sum_{N_1=0}^{\infty} e^{N_1 \mu/kT} Q(N_1, V_1, T) \quad (4.50)$$

Its value is just the “total statistical weight” of all possible states of system 1. Above all, it serves as the source function of thermodynamics.

THERMODYNAMICS IN THE GRAND CANONICAL ENSEMBLE

From the grand partition function we can easily derive expressions for the various thermodynamic observables. For instance, putting $z \equiv e^{\mu/kT}$ and $\beta \equiv 1/kT$ we find

$$P = \frac{kT}{V} \ln Z(z, V, T) \quad (4.51)$$

$$N(\equiv \langle N \rangle) = z \frac{\partial}{\partial z} \ln Z(z, V, T) = kT \frac{\partial \ln Z}{\partial \mu} \quad (4.52)$$

$$U(\equiv \langle E \rangle) = -\frac{\partial}{\partial \beta} \ln Z(z, V, T) = kT^2 \frac{\partial \ln Z}{\partial T} \quad (4.53)$$

As a rule the – permitted – fluctuations of the number of particles remain small; in particular we have $\Delta N/N \approx 1/\sqrt{N}$. Thus the grand ensemble is again equivalent to others ensembles of statistical mechanics.

[To do: applet with MD simulation, averages taken only over particles in a partial volume – i

same results!]

Example: Let us visit the ideal gas again. For the grand partition function we have

$$\begin{aligned} Z(z, V, T) &= \sum_N z^N \frac{V^N}{N!} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2} \\ &= \sum_N \frac{y^N}{N!} \text{ with } y \equiv Vz \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \end{aligned} \quad (4.54)$$

Therefore

$$Z = \exp \left[-zV \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \right] \text{ or } \ln Z = -zV \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \quad (4.55)$$

Using the formulae for internal energy and pressure we find

$$P = -kTz \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \text{ and } U = -kTz \frac{3V}{2} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \quad (4.56)$$

Consequently, $P = 2U/3V$ or

$$P = \frac{2}{3V} \frac{3NkT}{2} = \frac{N}{V} kT \quad (4.57)$$

in keeping with the phenomenological ideal gas equation.

SIMULATION IN THE GRAND CANONICAL ENSEMBLE: GCMC

The states within the grand ensemble may again be sampled in a random manner. Just as in the canonical Monte Carlo procedure we produce a sequence of microstates $\vec{\Gamma}_c(m)$, $m = 1, \dots, M$ with the appropriate relative frequencies. The only change is that we now vary also the number of particles by occasionally adding or removing a particle. The stochastic rule for these insertions and removals is such that it agrees with the thermodynamic probability of such processes. By averaging some quantity over the “Markov chain” of configurations we again obtain an estimate of the respective thermodynamic observable.

4.4 Problems for Chapter 4

EXERCISES:

4.1 Canonical phase space density: Considering the Maxwell-Boltzmann distribution in a classical 3D fluid, show that

a) with increasing kinetic energy E_k the density of states decreases as $\exp[-E_k/kT]$;

b) in spite of this the most probable kinetic energy is not zero.

Apply this to the general **canonical phase space**.

4.2 Grand canonical ensemble: Using the grand partition function, derive the mean number

of particles in one m^3 of an ideal gas at standard conditions.

TEST YOUR UNDERSTANDING OF CHAPTER 4:

4.1 Temperature and entropy: What happens if two systems are put in contact with each other such that they may exchange energy?

4.2 Changes of state: What is a quasistatic change of state?

4.3 Equipartition theorem: Formulate the equipartition theorem; to which microscopic variables does it apply? Give two examples.

4.4 Canonical ensemble: What are the macroscopic conditions that define the canonical ensemble? When is it equivalent to the microcanonical ensemble?

4.5 Grand canonical ensemble: What are the macroscopic conditions defining the grand canonical ensemble? What is needed to make it equivalent to the canonical and microcanonical ensembles?

Chapter 5

Statistical Quantum Mechanics

Counting the states in phase space takes some consideration if we want to apply quantum mechanics. We have already mentioned the fact that quantum particles are not distinguishable, and we took this into account in a heuristic manner, by introducing the permutation factor $1/N!$ in the partition function.

Another feature of the quantum picture is that state space is not continuous but consists of finite “raster elements”: just think of the discrete states of an ideal quantum gas. Finally, it depends on the “symmetry class” of the particles how many of them may inhabit the same discrete microstate. **Fermions**, with a wave function of odd symmetry, can take on a particular state only exclusively; the population number of a raster element in phase space can be just 0 or 1. In contrast, even-symmetry **bosons** may in any number share the same microstate.

What are the consequences of these additional counting rules for Statistical Thermodynamics? To seek an answer we may either proceed in the manner of Boltzmann (see Section 2.2) or à la Gibbs (see Chapter 4). For a better understanding we will here sketch both approaches.

5.1 Ideal quantum gas: Method of the most probable distribution

We consider a system of N independent particles in a cubic box with ideally elastic walls. In the spirit of the kinetic theory of dilute gases we explore the state space for the individual particles, which Boltzmann dubbed μ -space. In the quantum case this space is spanned by the quantum numbers $n_{x,y,z} = 1, 2, \dots$ pertaining to the momentum eigenstates having eigenvalues $\vec{p} \equiv (h/2L)\vec{n}$ and energies $E_{\vec{n}} = p^2/2m = (h^2/8mL^2)|\vec{n}|^2$.

Now we batch together all states having energies $E_{\vec{n}}$ in an interval $[E_j, \Delta E]$. The number of states in such a “cell” is named g_j , $j = 1, 2, \dots, m$. The values of the g_j are not important; they should only be large enough to allow the application of Stirling’s formula.

As before we try to answer the question how the N particles should best be distributed over the m cells. To do so we change the notation from the one used in Section 2.2, in that we denote the number of particles in cell j by f_j . The reason for using f_j is that “ n ” is reserved for the quantum numbers.

A specific distribution $\vec{f} \equiv \{f_j; j = 1, \dots, m\}$ of the N particles to the m cells is more probable if its **multiplicity** W is larger, meaning that we can allot the $\{f_j\}$ particles in more different ways to the $\{g_j\}$ states in each cell – always keeping in mind the Fermi or Bose rules:

$$\text{Fermi – Dirac : } W = \prod_{j=1}^m \binom{g_j}{f_j} \quad \text{Bose – Einstein : } W = \prod_{j=1}^m \binom{g_j + f_j - 1}{f_j} \quad (5.1)$$

To compare: the multiplicity given in Section 2.2, pertaining to a classical distribution (see equ. 2.20) would in our present notation read $W = \prod_{j=1}^m g_j^{f_j} / f_j!$.

The distribution f_j^* having the largest multiplicity may again be determined by Lagrange variation with the conditions $\sum_j f_j E_j = E$ and $\sum_j f_j = N$:

$$\delta \ln W - \beta \delta \left(\sum_j f_j E_j - E \right) + \gamma \delta \left(\sum_j f_j - N \right) = 0 \quad (5.2)$$

Writing $z \equiv e^\gamma$ we find for

$$\text{Fermi - Dirac : } f_j^* = \frac{g_j}{z^{-1} e^{\beta E_j} + 1} \quad \text{Bose - Einstein : } f_j^* = \frac{g_j}{z^{-1} e^{\beta E_j} - 1} \quad (5.3)$$

This is the most probable distribution of the particles upon the cells. Since g_j denotes the number of states in cell j , we have for the **average population number** of each state

$$\langle f_{\vec{n}} \rangle \equiv \frac{f_j^*}{g_j} = \frac{1}{z^{-1} e^{\beta E_j} \pm 1} \quad (+ \dots \text{Fermi; } - \dots \text{Bose}) \quad (5.4)$$

It is easy to interpret the Lagrange parameters β and z . As in Section 2.2 one compares the consequences of the population densities given above to empirical/thermodynamical facts, finding that β is related to temperature as $\beta = 1/kT$, and that $z = e^\gamma = e^{\mu/kT}$ is identical to the fugacity.

For a better understanding of this derivation, let us interpret its premises as a set of rules in a game of fortune, as we have done in Chapter 2. By running the applet *EFFRoulette* we may indeed play that game – for Fermi particles at least – and compare its outcome with the result just given.

These are the rules:

For non-interacting particles in a square box the μ -plane is spanned by integers n_x, n_y ; each quantum state is represented by a point. A specific state of a system of N fermions is represented by a set of N inhabited points on that plane.

To find the average (and also most probable!) distribution of particles on states,

- assign N particles randomly to the states on μ -plane
- make sure that the sum of the particle energies equals the given system energy,
- AND
- discard all trials in which a state is inhabited by more than one particle
- determine the mean number of particles in each state; sort the result according to the state energies

5.2 Ideal quantum gas: Grand canonical ensemble

We may derive the properties of a quantum gas in another way, making use of the (z, V, T) ensemble in Gibbsian phase space. Recalling the general definition of the grand partition function, $Z(z, V, T) \equiv \sum_{N=0}^{\infty} z^N Q(N, V, T)$, we now write Q as a **sum** (in place of an integral) over states:

$$Q(N, V, T) = \frac{m^{3N}}{N! h^{3N}} \int d\vec{\Gamma} \exp \left[-E(\vec{\Gamma})/kT \right] \implies \sum_{\{f_{\vec{n}}\}} \exp \left[-\sum_{\vec{n}} f_{\vec{n}} E_{\vec{n}}/kT \right] \quad (5.5)$$

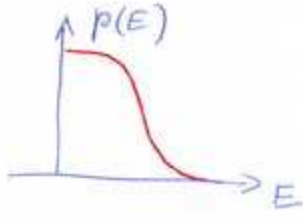


Figure 5.1: Playing the Fermi-Dirac game. [Code: EFRoulette]

The sum $\{f_{\vec{n}}\}$ is to be taken over all permitted population numbers of all states \vec{n} , again requiring that $\sum_{\vec{n}} f_{\vec{n}} = N$. The permitted values of $f_{\vec{n}}$ are: 0 and 1 for fermions, and $0, 1, 2, \dots, N$ for bosons. In this manner we arrive at

$$Z = \sum_{N=0}^{\infty} z^N \sum_{\{f_{\vec{n}}\}} \exp \left[- \sum_{\vec{n}} f_{\vec{n}} E_{\vec{n}} / kT \right] = \sum_{N=0}^{\infty} \sum_{\{f_{\vec{n}}\}} \prod_{\vec{n}} \left(z e^{-E_{\vec{n}}/kT} \right)^{f_{\vec{n}}} \quad (5.6)$$

It is easy to show that this is equal to

$$Z = \prod_{\vec{n}} \left[\sum_f \left(z e^{-E_{\vec{n}}/kT} \right)^f \right] \quad (5.7)$$

Now we can insert the possible values of f . We find that for

Fermions ($f = 0, 1$):

$$Z = \prod_{\vec{n}} \left[1 + z e^{-E_{\vec{n}}/kT} \right] \quad (5.8)$$

and for **Bosons** ($f = 0, 1, 2, \dots$):

$$Z = \prod_{\vec{n}} \frac{1}{1 - z e^{-E_{\vec{n}}/kT}} \quad (5.9)$$

Having secured the grand partition function, we can now apply the well-known formulae for pressure, mean particle number, and internal energy (see Section 4.3) to determine the thermodynamic properties of the system. The mean population number of a given state \vec{n} is

$$\langle f_{\vec{n}} \rangle \equiv \frac{1}{Z} \sum_N \sum_{\{f_{\vec{n}}\}} f_{\vec{n}} \exp \left[- \sum_{\vec{n}} E_{\vec{n}} f_{\vec{n}} \right] = - \frac{1}{\beta} \frac{\partial}{\partial E_{\vec{n}}} \ln Z \quad (5.10)$$

Inserting for Z the respective expression for Fermi or Bose particles we once more arrive at the population densities of equ. 5.4.

In the following sections we discuss the properties of a few particularly prominent fermion and boson gases.

5.3 Ideal Fermi gas

Figure 5.2 shows the fermion population density (see equ. 5.4 with the positive sign in the denominator.) It is evidently quite different from its classical counterpart, the Boltzmann factor. In particular, when the temperature is low then it approaches a step function: in the limit $kT \rightarrow 0$

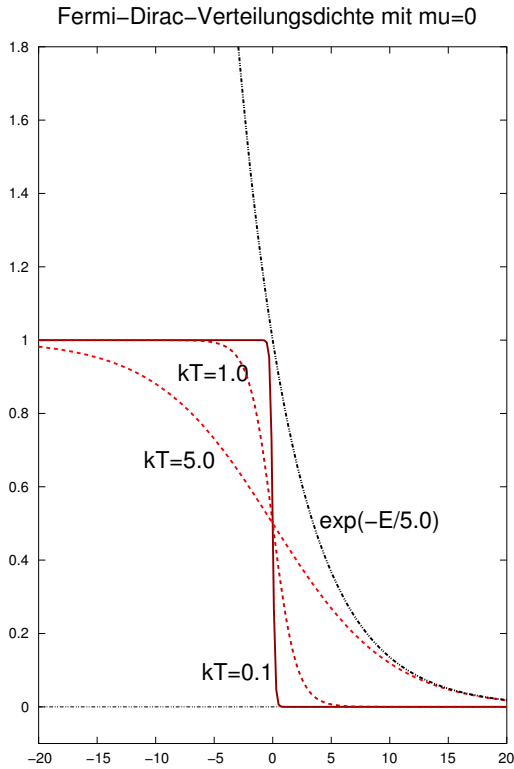


Figure 5.2: Mean population numbers of the states in a Fermi-Dirac system with $\mu = 0$. For comparison we have also drawn the classical Boltzmann factor $\exp[-E/kT]$ for a temperature of $kT = 5$.

all states having an energy below a certain value are inhabited, while the states with higher energies remain empty. The threshold energy $E_F = \mu$ is called “Fermi energy”.

It is only for high temperatures and small values of μ that the Fermi-Dirac density approaches the classical Boltzmann density; and as it does so it strictly obeys the inequality $f_{FD} < f_{Bm}$. For the purpose of comparison we have included a graph of $f_{Bm} = \exp[-E_{\vec{p}}/kT]$ in Figure 5.2.

ELECTRONS IN METALS

Conduction electrons in metallic solids may be considered as an ideal fermion gas – with an additional feature: since electrons have two possible spin states the maximum number of particles in a state \vec{p} is 2 instead of 1.

At low temperatures all states \vec{p} with $E_{\vec{p}} \leq E_F = \mu$ are populated. The number of such states is, as we can see from eq. 1.21,

$$N = \frac{8\pi}{3} V \left(\frac{2m\mu}{h^2} \right)^{3/2} \quad (5.11)$$

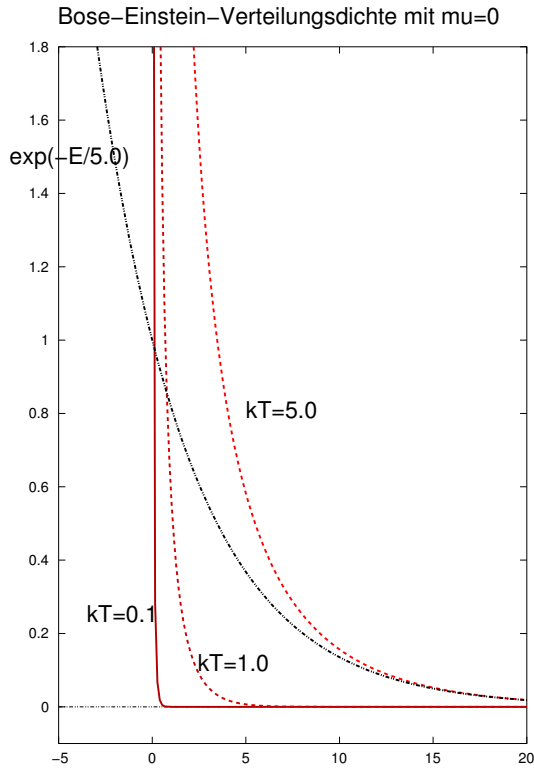


Figure 5.3: Mean population number of states in a Bose-Einstein system with $\mu = 0$. For comparison we include a graph of the classical density $\exp[-E/kT]$ at $kT = 5$.

Assuming a value of $\mu \approx 5 \cdot 10^{-19} J$ typical for conduction electrons we find $N/V \approx 3 \cdot 10^{27} m^{-3}$.

Applying the general formulae 4.51 and 4.53 we may easily derive the pressure and internal energy of the electron gas. We may also recapture the relation $PV = 2U/3$, consistent with an earlier result for the classical ideal gas.

5.4 Ideal Bose gas

For bosons the mean population number of a state is

$$\langle f_{\vec{p}} \rangle \equiv \frac{f_i^*}{g_i} = \frac{1}{z^{-1} e^{\beta E_i} - 1} \quad (5.12)$$

This function looks a bit like the Boltzmann factor $\propto e^{-E_{\vec{p}}/kT}$ but is everywhere larger than the latter. For small μ and large T we again find that $f_B \approx f_{Bm}$.

PHOTONS IN A BOX

A popular example for this type of system is a “gas” of photons. A box with an internal coating

of reflecting material in which a photon gas is held in thermodynamic equilibrium is often called a “black body”. Originally this name refers to a grain of black material that might be placed in the box to allow for absorption and re-emission of photons, thus enabling energy exchange and equilibration. In actual fact the internal walls of the box are never perfectly reflecting, rendering the insertion of a black body grain unnecessary.

Since the total energy in the box is conserved and the photons may change their energy upon absorption and reemission, the total number of photons is **not conserved**. This is tantamount to assuming $\mu = 0$.

To determine the energy spectrum of the photons we first calculate the number $d\Sigma$ of states within a small energy interval. These states will then be populated according to equ. 5.12. A simple geometric consideration – how many lattice points are lying within a spherical shell – leads to

$$d\Sigma = \frac{8\pi V p^2 dp}{h^3} = \frac{8\pi V \nu^2 d\nu}{c^3} \quad (5.13)$$

where $p = h\nu/c$ is the momentum pertaining to the energy $E = h\nu$. The total number of photons in the system is thus

$$N = \int \langle f_{\vec{p}} \rangle d\Sigma = \frac{8\pi V}{c^3} \int \frac{\nu^2}{e^{h\nu/kT} - 1} d\nu \quad (5.14)$$

and accordingly the number of photons in a frequency interval $[\nu, d\nu]$ is

$$n(\nu)d\nu = \frac{8\pi V}{c^3} \frac{\nu^2}{e^{h\nu/kT} - 1} d\nu \quad (5.15)$$

The amount of energy carried by these photons is the **spectral density** of black body radiation; it is given by

$$I(\nu)d\nu \equiv \frac{dE(\nu)}{d\nu} = E(\nu)n(\nu)d\nu = \frac{8\pi V}{c^3} \frac{h\nu^3}{e^{h\nu/kT} - 1} d\nu \quad (5.16)$$

For pressure and internal energy of the photon gas we find – in contrast to the classical and Fermi cases – the relation $PV = U/3$.

Providing an explanation for the experimentally measured spectrum of black body radiation was one of the decisive achievements in theoretical physics around 1900. Earlier attempts based on classical assumptions had failed, and it was Max Planck who, by the ad hoc assumption of a quantization of energy, could reproduce the correct shape of the spectrum. The subsequent efforts to understand the physical implications of that assumption eventually lead up to quantum physics.

5.5 Problems for Chapter 5

TEST YOUR UNDERSTANDING OF CHAPTER 5:

5.1 Population densities: Explain the significance of the population densities for Fermi-Dirac and Bose-Einstein statistics. Sketch these functions of energy at two temperatures.

5.2 Classical limit: Starting from the Fermi-Dirac density discuss the classical limit of quantum statistics.

5.3 Ideal quantum gases: Discuss two examples of ideal quantum gases.

Bibliography

- [ABRAMOWITZ 65] Abramowitz, M., Stegun, I. A., Hrsg.: Handbook of Mathematical Functions. Dover, New York, 1965.
- [ALDER 57] Alder, B. J. und Wainwright, T. E., J. Chem. Phys. 27 (1957) 1208.
- [ALDER 67] Alder, B. J., und Wainwright, T. E., Phys. Rev. Lett. 18 (1967) 988.
- [ALLEN 90] Allen, M.P., und Tildesley, D.J.: Computer Simulation of Liquids. Oxford University Press, 1990.
- [BINDER 87] Binder, K.: Applications of the Monte Carlo Method in Statistical Physics. Springer, Berlin 1987.
- [BINDER 92] Binder, K.: The Monte Carlo Method in Condensed Matter Physics. Springer, Berlin 1992.
- [HANSEN 69] Hansen, J.-P., und Verlet, L., Phys. Rev. 184 (1969) 151.
- [HOOVER 68] Hoover, W. G., und Ree, F. H., J. Chem. Phys. 49 (1968) 3609.
- [KALOS 86] Kalos, M. H., und Whitlock, P. A.: Monte Carlo Methods. Wiley, New York 1986.
- [LEVESQUE 69] Levesque, D., und Verlet, L., Phys. Rev. 182 (1969) 307; Phys. Rev. A2/6 (1970) 2514; - und Kärki, J., Phys. Rev. A7/5 (1973) 1690.
- [METROPOLIS 49] Metropolis, N., und Ulam, S., J. Amer. Statist. Assoc. 44 (1949) 335.
- [METROPOLIS 53] Metropolis, N. A., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., und Teller, E., J. Chem. Phys. 21 (1953) 1087.
- [PAPOULIS 81] Papoulis, Athanasios: Probability, Random Variables and Stochastic Processes. McGraw-Hill International Book Company, 1981.
- [POSCH 87] Holian, B. L., Hoover, W. G., Posch, H. A., Phys. Rev. Lett. 59/1 (1987) 10.
- [REIF 90] Reif, Frederick: Statistische Physik (Berkeley Physik Kurs 5). Vieweg, Braunschweig 1990.
- [VESELY 78] Vesely, F. J.: Computereperimente an Flüssigkeitsmodellen. Physik Verlag, Weinheim 1978.
- [VESELY 94] Vesely, F. J.: Computational Physics – An Introduction. Plenum, New York 1994; 2nd edn. Kluwer, New York 2001.