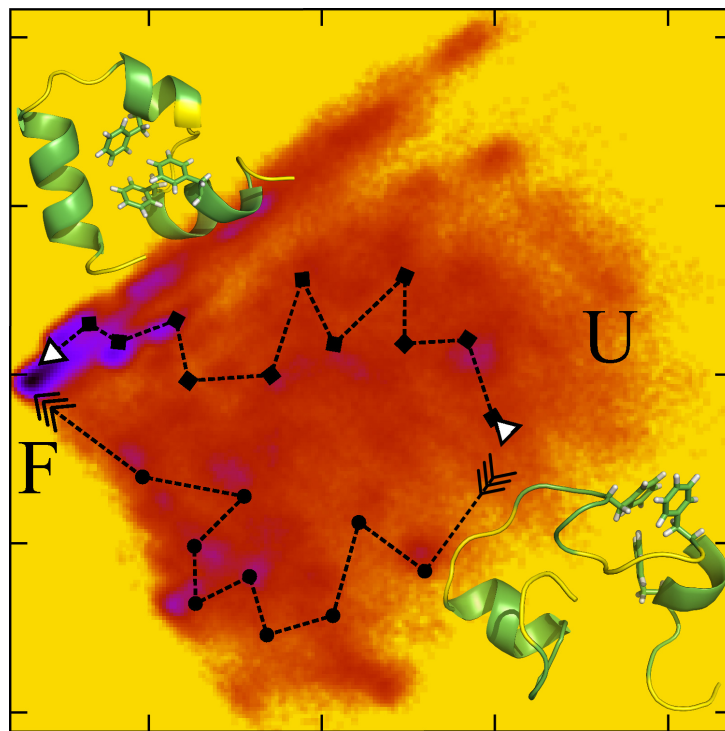# Classical Complex Systems

Professor Gerhard Stock

Biomolecular Dynamics, Institute of Physics
Albert Ludwigs University Freiburg

Winter Semester 2022/2023

Prolog

For most of the last century, the research of molecular physics focused on small molecules in the gas phase. The common thinking was that complex molecular systems such as liquids, nanostuctured materials or biomolecules can only be treated by simple empirical models that employ uncontrollable approximations. Indeed, neither straightforward quantum mechanics nor simple classical modeling is typically suited to describe the structure, dynamics and function of complex systems – new strategies are needed. Meanwhile, within about the last thirty years or so, is has become possible to theoretically describe even complex molecular systems such as proteins in a "first principle" manner. This is due to tremendous progress in the development of new theoretical and computational tools, the ever growing performance of computers, and the advent of highly sensitive experimental techniques such as multidimensional spectroscopy.

This lecture on *Classical Complex Systems* is aimed to introduce the main theoretical and computational concepts of this field to graduate students of Physics and Chemistry. After a general introduction (Section 1) and a brief review of statistical mechanics (Section 2), we are first concerned with the computational approach to complex systems (Section 3), before we consider the their theoretical treatment in Sections 4 and 5. The ordering reflects the course of the accompanying computer lab lectures, which most naturally start with the simulation of simple N-particle systems before considering, e.g., the stochastic modeling of complex systems.

Although the manuscript is self-contained and aims to lead from simple to more involved physical problems, it can certainly not replace a good textbook on these matters. A general introduction to the concepts of modern statistical mechanics is given by the monograph by David Chandler [Chandler, 1987] and the beautiful book of Ken Dill [Dill and Bromberg, 2010]. Well-done overviews of computer simulations are the books Leach [Leach, 1996] and Frenkel [Frenkel and Smit, 2002]. For the theory part, I favor the monographs Zwanzig [Zwanzig, 2001] and [Nitzan, 2006].

I thank all the students who over the years helped to improve the manuscript by asking questions and making suggestions. I also thank Peter Hamm and Rainer Hegger for many inspiring discussion and for numerous figures, slides, and other help they provided. Florian Sittel, Carsten Burgard, Simon Gramatte and Marion Furtwängler-Fritz helped to get the manuscript typed and illustrated. As it is still far from perfect, please mail corrections and comments to stock@physik.uni-freiburg.de.

Freiburg, Fall 2022

Gerhard Stock

Contents

Part I

Introduction: Classical and Statistical Mechanics

Chapter 1

Overview

1.1 First principles description: From quantum to classical mechanics

Microscopic systems such as polyatomic molecules *per se* require a quantum-mechanical description, which becomes quite laborious with increasing system size. To facilitate the computational treatment of large molecular systems, we would rather use classical mechanics which scales linearly with system size. But, under what circumstances would a classical formulation work for microscopic systems? Starting from a proper quantum-mechanical formulation, in the following we introduce a series of assumptions and approximations which pave the way to a classical description of atomistic systems.

We know how to describe the dynamics of systems with a few (say N) particles. In classical mechanics, for example, we construct the Hamiltonian function

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + V\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\right) \tag{1.1}$$

and solve the $2 \cdot 3N$ Hamilton equations ($i = 1, \ldots, 3N$)

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i} \tag{1.2}$$

or equivalently the $3N$ (second-order) Newton equations

$$m\ddot{x}_i = \dot{p}_i = -\frac{\partial V}{\partial x_i} \; . \tag{1.3}$$

In quantum mechanics, we construct the corresponding Hamilton Operator $\hat{H}$ and solve Schrödinger's equation for the wave function Ψ

$$\hat{H}\Psi\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\right) = E\Psi\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\right) . \tag{1.4}$$

In practice, we wish to apply this to molecules with

M electrons $\quad r = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_M)$

$\quad N$ nuclei $\quad R = (\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)$

and the Hamiltonian

$$H = T_e(r) + T_n(R) + U(r, R).$$

In the *Born-Oppenheimer* ansatz, we first solve Schrödinger's equation for fixed nuclei $R = R_0 = \text{const}$ (hence $T_n = 0$)

$$[T_e + U(r, R_0)] \Psi_k(r, R_0) = V_k(R_0) \Psi_k(r, R_0), \tag{1.5}$$

which yields as a function of the parameter R_0 the
- adiabatic potential-energy surfaces $V_k(R_0)$
- adiabatic electronic wave functions $\Psi_k(R_0)$

Restricting ourselves to the electronic ground state, $k = 0$, and by solving equation (6.6) for many nuclear geometries $R = R_0$, we can construct the *electronic potential-energy surface*

$$V_0(R_0) \equiv V(R) = V(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N).$$

For example, for a diatomic molecule such as H_2 with a single internal coordinate (the H-H distance), the potential-energy surface typically would look like in Fig. 1.1. It
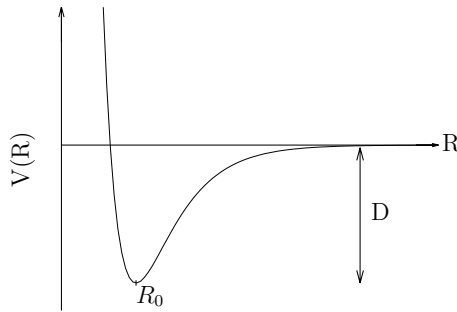

Figure 1.1: Potential energy curve obtained for the electronic ground state of a diatomic molecule, revealing bond length R_0 and dissociation energy D.

- is repulsive for small R,
- has minimum at $R_0 \rightarrow$ chemical bond,
- and leads to $V \xrightarrow{R \rightarrow \infty} 0$, i.e., dissociation for $E > D$.

4

To describe the motion of the nuclei, in a second step, we solve Schrödinger's equation for the nuclear wave function $\phi(R)$ (pertaining to the electronic ground state Ψ_0)

$$(T_n(R) + V(R))\,\phi(R) = E\phi(R),\tag{1.6}$$

where possible couplings to excited electronic potential-energy surfaces $V_{k\neq 0}$ can usually be neglected (Born-Oppenheimer approximation). This quantum-mechanical approach is often referred to as "*ab initio*" (or "*first principles*") approach.

What kind of systems can we describe this way? Using commercial "quantum-chemical" programs such as *Gaussian*, nowadays one routinely solves the *electronic* Schrödinger equation (6.6) for, say ≈ 100 atoms for a few points R. (*Question:* Are a few points enough for our purposes?) To treat larger systems for many points, one often employs "building block" models to construct the potential energy function. For example, to describe the motion of N identical atoms, we first solve the electronic problem for $N = 2$, and obtain an interaction potential $V(R_{12}) = V_{12}(|\boldsymbol{R}_1 - \boldsymbol{R}_2|)$ of the two particles as shown in Fig. 1.1. Now, if we assume that only *pairwise interactions* exist between the particles, we can approximate the potential energy of the N-atom system by

$$V(\boldsymbol{R}_1, \ldots \boldsymbol{R}_N) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}V_{ij}(R_{ij}) \equiv \sum_{i<j}V_{ij}(R_{ij}).\tag{1.7}$$

Question: Why is Eq. (1.7) an approximation?

In a similar vein, we can approximate the potential energy of any N-particle system (say, a complex molecule such as a protein) by using the ansatz of pairwise interactions. That is, by solving the electronic Schrödinger equation to obtain interaction potentials of small fragments of the system and assuming that the total potential energy is given as sum of all these interaction potentials, we can obtain the potential energy function of a many-particle system via a "*divide and conquer*" strategy. The resulting potential energy model is usually termed empirical "*force field*" (because its derivative give the forces acting on the particles). We will discuss such force fields in some detail in Part II of this class.

Having successfully constructed a suitable potential energy model $V(R)$, we still need to describe the motion of the nuclei on $V(R)$. Unfortunately, the nucleic Schrödinger equation (6.6) can (without further approximations) only be solved for a few coordinates, since the effort scales with $\sim 10^N$.

(*Question:* Why is that?)

In many cases of interest, though, it turns out sufficient to treat the motion of the nuclei classically via

$$M_i \ddot{R}_i(t) = -\frac{\partial V(R)}{\partial R_i}. \tag{1.8}$$

The classical limit, $\hbar \to 0$, is typically given for
- heavy particles (having a small quantum excitation energies ε) and

- high temperatures T (such that the system is highly excited).

Problem: Motivate the classical limit by discussing the heat capacity of a quantum and a classical harmonic oscillator.

In particular, classical approximations are often valid when we consider *highly averaged* observables, e.g., the mean distance between two particles. On the other hand, classical mechanics will fail if we want to describe *in detail* the motion of *light* particles at *low* temperatures (enabling tunneling and other wave-related quantum effects). For most situations, it is meanwhile well understood when we can approximately describe nuclear motion via classical mechanics. In practice, e.g., for biomolecular simulations, quantum effects are not a big issue.

Here we wish to describe large molecular systems such as
- liquids (e.g., water), solid-state systems (e.g., crystals)

- solvated molecules, such as a protein in aqueous solution

- molecules on surfaces, nanostructures, ...

and their properties, such as
- structure

- thermodynamic behavior, e.g., $C_V = \frac{\partial U}{\partial T}$

- dynamics, such as
  - phase transitions
  - aggregation, folding
  - behavior after excitation

which is typically discussed for $N \sim 10^{23}$ particles.

Hence, we might want to consider an approach using *statistical mechanics*.

1.2   Statistical description and probability

Usually we are not interested in the detailed motion of 10^{23} particles but rather in the *average* behavior of the particles. Tagging a single H_2O molecule in liquid water, we find that it perform a quite complicated random-looking walk:
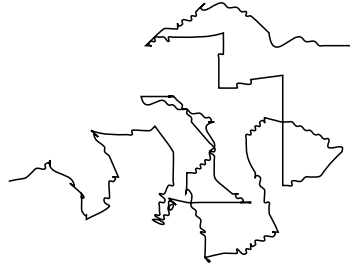
Figure 1.2: Random walk of a tagged molecule in solution.

It looks random,

- because we do not know the motion of the other molecules

- because of classical chaos

Typically, however, we do not watch a single particle, but we monitor the average of certain quantities. A simple example is the total energy of an isolated system

$$E_{\text{tot}} = \sum_{i=1}^{N} E_i\left(t\right) = \text{const},$$

which is constant, although the $E_i\left(t\right)$ are not. Another example is the mean velocity of a particle

$$\langle v \rangle = \frac{1}{N} \sum_{i=1}^{N} v_i.$$

From kinetic gas theory we know that at a temperature T the *distribution* of v is given by

$$W\left(v\right) \sim v^2 e^{-mv^2/2kT}, \tag{1.9}$$

the so-called Maxwell-Boltzmann distribution, where k denote the Boltzmann constant.
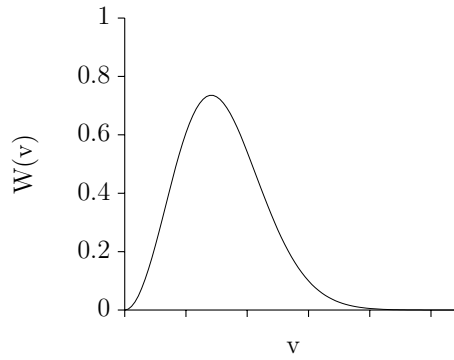
Figure 1.3: Maxwell-Boltzmann velocity distribution.

Being a probability distribution, $W(v)$ is normalized

$$\int_0^\infty W(v)\,dv = 1, \tag{1.10}$$

has the mean value

$$\langle v \rangle = \int_0^\infty v \cdot W(v)\,dv, \tag{1.11}$$

and the mean (squared) width

$$\left\langle \delta v^2 \right\rangle := \left\langle (v - \langle v \rangle)^2 \right\rangle = \left\langle v^2 \right\rangle - \langle v \rangle^2. \tag{1.12}$$

$W(v)$ and its moments $\langle v^n \rangle$ are statistically averaged quantities.

To summarize, we have found that to characterize macroscopic systems we need to adopt a statistical description, which is based on probability, which essentially means lack of knowledge.

Example: Ideal gas simulation

Given is a box in $2D$ (for easy representation) filled with N molecules. At time $t = 0$, all molecules start with the same velocity but random direction. Having a finite radius $R > 0$, the molecules scatter elastically, thus changing their velocity. We are interested in the velocity distribution, which is obtained via a *histogram* like the following:

8
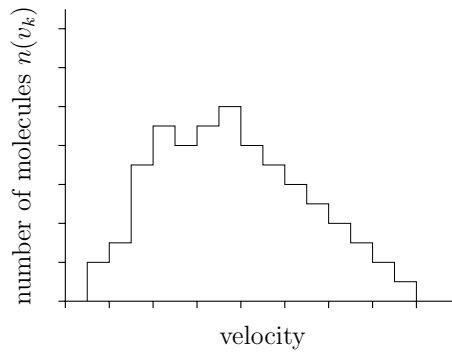
Figure 1.4: Histogram of the velocity distribution.

$$W\left(v_k\right) = \frac{n\left(v_k\right)}{N_{\text{time}} \cdot N_{\text{particles}}}$$

Observation of simulation:

- after some time, $W\left(v\right)$ looks like a Maxwell distribution

- for $R = 0$, the molecules do not interact $\Rightarrow v_i\left(t\right) = v_i\left(0\right)$

To reproduce numerically the statistical theory leading to the Maxwell distribution

- we need to *average* over many molecules and/or long time
  (good statistics $\hat{=}$ lots of data)

- the molecules need to interact (even for an ideal gas!)

Properties of probabilities

1. Definition of probability:

$$p_i = \lim_{N \to \infty} \frac{N_i}{N},$$

where $\sum_i N_i = N$ $\qquad \Rightarrow$ $\begin{array}{l} 0 \leq p_i \leq 1 \\ \sum_i p_i = 1 \end{array}$

2. Calculus of probabilities:
   - *Addition*: If outcomes $A, B, \ldots$ with probabilities $p_A, p_B, \ldots$ are *mutually exclusive*, then the probability of observing either A or B or $\ldots$ is $p_A + p_B + \ldots$
   - *Multiplication*: If outcomes $A, B, \ldots$ are *independent*, then the probability to observe A and B and $\ldots$ is $p_A \cdot p_B \cdot \ldots$

3. Maximal information is the probability distribution

$$p_i \quad \text{(for discrete systems)}$$

$$p(x) \quad \text{(for continuous systems)}$$

e.g., dice with $p_i = {}^1\!/_6$:
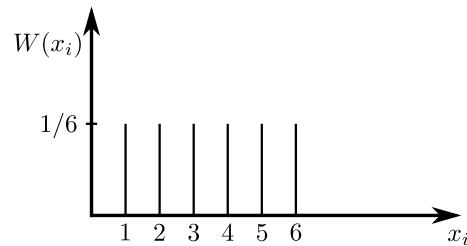
Figure 1.5: Equal distribution obtained for a six-sided dice.

or a Gaussian distribution:
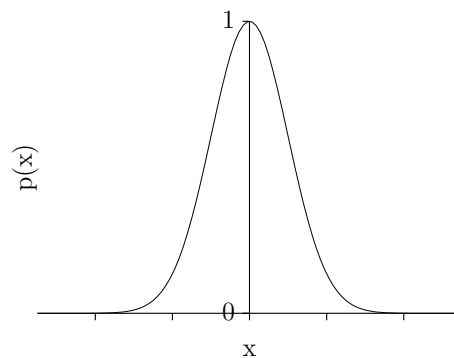
Figure 1.6: Probability distribution obtained for a Gaussian system.

4. Probability distributions are characterized by their moments

$$\langle x^n \rangle = \int x^n p(x) \, dx$$

with

$n = 0$ normalization

$n = 1$ mean

$n = 2$ variance

$n = 3$ asymmetry

5. A *stochastic process* describes the time evolution of a random variable. The mother of all random processes is the random walk, which we are going to consider next.

10

1.3 Stochastic processes: The random walk

A random walk describes the random motion of objects (e.g., particles) in one or more dimensions. For example, the walk of a (completely) drunken man along a street (and many other examples like diffusion, heat conduction, etc). In $1D$, each step of unit length in $+x$ or $-x$ direction has the same probability

$$p_{+x} \equiv p = {}^1\!/2 = p_{-x} \equiv q$$

After N steps, m steps are taken in $+x$ direction. Equivalently, $(N-m)$ steps are taken in $-x$ direction.

We aim to calculate the probability distribution of m-values $P(m,N)$ for an N-step walk. As an example, let $N=5$ and consider the path $+--++$ with total probability $pqqpp = p^3 q^2$, i.e., $m = 3$. There are several paths that lead to $m = 3$: There are $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$ ways to arrange 5 items $abcde$. If m items (and thus also $N-m$ items) are identical, the $m!$ (and $(N-m)!$) permutations are identical and we need to divide by $m!$ and $(N-m)!$. Given that the probability of a single permutations is $p^3 q^2$, we have

$$P(3,5) = \frac{5!}{2!3!} p^3 q^2 = \binom{5}{3} p^3 q^2.$$

I.e., in general we have a *binomial distribution* (with $p = q = 1/2$)

$$P(m,N) = \frac{N!}{m!\,(N-m)!} \left(\frac{1}{2}\right)^N. \tag{1.13}$$

We can approximate $P(m,N)$ for large N by using Stirling's approximation

$$\ln n! \approx n \ln n - n \qquad \text{(for large n)}$$

and consider

$$
\begin{aligned}
\ln P &\approx \ln N! - \ln m! - \ln(N-m)! + N \ln \frac{1}{2} \\
&= N \ln N - N - m \ln m + m - (N-m) \ln(N-m) + (N-m) + N \ln \frac{1}{2} \\
&= N \ln N - m \ln m - (N-m) \ln(N-m) + N \ln \frac{1}{2}.
\end{aligned}
$$

We then make a Taylor expansion of $\ln P$ around the most probable end point m^* for fixed N:

$$\ln P(m) = \ln P(m^*) + \left.\frac{d \ln P}{dm}\right|_{m^*} (m - m^*) + \frac{1}{2} \left.\frac{d^2 \ln P}{dm^2}\right|_{m^*} (m - m^*)^2$$

To find m^*, we calculate

$$\begin{aligned}
\frac{d\ln P}{dm} &= -\left(\ln m + \frac{m}{m}\right) + \ln(N-m) - \frac{N-m}{N-m}(-1) \\
&= \ln(N-m) - \ln m \\
&= \ln\frac{N-m}{m} \overset{!}{=} 0
\end{aligned}$$

$$\rightarrow \frac{N-m^*}{m^*} = 1 \quad ; \quad m^* = \frac{N}{2}$$

$$\begin{aligned}
\left.\frac{d^2\ln P}{dm^2}\right|_{m^*} &= \frac{1}{N-m^*}(-1) - \frac{1}{m^*} \\
&= -\frac{1}{N-\frac{N}{2}} - \frac{2}{N} = -\frac{4}{N}
\end{aligned}$$

Put together, we find

$$P(m,N) = P^* e^{-2(m-m^*)^2/N},$$

that is, we obtain a Gaussian distribution from the binomial distribution for large N.

Next, we convert to the number of forward steps m to x, i.e.,

$$\begin{aligned}
x &= m - (N-m) \quad \text{(number of forward} - \text{reverse steps)} \\
&= 2m - N
\end{aligned}$$

Since $m^* = \frac{N}{2}$, $x^* = 0$, i.e., most likely we end at $x = 0$. Using

$$m - m^* = \frac{(x+N)}{2} - \frac{N}{2} = \frac{x}{2},$$

we get

$$P(x) = P^* e^{-x^2/2N}.$$

The normalization condition (using $\int_{-\infty}^{\infty} e^{-ax^2}dx = \sqrt{\frac{\pi}{a}}$)

$$1 = \int_{-\infty}^{\infty} P(x)\,dx = P^* \int_{-\infty}^{\infty} e^{-x^2/2N}dx = P^*\sqrt{2\pi N}$$

yields then the final result

$$P(x) = (2\pi N)^{-1/2}\, e^{-x^2/2N}. \tag{1.14}$$

Calculating the fluctuations via the mean square (using $\int_{-\infty}^{\infty} x^2 e^{-x^2/2N}dx = N\sqrt{2\pi N}$)

$$\langle x^2\rangle = \int_{-\infty}^{\infty} x^2 P(x)\,dx = N,$$

we find that the root mean square distance traveled by the walker is

$$\left\langle x^2 \right\rangle^{\frac{1}{2}} = N^{\frac{1}{2}}.$$

Starting at time $t = 0$, the walker ends at $t = \Delta t \cdot N$, Δt being the time step. Thus

$$\left\langle x^2\left(t\right) \right\rangle \sim t \quad \text{(diffusion)} \tag{1.15}$$

Comparing to the free motion of a particle with constant velocity v, we have

$$x = vt, \; x^2\left(t\right) \sim t^2 \quad \text{(ballistic motion)} \tag{1.16}$$

i.e. due to its random nature, diffusion is significantly slower!

Discussion: The random walk and its properties are very general and apply to many situations in physics and statistics.

- *Diffusion* is an important transport process in gases, liquids and soft matter (e.g., cells) and a standard example of the random walk. Consider Fick's law of the particle density $P(x,t)$ (in one-dimensional notation)

$$J(x,t) = -D\frac{\partial P(x,t)}{\partial x}, \tag{1.17}$$

  stating that current $J(x,t)$ is proportional to gradient $\frac{\partial P(x,t)}{\partial x}$, with D being the diffusion constant. Employing the continuity equation

$$\frac{\partial P(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = 0,$$

  we obtain the diffusion equation

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial^2 P}{\partial x^2}. \tag{1.18}$$

  Insertion of Eq. (1.14) shows, that a random walk with $N = t/\Delta t$ and $D = x_0^2/2\Delta t$ (with time step Δt and step size x_0) is a solution of the diffusion equation. Remarkably, we have thus shown that a microscopic theory is the solution of phenomenological transport law.

- Many physical processes, such as particle diffusion, heat transport, motion of biomolecules, etc. are diffusive, even though the underlying microscopic behavior may be quite different. This independence of the statistical theory of the microscopic behavior is called "*universality*".

- Random walks are "scale invariant", i.e., they look the same on different length scales as long as N is still large.
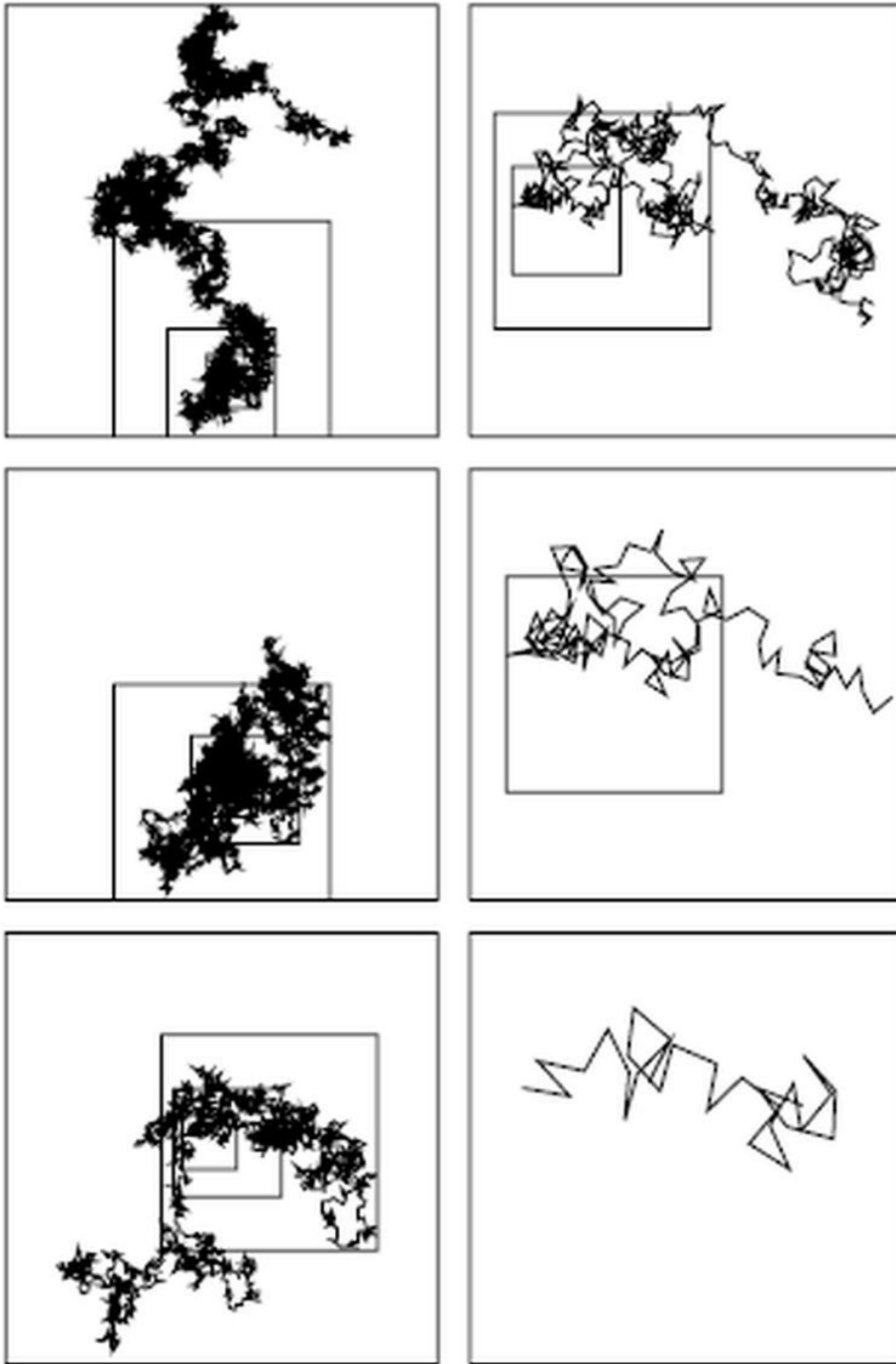
Scale invariance of a two-dimensional random walks. Taken from J. P. Sethna: Statistical Mechanics: Entropy, Order Parameters, and Complexity (2006)

1.4 Computer simulations: Averages and ergodicity

A random walk can easily be simulated on a computer. To do so, we need a *random number generator*. For example,

```
do i=1..N
  r(i)=Ran()
enddo
```

gives you a sequence of N random numbers $r(i)$ with $0 \leq r(i) \leq 1$.

A random path is simulated by the following algorithm:
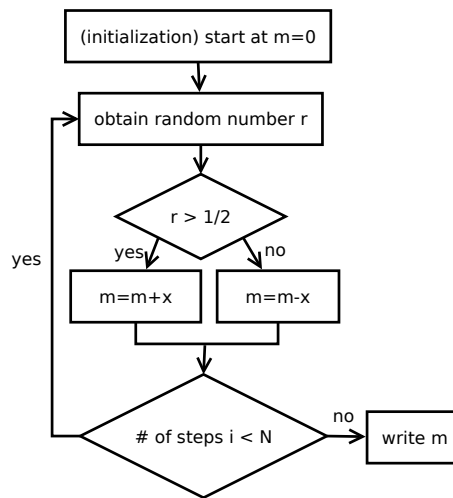
Figure 1.7: Flow chart of a simple random walk algorithm.

By running many of these paths, we obtain the distribution $P(m, N)$ via an "*ensemble average*." The scheme is a very simple example of a *Monte-Carlo* algorithm which es an extremely powerful method. The name refers to the fact, that you compute something by throwing dice or via random numbers.

Example 1: Calculation of π.

Another example of a simple Monte-Carlo application is the calculation of π via random counting of points in a unit square and a unit circle

$$\frac{N_S}{N_C} = \frac{1^2}{\pi \left(\frac{1}{2}\right)^2} = \frac{4}{\pi}$$

with N_S being the total number of 'moves' and N_C the number of 'moves' which end in the unit square.

Example 2: Traffic jam

- N cars go in a circle (particles in one dimension).

- Each driver wants to go at some maximum speed, however, if the car in front slows down, the driver breaks with some random time delay.

- Considering the velocity distribution of the cars, we find for $N = 100$ steady traffic and for $N = 200$ a traffic jam with a transition at $N = 170$.

Postulate of statistical mechanics

The simulation approach is based on the plausible *assumption* that the experimentally measured value is given by an *average* over all *microstates*. In the above example of ideal gas, a microstate is defined by its molecular

$$
\begin{array}{ll}
\text{positions} & x_i(t) \\
\text{and velocities} & v_i(t) \text{ with } i = 1\ldots N.
\end{array}
$$

The theoretical foundation of statistical mechanics is based on a single *postulate*:

All microstates of the same energy are equally likely.

A counter example is given by a few non-interacting molecules. Hence the postulate requires

- many particles, observed for a long time,

- some particle-particle interaction, and a

- random character, e.g., via chaos.

Time and ensemble average

There are two ways to obtain an average $\langle \dots \rangle$ in order to calculate the mean of an observable x (e.g., position, energy, etc).

$$\langle x \rangle = \sum_j p_j x_j$$

Given an equilibrium simulation, a *time average* is performed by measuring x for a single system as a function of time t:
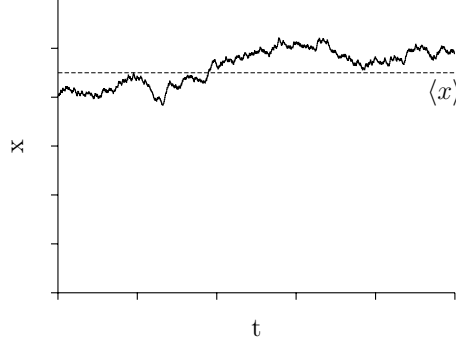
Figure 1.8: Calculation of the mean $\langle x \rangle$ via a time average of the trajectory $x(t)$.

$$\langle x \rangle_T = \frac{1}{T} \int_0^T x(t)\, dt \tag{1.19}$$

where the measurement time T needs to be long compared to the (slowest) timescale τ_{fluc} of the fluctuations of x. Discretizing the time as $t_n = \delta t n$ with $\delta t = \frac{T}{N}$, we get

$$
\begin{aligned}
\langle x \rangle_T &= \frac{1}{N_t} \sum_{n=1}^{N_t} x(t_n) \\
&= \sum_j \frac{n_j}{N_t} x_j \\
&= \sum_j p_j x_j
\end{aligned}
$$

where n_j denotes the number of time steps with $x = x_j$.

An *ensemble average* means that we measure a large number N of identical systems for a time that is short compared to t_{Fluk}

$$\langle x \rangle_E = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{1.20}$$

17

If for large N and long T both averages coincide

$$\langle x \rangle_T = \langle x \rangle_E \tag{1.21}$$

we call the system ergodic. This will be the case if the system visits all microstates during the time average and also averages over microstates in the ensemble average.

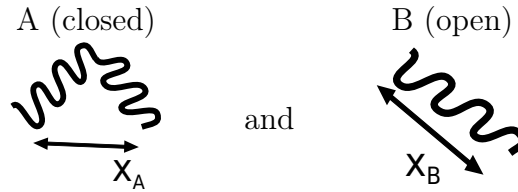But: Due to the enormous number of possible microstates of a macroscopic system, usually only a (very small) part of all microstates are sampled, while nevertheless (6.18) holds. To illustrate this, consider a molecule in solution which adopts the structures

A (closed)   B (open)

and

X_A X_B

which have different end-to-end distances x. A typical time trace of x may look like
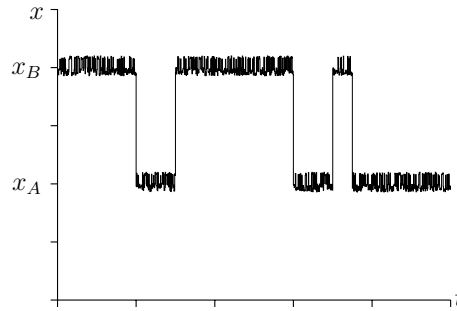
Figure 1.9: Typical time trace of a measured distance.

Hence, to obtain a converged mean value $\langle x \rangle$, we need to average over many $A \leftrightarrow B$ transitions either in time or via an ensemble. While the time evolution of x certainly depends on all other degrees of freedom of the system (molecule and solvent), the main effect of these degree of freedom is to make x

- undergo small fluctuations in a stable state

- jump from $A \leftrightarrow B$ once in a while.

Hence, to satisfy (6.18) we do not need to sample all microstates, but only enough that the evolution of x is described well.

The above examples are so simple, that one can also easily calculate their outcome analytically. Computational simulations, however, are much more powerful.

18

Simulation approach

Starting from the quantum mechanical formulation (6.6), we try to find appropriate approximations such that a practical calculation of the desired system and dynamics become feasible. This usually includes the following steps:

1. **Construct a model for $V(R) = V(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)$**
   The vibrations in a crystal (the phonons) are, for example, well described by coupled harmonic oscillators. This model is often referred to as the *force field* because it directly gives (via $F_i = -\frac{\partial V}{\partial R_i}$) the forces acting on the particles.

2. **Describe the nuclear motion classically**
   This can be done, for example, via equations (2.2) or (6.4) instead of the quantum mechanical equation (6.7). Even though atoms are microscopic particles, this works surprisingly well in many cases. To simulate an external system, we, of course, can not consider 10^{23} particles, but we can use tricks like periodic boundary conditions.

3. **Perform time or ensemble average** to calculate observables of interest.

In principle, both approximations can be systematically improved. We can, for example, use a more realistic model $V(R)$ or introduce quantum mechanical corrections to the nuclear motion, or simulate a larger system. Despite of its approximate nature, this approach is often called a "first principle" approach.

Exercises

Exercise I: Velocity Distributions

Exercise II: Random Walk

Exercise *: Contagion Spread Simulation

Chapter 2

Statistical mechanics in a nutshell

The purpose of this chapter is not to comprehensively review statistical mechanics, but to focus on the role that the "driving forces" *energy*, *entropy* and *free energy* play in molecular dynamics. For a comprehensive treatment, see for example the beautiful book of Ken Dill [Dill and Bromberg, 2010].

2.1   Entropy

The second law of thermodynamics says, somewhat loosely stated, that every isolated system approaches a state of *maximum disorder*. To illustrate this, consider the following experiment:

(a)   (b)
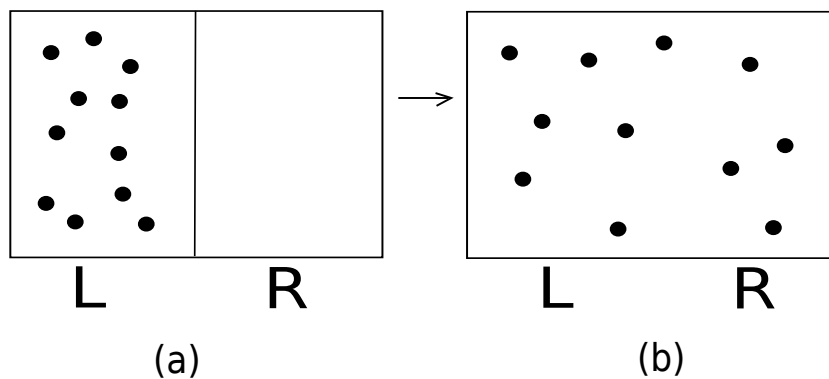
Why do the molecules disperse to both sides? Recall that

- both cases (a) and (b) have the same energy and

- both cases - seen as microstates - are equally likely.

So, what is the driving force for this effect?

To calculate the probability of a specific *microstate*, we assume that the probability of particle i to be on the left side of the box is $p_{i,L} = \frac{1}{2}$. Hence, the probability that *all* N particles are left is

$$p_L = \left(\frac{1}{2}\right)^N.$$

For $N = 10^{23}$, $p_L \approx 10^{-10^{22}}$, i.e., this case *never* happens!

Problem: Illustrate this by estimating the probability that this occurs during the life time of the universe.

According to the fundamental postulate of statistical mechanics, the probability of any other microstate is the same. Assuming that l molecules are left, we find

$$p_L = \left(\frac{1}{2}\right)^l$$

and

$$p_R = \left(\frac{1}{2}\right)^{N-l},$$

thus

$$p_L \cdot p_R = \left(\frac{1}{2}\right)^N.$$

An obvious *macroscopic* observable to consider is the number of molecules on the left, *disregarding which one*. In direct analogy to the random walk studied above, this probability is given by the binomial distribution

$$p_l = \frac{N!}{l!\,(N-l)!} \left(\frac{1}{2}\right)^N = \binom{N}{l}\left(\frac{1}{2}\right)^N. \tag{2.1}$$

Using the results taken from the random walk, we find that the mean particle number is (as expected)

$$\langle l \rangle = \sum_{l=0}^{N} l \cdot p_l = \frac{N}{2}. \tag{2.2}$$

The width of the distribution is

$$\Delta l = \sqrt{\langle (l - \langle l \rangle)^2 \rangle} = \sqrt{\frac{N}{2}}$$

and the relative width

$$\frac{\Delta l}{\langle l \rangle} \sim \frac{1}{\sqrt{N}}. \tag{2.3}$$

This means, we obtain

- a broad distribution for a few particles

- a very sharp distribution for many (e.g., 10^{23}) particles
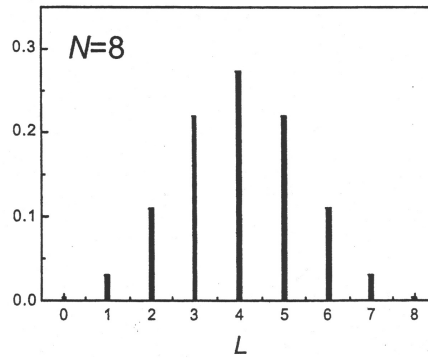
Figure 2.1: Distribution q_L of a system of 8 particles. From [?]
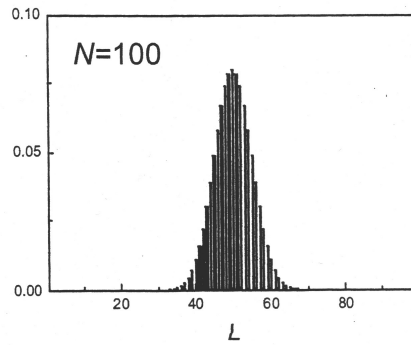
Figure 2.2: Distribution q_L of a system of 100 particles. From [?].

From equations (2.1) and (2.2) we learn that the probability distribution p_l has its maximum at $\langle l \rangle = \frac{N}{2}$. This is simply because this case provides the largest number of possibilities to divide up the N particles. In other words:

The macrostate that contains the largest number of microstates is the most likely one.

Since there are way less "ordered" microstates [as in (a)] than disordered microstates [as in (b)], this reasoning serves to explain the above observation of the dispersing particles and the second law of thermodynamics in general. In other words, starting in an ordered state (a) (a "nonequilibrium state"), the system evolves into a disordered state (b) (the "equilibrium state") since there are many more possibilities to create disorder than there are possibilities to create order.

Note that we did not make any assumptions about the interactions, but only used combinatory arguments. This generality makes statistical mechanics a universal description.

22

Definition of entropy

We wish to define a quantity that measures disorder and can be viewed as a driving force. To this end, we consider an isolated system with

- constant energy E

- constant volume V

- constant number of particles N

(a so-called *microcanonical ensemble*) and define $\Omega(E, V, N)$ as the number of microstates for a system with given E, V, N. The probability of a single microstate then is $p = 1/\Omega$. Following Boltzmann, entropy is defined by

$$\boxed{S = k \ln \Omega,} \tag{2.4}$$

where k is a constant. In thermodynamics, k is Boltzmann's constant.

Why $\ln \Omega$ and not Ω? We wish to define entropy as an *extensive* (or *additive*) function, i.e., the total entropy of two (non-interacting) subsystems A, B is

$$S_{\text{total}} = S_{AB} = S_A + S_B$$

(just like the energy is extensive). For independent subsystems we have

$$\Omega_{AB} = \Omega_A \cdot \Omega_B$$

and thus

$$S_{AB} = k \ln \Omega_{AB} = k \ln \Omega_A \Omega_B = S_A + S_B,$$

which explains the definition of the entropy via $\ln \Omega$. At $T = 0$, we expect the system to be in its ground state, i.e., $\Omega = 1$ (if it is not degenerated). Hence

$$S(T = 0) = 0$$

(sometimes called third law of thermodynamics).

Let us now calculate the entropy for a generic example. To this end, we consider a system of N particles with constant total energy E. We assume that $n_1, n_2, \ldots$ are the number of particles in the energy levels $\epsilon_1, \epsilon_2, \ldots$ of the system. Hence

$$N = \sum_i n_i \quad \text{total number of particles,}$$

$$E = \sum_i n_i \epsilon_i \quad \text{total energy.}$$

Using the same arguments as above, the number of microstates for given $\{n_i\}$ is

$$\Omega\left(n_1, n_2, \ldots\right) = \frac{N!}{\Pi_i n_i!}. \tag{2.5}$$

Using Stirling's approximation

$$\ln n! \approx n \ln n - n \qquad \text{(for large n)},$$

we get for the entropy

$$\begin{aligned}
\frac{S}{k} &= N \ln N - N - \sum_i n_i \ln n_i - \sum_i n_i \\
&= N \ln N - \sum_i n_i \ln n_i.
\end{aligned}$$

Introducing the population probability of energy level ϵ_i

$$p_i = \frac{n_i}{N}$$

we obtain

$$\begin{aligned}
\frac{S}{k} &= N \ln N - N \sum_i p_i \ln N p_i \\
&= N \ln N - N \ln N \underbrace{\sum_i p_i}_{=1} - N \sum_i p_i \ln p_i \\
&= -N \sum_i p_i \ln p_i.
\end{aligned}$$

Hence, the entropy per particle is

$$\boxed{S = -k \sum_i p_i \ln p_i.} \tag{2.6}$$

Limiting cases:

- total order: e.g., $p_1 = 1, p_{i \neq 1} = 0 \to S = 0$

- total disorder: $p_i = \text{const} \to S = \text{max}$

The principle of maximum entropy predicts the probability distribution

To show that the second law of thermodynamics - or more general the principle of maximum entropy (Jaynes 1957) - is of practical use, we use it to predict the probability distribution of some simple systems.

As first example, consider an unbiased t-sided dice. As the probabilities must sum to one, we have

$$\sum_{i=1}^{t} p_i = 1. \tag{2.7}$$

Hence, we seek the distribution $(p_1, p_2, \ldots, p_t)$ that causes the entropy function

$$S(p_1, p_2, \ldots, p_t) = -k \sum_i p_i \ln p_i \overset{!}{=} \max \tag{2.8}$$

to be at its maximally possible value, subject to condition (2.7). Here, we set $k = 1$. Using the Lagrange multiplier method, with

$$g(p_1, \ldots, p_t) = \sum_{i=1}^{t} p_i - 1 = 0,$$

we have the condition

$$\left(\frac{\partial S}{\partial p_j}\right) - \alpha \left(\frac{\partial g}{\partial p_j}\right) = 0, \text{ for } j = 1 \ldots t. \tag{2.9}$$

Therefore

$$-1 - \ln p_j - \alpha = 0$$
$$\Rightarrow p_j = e^{-1-\alpha}.$$

Using that

$$\frac{p_j}{\sum_i p_i} = \frac{e^{-1-\alpha}}{te^{-1-\alpha}} = \frac{1}{t},$$

we find a flat distribution, i.e., *if there is no bias all outcomes are equally likely.*

As a second example, we again roll a die with t sides. After N rolls, we obtain a *total score* E, which is given by the sum of the N individual scores (i.e., the number $i = 1, 2, \ldots, t$ of the upside face). Now, we only consider rolls that give some chosen value of E. Hence we have

$$E = \sum_{i=1}^{t} \epsilon_i n_i \quad \begin{array}{l} \text{with } \epsilon_i \text{ being the score when face } i \text{ shows up} \\ \text{and } n_i \text{ being the number of times } \epsilon_i \text{ appears.} \end{array}$$

and the average score per roll is

$$\langle \epsilon \rangle = \frac{E}{N} = \sum_i \epsilon_i p_i \tag{2.10}$$

with $p_i = n_i/N$.

Now, what is the distribution that satisfies

- maximum entropy (2.8)

- the normalization condition (2.7)

- the average score condition (2.10)?

Writing

$$g(p_1, \ldots, p_t) = \sum_{i=1}^{t} p_i - 1 = 0$$

$$h(p_1, \ldots, p_t) = \sum_i p_i \epsilon_i - \langle \epsilon \rangle = 0$$

we get from the Lagrange formalism

$$
\begin{aligned}
\left(\frac{\partial S}{\partial p_j} \right) - \alpha \left(\frac{\partial g}{\partial p_j} \right) - \beta \left(\frac{\partial h}{\partial p_j} \right) &= 0 \\
-1 - \ln p_j - \alpha - \beta \epsilon_j &= 0 \\
p_j = e^{-1-\alpha-\beta\epsilon_j} &= \frac{e^{-1-\alpha-\beta\epsilon_j}}{\sum_i e^{-1-\alpha-\beta\epsilon_i}},
\end{aligned}
$$

where we divided by $\sum_i e^{-1-\alpha-\beta\epsilon_i} = 1$, to eliminate α. This yields the Boltzmann population probability

$$p_j = \frac{e^{-\beta\epsilon_j}}{\sum_i e^{-\beta\epsilon_i}}. \tag{2.11}$$

Hence, by using the principle of maximum entropy, we have derived the Boltzmann distribution $\sim e^{-\beta\epsilon_i}$. The normalization factor is the partition function

$$Q = \sum_i e^{-\beta\epsilon_i}. \tag{2.12}$$

Remarks:

- The derivation above holds for any case with a constraint of the kind (2.10). In thermodynamics, we identify

$$E \ \hat{=} \ \text{energy}$$
$$\beta \ \hat{=} \ \frac{1}{kT} \ \text{(inverse) temperature}$$

  The latter is done by equating the entropy change dS

  - from the statistical definition, [Eq. (2.6)]

  $$dS \ = \ -k \sum_j \ln p_j \, dp_j$$

  - and from the thermodynamical definition

  $$dS \ = \ \frac{\delta q}{T},$$

  where δq is the transferred heat.

- The partition function Q is a central quantity, from which all observable quantities of the system can be calculated at equilibrium.

2.2 The partition function

As a first example of the usability of the partition function Q, we have a look at the average value of the energy

$$
\begin{aligned}
\langle E \rangle &= \sum_j p_j \left(N, V, \beta \right) \epsilon_j \left(N, V \right) \\
&= \sum_j \frac{\epsilon_j \left(N, V \right) e^{-\beta \epsilon_j (N,V)}}{Q \left(N, V, \beta \right)}.
\end{aligned}
$$

Using

$$
\left(\frac{\partial \ln Q \left(N, V, \beta \right)}{\partial \beta} \right)_{N,V} = \frac{1}{Q \left(N, V, \beta \right)} \sum_j \left[-\epsilon_j \left(N, V \right) \right] e^{-\beta \epsilon_j}
$$

we get

$$
\langle E \rangle = - \left(\frac{\partial \ln Q}{\partial \beta} \right)_{N,V}. \tag{2.13}
$$

Hence, the partition function can be used to calculate the average energy. Alternatively, one can use

$$
\frac{\partial f}{\partial T} = \frac{\partial f}{\partial \beta} \frac{\partial \beta}{\partial T} = \frac{\partial f}{\partial \beta} \left(-\frac{1}{kT^2} \right)
$$

to express the energy as

$$
\langle E \rangle = kT^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V}.
$$

Example:
Calculate $\langle E \rangle$ for a proton in the magnetic field B_z. Protons have a spin of $\pm\frac{1}{2}$ with the energy $E_\pm = \pm\frac{1}{2}\hbar\gamma B_z$, where γ denotes the gyromagnetic ratio.

$$
\begin{aligned}
Q \left(T, B_z \right) &= e^{\beta \hbar \gamma B_z / 2} + e^{-\beta \hbar \gamma B_z / 2} \\
&= 2 \cosh \left(\frac{\beta \hbar \gamma B_z}{2} \right)
\end{aligned}
$$

$$
\begin{aligned}
\langle E \rangle &= - \left(\frac{\partial \ln Q}{\partial \beta} \right)_{B_z} = -\frac{1}{Q \left(\beta, B_z \right)} \frac{\partial Q}{\partial \beta} \\
&= -\frac{\hbar \gamma B_z}{2} \frac{\left(e^{\beta \hbar \gamma B_z / 2} - e^{-\beta \hbar \gamma B_z / 2} \right)}{\left(e^{\beta \hbar \gamma B_z / 2} + e^{-\beta \hbar \gamma B_z / 2} \right)} \\
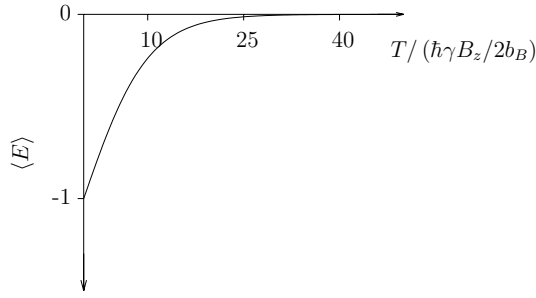&= -\frac{\hbar \gamma B_z}{2} \tanh \left(\frac{\beta \hbar \gamma B_z}{2} \right)
\end{aligned}
$$

Figure 2.3: $\langle E \rangle$ of a proton in a magnetic field.

$$\langle E \rangle \quad \overset{T=0}{\to} \quad \frac{-\hbar \gamma B_z}{2} \quad \text{(all protons in ground state)}$$

$$\langle E \rangle \quad \overset{T \to \infty}{\to} \quad 0 \quad \begin{cases} 50\% & \text{in spin up} \\ 50\% & \text{in spin down} \end{cases}$$

This system is an example of a *two-state model*.

As a second example, we consider the calculation of the entropy S via Q. With

$$p_j \quad = \quad p_j\left(N, V, \beta\right) = \frac{e^{-\beta \epsilon_j (N,V)}}{Q\left(N, V, \beta\right)}$$

we get

$$
\begin{aligned}
S \quad &= \quad -k \sum_j \frac{e^{-\beta \epsilon_j}}{Q} \left(-\beta \epsilon_j - \ln Q\right) \\
&= \quad \underbrace{\beta k}_{=\frac{1}{T}} \underbrace{\sum_j \frac{\epsilon_j e^{-\beta \epsilon_j}}{Q}}_{=\langle E \rangle} + \frac{k \ln Q}{Q} \underbrace{\sum_j e^{-\beta \epsilon_j}}_{=Q} \\
&= \quad \frac{\langle E \rangle}{T} + k \ln Q.
\end{aligned}
\tag{2.14}
$$

By using

$$\langle E \rangle = kT^2 \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V}$$

we therefore obtain

$$S = kT \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} + k \ln Q. \tag{2.15}$$

29

2.3 Ensembles and thermodynamic potentials

In thermodynamics, we distinguish between three main cases: isolated, closed and open systems.

Isolated system: microcanonical ensemble

$$\text{There is no exchange of energy with the environment,} \quad dE = 0$$

$$\text{also, there is no exchange of particles with the environment,} \quad dN = 0$$

$$\text{and the volume is fixed.} \quad dV = 0$$

The resulting thermodynamical state is described by the quantities E, N, V. It is also referred to as *microcanonical ensemble*.

A suitable quantity to describe an isolated system is the entropy, because

- it also depends on N, V, E

- it satisfies a variation principle, i.e., it becomes maximal at equilibrium.

The entropy is therefore called the thermodynamical potential of the microcanonical ensemble.

Fundamental thermodynamic equations

To connect the general statistical considerations above with thermodynamics, we briefly introduce some basic thermodynamical definitions. Using the *internal energy*

$$E = \sum_i n_i \epsilon_i,$$

we can write for the entropy

$$S = S\left(E, V, N\right) \text{ independent variables } E, V, N$$

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN. \tag{2.16}$$

Alternatively, we can write

$$E = E\left(S, V, N\right)$$

$$dE = \underbrace{\left(\frac{\partial E}{\partial S}\right)_{V,N}}_{=:T} dS + \underbrace{\left(\frac{\partial E}{\partial V}\right)_{S,N}}_{=:-P} dV + \underbrace{\left(\frac{\partial E}{\partial N}\right)_{S,V}}_{=:\mu} dN$$

which defines temperature (T), pressure (P) and chemical potential (μ).

$$\boxed{dE = TdS - PdV + \mu dN.} \tag{2.17}$$

(If there are several species, the last term generalizes to $\sum_j \mu_j dN_j$.)

Rearranging (2.17) yields

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN,$$

describing the tendency of energy exchange $(1/T)$, volume exchange (p/T) and particle exchange (μ/T).

Closed system: canonical ensemble

Here the system is in contact with the environment, often called the *bath* with a temperature T, ergo $T = \mathrm{const}$.
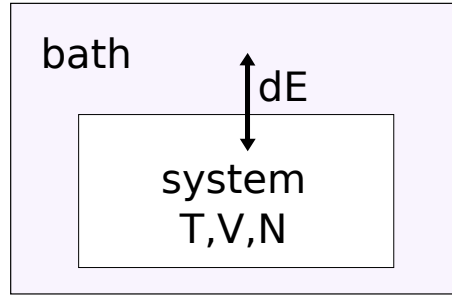

Figure 2.4: Closed system embedded in a heat bath.

- there is energy exchange between system and bath, $dE \neq 0$

- but no exchange of particles $\rightarrow dN = 0$

Therefore, the thermodynamical state is described by N, V, T. What is the corresponding thermodynamical potential that is maximal/minimal to obtain the equilibrium state?

To derive this quantity, the so-called *free energy*, we consider a 'total system' that comprises the 'system' and the 'bath'. For the total system, we find

$$\begin{align}
dS_{\mathrm{tot}} &= dS_S + dS_B \geq 0 & \text{(entropy is extensive)} & \quad (2.18) \\
dE_B + dE_S &= 0 & \text{(total system is isolated)} & \quad (2.19) \\
dS_B &= \frac{1}{T}dE_B + \frac{p}{T}dV_B - \frac{\mu}{T}dN_B \overset{V,N=\mathrm{const}}{=} \frac{1}{T}dE_B & & \quad (2.20)
\end{align}$$

Using (2.19), we find

$$dS_B = -\frac{1}{T}dE_S$$

and with (2.18) we get

$$dS_S - \frac{dE_S}{T} \geq 0,$$

that is an expression that depends only on system variables. In a closed system, one apparently wants to minimize the quantity

$$dE_S - TdS_S \leq 0. \tag{2.21}$$

To this end, we define the *Helmholtz free energy* as

$$F = E - TS, \tag{2.22}$$

Hence

$$dF \underset{T=\text{const}}{\overset{=}{=}} \begin{array}{l} dE - TdS - SdT \\ dE - TdS. \end{array}$$

Due to equation (2.21) the free energy F is at a minimum, if a system with constant (T, V, N) is in equilibrium. $F = E - TS$ shows that there is a balance between internal energy and entropy, the position of this balance is determined by the temperature T. That is, the system wants to both

- minimize E

- maximize S.

At high temperatures, the entropy dominates.

Using equation (2.14), $ST = E + kT \ln Q$, we can also define the free energy in terms of the partition function via

$$F = -kT \ln Q, \tag{2.23}$$

which looks quite similar to Boltzmann's definition of the entropy, $S = k \ln \Omega$.

Other thermodynamical potentials

- **Gibbs free energy**

  The above introduced quantity $F = F(N, V, T)$ is the so-called *Helmholtz free energy*, which refers to the case that $V = \text{const}$. In many cases it is more realistic to assume that the pressure $p = \text{const}$. The corresponding thermodynamical potential is the *Gibbs free energy* defined by

  $$\boxed{G = F + pV = E - TS + pV,} \tag{2.24}$$

  where the additional term describes the work to change the volume V.

- **Open system: grand canonical ensemble**

  In this case, energy and particles are exchanged. Hence T, V and the chemical potential μ are constant and

  $$\boxed{\phi = F - \mu N = E - TS - \mu N} \tag{2.25}$$

  is the corresponding thermodynamical potential.

Example: Dimerization

As an example for the balance between internal energy and entropy, we consider a simple model for dimerization, consisting of two gas particles in a test tube with a discrete volume of V lattice sites [Dill and Bromberg, 2010]. To identify the stable (or equilibrium) state, i.e., the one with minimum free energy, we compare the only two possible states of the system: the dimer state and the dissociated monomer state.
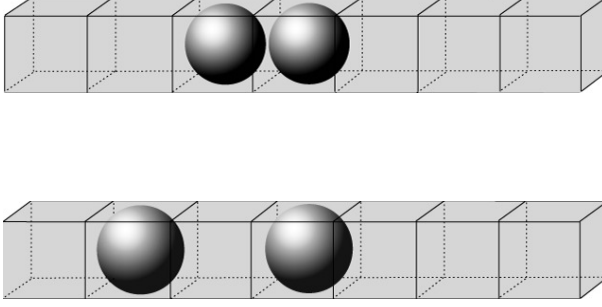

Figure 2.5:
The dimer state (top) and the dissociated state (bottom) of the two-particle model. From [Dill and Bromberg, 2010].

As there are $W_{\mathrm{dimer}} = V - 1$ possible placements of a dimer on a linear lattice of V sites, the resulting free energy for the dimer is

$$F_{\mathrm{dimer}} = E_{\mathrm{dimer}} - TS_{\mathrm{dimer}} = -\epsilon - kT \ln (V - 1),$$

where $\epsilon > 0$ defines a "bond energy" holding the dimer together. For the (dissociated) monomers, on the other hand, the number of possible placements is

$$
\begin{aligned}
W_{\mathrm{monomer}} &= W_{\mathrm{total}} - W_{\mathrm{dimer}} \\
&= \frac{V!}{(2!)\,(V-2)!} - (V - 1) \\
&= \left(\frac{V}{2} - 1\right)(V - 1),
\end{aligned}
$$

as the particles may *not be adjacent* to each other. Combining this with the fact that there is no bond energy for the monomer, the resulting free energy is

$$
\begin{aligned}
F_{\mathrm{monomer}} &= E_{\mathrm{monomer}} - TS_{\mathrm{monomer}} \\
&= -TS_{\mathrm{monomer}} \\
&= -kT \ln \left[\left(\frac{V}{2} - 1\right)(V - 1)\right].
\end{aligned}
$$

Figure (2.3) shows the free energy F for the monomers and the dimer as a function of temperature T. At low temperatures the dimer is more stable, while dissociation is favored at high T. At $T = T_0$, monomers and the dimer are equally stable.
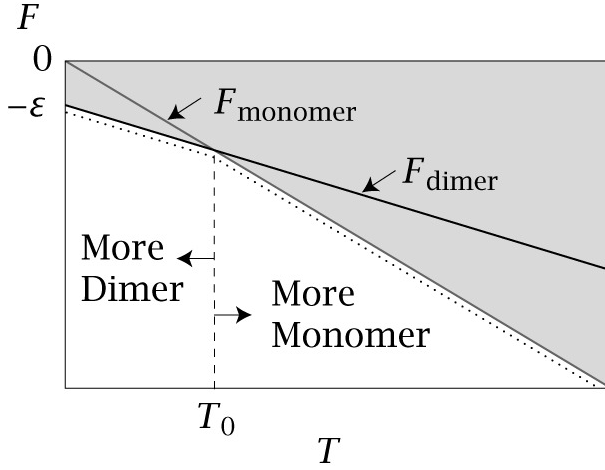
Figure 2.6:
Free energy F for the monomers and the dimer as a function of temperature T. The dotted line indicates if the system is in the dimer state (low temperatures) or rather in a monomer state (high temperatures). From [Dill and Bromberg, 2010].

2.4 Free energy landscapes

We have discussed in Chapter 1 for the example of an isolated molecule that the dynamics of nuclei can be illustrated in terms of the motion on a potential-energy surface (see Fig. 1.1). To generalize this idea to a system embedded in a heat bath, we now introduce the concept of a free energy surface or free energy landscape.

Reduced distribution functions

The probability distribution of N particles

$$P(\boldsymbol{r}) = P(\boldsymbol{r}_1, ..., \boldsymbol{r}_N)$$

does not factor in single-particle functions, because in general the potential energy $U(\boldsymbol{r})$ couples all the coordinates. Nonetheless, it is instructive to discuss distribution functions for a small number of particles. E.g., for two particles at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, we obtain the two-particle distribution

$$P(\boldsymbol{r}_1, \boldsymbol{r}_2) = \int d\boldsymbol{r}_3 ... \int d\boldsymbol{r}_N P(\boldsymbol{r}).$$

Usually, we are interested in the distribution function $\rho(\boldsymbol{r}_1, \boldsymbol{r}_2)$ for finding a particle (any one) at $\boldsymbol{r}_1$ and any other one at $\boldsymbol{r}_2$. Since there are N ways for picking the first particle, $N-1$ ways the second, we obtain

$$\rho(\boldsymbol{r}_1, \boldsymbol{r}_2) = N(N-1)P(\boldsymbol{r}_1, \boldsymbol{r}_2).$$

Furthermore, we may be interested in the one-particle density

$$P(\boldsymbol{r}_1) = \int d\boldsymbol{r}_2 ... \int d\boldsymbol{r}_N P(\boldsymbol{r}),$$

which is constant for a isotropic fluid, $P(\boldsymbol{r}) = \rho = N/V$. Since the particles are uncorrelated for an ideal gas, the two-particle distribution is

$$\rho(\boldsymbol{r}_1, \boldsymbol{r}_2) = \rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2) = \frac{N(N-1)}{V^2} \approx \rho^2.$$

As an important quantity to characterize liquids, we define the *radial distribution function* as

$$g(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{\rho(\boldsymbol{r}_1, \boldsymbol{r}_2)}{\rho^2}.$$

For an isotropic fluid, this function depends only on the distance $r = \mid \boldsymbol{r}_1 - \boldsymbol{r}_2 \mid$, yielding

$$g(r) = g(\mid \boldsymbol{r}_1 - \boldsymbol{r}_2 \mid) = \frac{N(N-1)}{\rho^2} P(\boldsymbol{r}_1, \boldsymbol{r}_2). \tag{2.26}$$

Note that

$$g(0) = 0 \quad \text{(for } r \leq \text{ diameter of particles)}$$
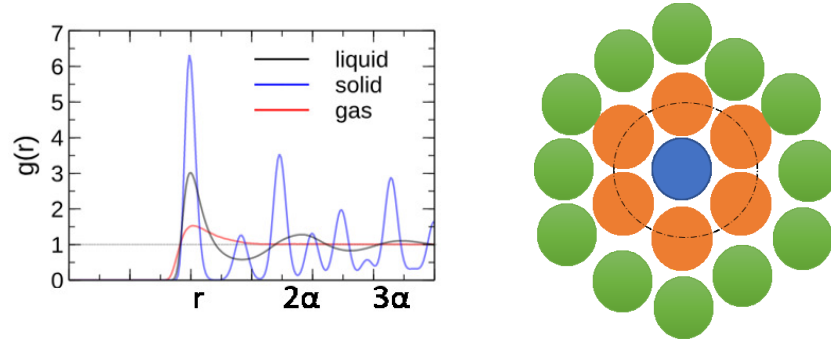$$g(r \to \infty) = 1 \quad \text{(for } r \gg \text{ range of particle-particle interaction)}$$

Figure 2.7: (Left) Simulated radial distribution function $g(r)$ for solid, liquid and gaseous argon, where $\alpha = 3.4\text{Å}$ represents the van der Waals diameter. While $g(r)$ is almost constant for gases, we see smooth maxima associated with solvation shells for liquids, and sharp peaks related to the crystal structure. (Right) Simple liquid structure, indicating first and second solvation shells. Sources: https://upload.wikimedia.org/wikipedia/commons/b/b8/ and https://nptel.ac.in/courses/103103036/module3/lec18/

Potential of mean force

We now want to make the connection between the reduced distribution function $g(r)$ and a free energy curve along coordinate r. To this end, we need to show that that $g(r)$ is related to reversible work $W(r)$ for a process in which two tagged particles are moved from infinite separation to distance r via

$$g(r) = e^{-\beta W(r)}.$$

Since $W(r) = W(r; \beta, \rho)$, the process is performed reversibly at constant NVT. Hence $W(r)$ is the change in Helmholtz free energy, $W(r) = F(r)$.

Proof: Consider the solvent averaged force between a pair of particles, say 1 and 2. Performing the average over all configurations with particles held fixed at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, the averaged force is

$$-\left\langle \frac{dU(\boldsymbol{r})}{d\boldsymbol{r}_1} \right\rangle_{\boldsymbol{r}_1, \boldsymbol{r}_2 \text{ fixed}} = -\frac{\int d\boldsymbol{r}_3...d\boldsymbol{r}_N \, (dU/d\boldsymbol{r}_1) \, e^{-\beta U}}{\int d\boldsymbol{r}_3...d\boldsymbol{r}_N \, e^{-\beta U}} = \frac{kT \frac{d}{d\boldsymbol{r}_1} \int d\boldsymbol{r}_3...d\boldsymbol{r}_N \, e^{-\beta U}}{\int d\boldsymbol{r}_3...d\boldsymbol{r}_N e^{-\beta U}}$$

$$= kT \frac{d}{d\boldsymbol{r}_1} \ln \int d\boldsymbol{r}_3...d\boldsymbol{r}_N \, e^{-\beta U}.$$

Using that

$$\ln g(\boldsymbol{r}_1, \boldsymbol{r}_2) = \ln \left[\frac{1}{\rho^2} N(N-1) \int d\boldsymbol{r}_3...d\boldsymbol{r}_N \, \frac{e^{-\beta U}}{\int d\boldsymbol{r}_1...d\boldsymbol{r}_N \, e^{-\beta U}} \right]$$

$$= \ln \int d\boldsymbol{r}_3 - d\boldsymbol{r}_N e^{-\beta U} + \ln const.,$$

we obtain for the averaged force

$$\langle \boldsymbol{K} \rangle = -\left\langle \frac{d}{d\boldsymbol{r}_1} U(\boldsymbol{r}) \right\rangle_{\boldsymbol{r}_1, \boldsymbol{r}_2 \text{ fixed}} = \frac{d}{d\boldsymbol{r}_1} kT \ln g(\boldsymbol{r}_1, \boldsymbol{r}_2),$$

with potential of mean force $-kT \ln g(\boldsymbol{r}_1, \boldsymbol{r}_2)$. Hence the force between particles 1 and 2, averaged over the equilibrium distribution of all other particles, is given by the gradient of $-kT \ln g(|\boldsymbol{r}_1 - \boldsymbol{r}_2|) = -kT \ln g(r)$. Integration of the averaged force yields the reversible work $W(r) = F(r)$ which is equivalent to the free energy $F(r)$,

$$F(r) = -kT \ln g(r). \tag{2.27}$$

Hence the radial distribution function $g(r)$ represents a simple example of a free energy landscape (cf. Eq. (4.12)), with r being a reaction coordinate to describe the structure of liquids.
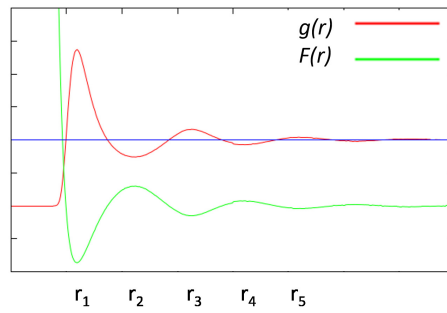
Figure 2.8: Radial distribution function $g(r)$ (red) and associated free energy curve (green). The minima at r_1, r_3, r_5 indicate metastable states, while the maxima at r_2, r_4 correspond to energy barriers.

Exercises

Exercise III: Entropy Production
Exercise IV: Free Energies