

PHYS-5500, Statistical Mechanics

Juha Javanainen

August 20, 2014

Contents

1	Introduction	5
1.1	Why this book	5
1.2	Practicalities (updated by S. Yelin)	6
2	Crash course on thermodynamics	7
2.1	Definitions	8
2.2	Laws of thermodynamics	12
2.2.1	Second law	13
2.2.2	Third law	14
2.2.3	Zeroth law	14
2.3	Building the machinery of thermodynamics	15
2.3.1	Thermodynamic potentials	15
2.3.2	Response functions	17
2.3.3	Varying particle number	18
2.4	Examples of thermodynamics	20
2.4.1	Classical ideal gas	20
2.4.2	Connection between heat capacities	22
2.4.3	Heat capacity at low temperature	23
2.5	Phase transitions	23
2.6	Local equilibrium and stability	25
2.7	Exercises	27
3	Statistical mechanics: preliminaries	33
3.1	Quantum versus classical	33
3.2	Demonstrating state counting	34
3.3	Simple statistical mechanics for spins	36
3.4	Exercises	38
4	Quantum mechanics primer	39
4.1	Topical review	41
4.2	Combining quantum systems	47
4.3	Density operator	49
4.4	Postulates of quantum mechanics revisited	53
4.5	Quantum operations	57

4.6	Relaxation terms in time evolution	61
4.7	Exercises	62
5	Quantum statistical mechanics	67
5.1	Ensembles in quantum statistical mechanics	68
5.1.1	Microcanonical ensemble	68
5.1.2	Canonical ensemble	71
5.1.3	Grand canonical ensemble	75
5.2	Exercises	78
6	Ideal quantum gases	81
6.1	State counting and statistics	81
6.2	Grand partition function	82
6.3	Ideal Fermi gas	84
6.3.1	High temperature/low density limit	86
6.3.2	Low temperature/high density	87
6.4	Ideal Bose gas	89
6.4.1	Bose-Einstein condensation	89
6.5	Massive ideal gases	91
6.5.1	Conduction electrons in a metal	92
6.5.2	Bose-Einstein condensation	92
6.6	Ideal gases of bosonic excitations	93
6.6.1	Photons	94
6.6.2	Phonons	96
6.7	Exercises	98
7	Classical limit of statistical mechanics	103
7.1	Probabilities and classical states	104
7.2	Going to the classical limit	107
7.2.1	Classical statistical mechanics	108
7.2.2	Few-particle distribution functions	110
7.3	Kinetic theory	113
7.3.1	Boltzmann equation	113
7.4	Exercises	118
8	Interacting gas	125
8.1	Virial expansion	125
8.2	Exercises	129
9	Mean field theories	131
9.1	Model of magnetic system	131
9.1.1	Ising model	131
9.1.2	Spontaneously broken symmetry	135
9.2	van der Waals equation of state	137
9.3	Exercises	141

10 Fluctuations and response	145
10.1 Linear response	145
10.1.1 Perturbation theory	146
10.2 Correlation functions	155
10.2.1 Scattering measures correlation functions	156
10.3 Fluctuation-dissipation theorem	158
10.4 Exercises	161

Chapter 1

Introduction

1.1 Why this book

What you have on hand is a highly sketchy and experimental on-the run textbook for PHYS-5500, Statistical Mechanics by Juha Javanainen. There are dilemmas that have led to this attempt at a book.

First, there are two different general styles to approach statistical mechanics. One might start counting quantum states, then develop quantum statistical mechanics and thermodynamics on such a foundation. This is the approach of the **Kittel-Kroemer** textbook. The most traditional track is to do thermodynamics first, then classical statistical mechanics, then quantum statistical mechanics. **Huang's** textbook is organized in this manner. The problem is, I am in favor of yet another scheme: First do thermodynamics. It may be formulated “axiomatically” based on a few fairly obvious empirical facts, gives a surprising number of answers in its own right, and keeps on working perfectly even in conjunction with quantum statistical mechanics. Then it is time for quantum statistical mechanics, and finally comes classical statistical mechanics. The motivation for the reordering is that classical statistical mechanics may be derived from quantum statistical mechanics, but an attempt to formulate classical statistical mechanics without the intervening quantum statistical mechanics ends up in deep conceptual trouble. For instance, how do you count the (uncountable) states of your usual classical-mechanics system?

Second, statistical mechanics even at the level of a core course is a potentially very broad subject. We used to have two courses for statistical mechanics and related matters, which got pared down to one. What is needed here is a precise selection of topics and an organization of the book so that it efficiently supports the selection of topics.

I do not know of any textbook in statistical mechanics that supports my reordering of the material or the topics. Hence this book.

1.2 Practicalities (updated by S. Yelin)

This course is supposed to run MW 10:30-11:45 here in P-121. I have no office hours, but I am available most of Monday and Wednesday. It might be a good idea to check via email whether I am around. My email is syelin@phys.uconn.edu, and the office and cell phones are 6x4899, 617-669-2471. You are encouraged to contact me any way that works for you — but when in doubt, call!

The course is defined by the class material. From Juha, I have inherited a rather tightly laid path through what, I think, are the core concepts of statistical mechanics as it is *applied* by physicists; people who study statistical mechanics for a living are a different breed altogether. The main exception is the section on quantum mechanics. This not only introduces the basic idea of the density operator as needed in statistical mechanics, but on the side also presents a romp through quantum mechanics as it is currently used in quantum information theory.

The course will have two what I call midterms, each 25% of the grade, and a final worth 50%. There is also **mandatory** homework, an assignment every week. Initially the homework is not graded, but if in my opinion the students do not give it their best shot (e.g., if the turnout drops to under 90%), all or part of the homework will be graded and the total grading formula will be changed in some manner. I will keep your homework solutions, so **make a copy for yourselves before turning them in!**

This book shows a tendency toward abstract reductionism, so that on the basis of the book (and the lectures) alone it may not be very easy to get started with statistical mechanics problems. Correspondingly, the homework is an essential part of the course – in particular, if you plan to pass the StatMech prelim exam. It has evolved into the present form over many years and is tightly tied to the lectures, but I do not shy away from **introducing new concepts in the problems** as well. At the end of each chapter this book has problems, for the most part the same ones that have been circulated in homework and tests for the past five years. Some of them are trivial, others are harrowing. At any rate, homework solutions (which I sketch in class, to an extent that depends on the success of the students in each problem) are in circulation. This is one reason why I am reluctant to grade; the grade for homework would be unfair. However, I will grade if I have to. Also, please do **your** best with the homework; a more or less verbatim copy of any last year's solution does not reflect well on you, nor does it provide the necessary practice for tests and prelim exams.

A warning (by J. Javanainen), born from experience: *I need to conclude with a warning. Although the subject matter of this course is nothing particularly difficult, students in their first or even second go-around in statistical mechanics will not see it that way. There is a seeming ad-hoc character to this discipline that makes it very hard for a beginner to apply even the simplest principles to new situations. If it is any consolation, with enough experience such problems will go away. In the interim, however, it is extremely important that you **stay constantly up-to-date with the course.** You lose it, you lose it good!*

Chapter 2

Crash course on thermodynamics

Thermodynamics is a discipline that can be postulated on the basis of empirical observations, and also “derived” from statistical mechanics. Either way, experience shows that thermodynamics and statistical mechanics can, and frequently will, be mixed freely. Thermodynamics is therefore a standard tool even in most sophisticated statistical mechanics considerations, and everyone should be familiar with the basics.

For the present purposes we regard thermodynamics as an *axiomatic* theory that gives *relations* between observable properties of *macroscopic* bodies. The qualifications are important. Thermodynamics is traditionally applied to objects that are large enough to be seen and detected without the aid of any instruments. Thermodynamics gives relations between measurable properties, but will not predict the value of the measurable quantities from a microscopic foundation. That is the domain of statistical mechanics. Hence, to start cranking the machinery of thermodynamics, you need some values either from measurements, or from statistical mechanics. Finally, for the present purposes thermodynamics is an abstraction extracted from observations. One could, in principle, build a mathematics style structure with definitions, axioms, propositions, lemmata, proofs, and whatnot. We go to some extent in this direction, but not all the way; full rigor would make the scheme hard to follow. Be warned that the exposition here deliberately has some logical flaws, such as forward references and assumptions of preexisting knowledge about the subject matter. Things simply are easier that way.

2.1 Definitions

State variable, a.k.a. thermodynamic parameter

A *state variable* or *thermodynamic parameter* is a measurable macroscopic property of the system, such as volume (V), pressure (p), temperature (T), entropy (S), particle number (N) [ahem, measurable], chemical potential (μ), internal energy (U), and so on. A crucial observation is that these parameters come in two types, *extensive* and *intensive*:

extensive: V, S, N, U, \dots

intensive: p, T, μ, \dots

The mental image here is that if we attempt to enlarge or shrink the system, a different-size copy that looks most like the original system is such that upon the change of size the extensive variables increase or decrease directly proportionally to the number of particles, and the intensive variables remain constant.

The distinction between extensive and intensive does not always apply straightforwardly as stated. For instance, think of the gas in the atmosphere bounded by gravity. What is the volume of the atmosphere in the first place? And if you double the number of air molecules under the same gravity, keeping the temperature constant, the atmosphere will still reach approximately as high or low up as before, so that the volume is basically unchanged. However, if you take a volume of air such that gravity has no substantial effect on the air density across the volume, then the concepts of extensive and intensive variables still apply *locally* in that particular volume. Also, if the property of being intensive or extensive applies in the first place, it is a property of the variable that does not depend on the system or its particular state.

Of course, the state of a system is, for the present purposes, construed as the collection of the values of those state variables that have a well-defined value.

Thermodynamic equilibrium

We know from experience that, when a macroscopic system is left alone, its time evolution (if any) eventually ceases. Thereafter, in the absence of changes in the external conditions acting on the system, the system shows no time dependence, at least on macroscopic size scales. Such a steady state is called *thermodynamic equilibrium*. This statement also has an important alternative form: To effect a change of state in an equilibrium system, you have to change the external conditions on the system.

In thermodynamics one occasionally also needs to think of nonequilibrium systems, a sticky proposition as virtually all of the mathematical machinery of thermodynamics only applies to the equilibrium. The usual mental image that, nonetheless, allows us to consider nonequilibrium systems runs as follows. We assume that the system is broken up to pieces, small but still large enough to warrant a thermodynamics discussion, and assume that each piece is in equilibrium in the sense that, if placed in a proper environment, there would be no time evolution. It just so happens that the other pieces of the system, while in

themselves also equilibrium systems in this same sense, do not constitute the proper environment. So, both the pieces and the system as a whole keep on evolving.

A frequent sign of a nonequilibrium system is the presence of currents. Consider, for instance, a solid block whose temperature is not constant throughout the block. We know from experience that in equilibrium the temperature would be constant throughout the block. The way the system responds to the nonequilibrium state is to set up heat currents that attempt to equilibrate the temperature everywhere to the same value.

Equation of state

Experience tells us that in an equilibrium system there are usually only very few thermodynamic parameters that can be varied independently. This is because there are relations, called *equations of state*, that the variables have to satisfy. An example is given by Boyle's law for the ideal gas

$$pV = Nk_B T; \text{ or } pV - Nk_B T = 0. \quad (2.1)$$

Here k_B (often just the plain k) is called the Boltzmann constant, and equals $k_B = 1.38 \times 10^{-23} \text{ J/K}$. In thermodynamics and statistical mechanics temperature is, by default, expressed on an absolute scale, such that the melting point of ice at atmospheric pressure equals 273.15 K.

Given the concept for the equation of state, we may now give a mathematical definition for the extensive and intensive variables. Thus, let A_1, \dots, A_n be extensive variables and a_1, \dots, a_m be intensive, and let

$$f(A_1, \dots, A_n; a_1, \dots, a_m) = 0 \quad (2.2)$$

be a valid equation of state, then

$$f(\lambda A_1, \dots, \lambda A_n; a_1, \dots, a_m) = 0 \quad (2.3)$$

also holds for every $\lambda > 0$. Boyle's law gives an explicit example.

Thermodynamic transformation; reversible transformation

The phrase *thermodynamic transformation* simply means that the state, of the system, i.e., measurable values in the system, changes; or even does not change if the transformation happens to be the trivial identity transformation. In thermodynamics particularly important are the reversible transformations. In a *reversible transformation* the system remains in an equilibrium during the entire change of the state. In mathematical terms this means that the equations of state are satisfied during the entire transformation. The analysis of a reversible transformation may therefore be based on the equations of state.

Of course, if the external conditions on the system remain fixed, by the very definition the state of an equilibrium system does not change. The only way to effect a nontrivial reversible transformation is to change the external

conditions on the system. The hallmark of a reversible transformation is that if the variation of the external conditions were suddenly halted, the evolution of the system would instantly cease. Reversible transformations generally result if the external conditions are varied “slowly enough,” whatever that may mean in practice.

In equilibrium there are only a few independent state variables. For instance, p and V are sufficient to determine also T , and it turns out, all other relevant thermodynamic variables, for a given amount (fixed N) of an ideal (monatomic) gas. This means that an equilibrium state of the given amount of ideal gas is completely specified by giving the position of the system in the pV plane. As all states in a reversible transformation are equilibria, an arbitrary reversible transformation between an initial state A and the final state B is also completely specified by the trajectory (path) of the system in the pV plane.

Let now F stand for an arbitrary state variable. Since F is a function of the state only, its value in the initial and final equilibrium states are unique, and moreover the change of F over a path from A to be B is independent of the path:

$$F(B) - F(A) = \int_C dF, \quad (2.4)$$

no matter what path C .¹ This even holds if the transformation is not reversible, as long the initial and final states are equilibria and the value of the variable exists throughout the transformation. Moreover, the reverse holds: If the integral of the right-hand side of (2.4) is independent of the path for any two given endpoints, the right-hand side defines a quantity F that satisfies the left-hand side of the equation uniquely up to an additive constant.

As an example, it was noted long ago that in a reversible process the heat brought into the system (dQ) and the temperature (T) satisfy the peculiar condition that $\int dQ/T$ is independent of the path between the initial and final state of the transformation. The integral

$$S(B) - S(A) = \int_A^B \frac{dQ}{T} \quad (2.5)$$

therefore defines a state variable S called *entropy* uniquely up to an overall additive constant. Entropy does not have an obvious mechanistic meaning such as, say, energy has, and to a physicists it may therefore appear somewhat mysterious. However, to someone designing steam plants, entropy is just as tangible and real as energy.

A process that is not reversible is called *irreversible*. Now, even if the machinery of thermodynamics only really works with reversible processes and many if not most important processes in technology and nature are irreversible, all is not lost. Often the initial and final states of interest are still equilibria, even if somewhere along the line the system goes irreversible. A physicist or engineer

¹The curve C must be sufficiently regular, or course. From now on we rarely mention this type of mathematical fine points unless they may actually become issues in physics.

analyzing the situation then imagines (makes up!) a reversible path and calculates, which will give all relations between the initial and final states correctly.

From now on, unless otherwise stated, all thermodynamic processes are assumed to be reversible.

Sidebar on integrability

In physics it often occurs that we are supposed to know alleged derivatives $g_1(x_1, \dots, x_n)$, $g_2(x_1, \dots, x_n)$, \dots , of a quantity $F(x_1, \dots, x_n)$ with respect to the variables x_1, \dots, x_n . Well, are g_i partial derivatives? And if so, how can one construct the corresponding quantity F from the derivatives?

Suppose that, indeed, there exists an F such that

$$g_i = \frac{\partial F}{\partial x_i}. \quad (2.6)$$

Mathematical pathologies aside, one must be able to exchange the order of derivatives in the second derivative, so that such partial derivatives must satisfy

$$\frac{\partial g_i}{\partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i} = \frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial g_j}{\partial x_i} \quad (2.7)$$

for all i and j .

But this necessary condition is also sufficient. Suppose one has a set of functions $g_i(x_1, \dots, x_n)$, $i = 1, \dots, n$, such that for all $i, j = 1, \dots, n$ the *integrability conditions*

$$\frac{\partial g_i}{\partial x_j} = \frac{\partial g_j}{\partial x_i} \quad (2.8)$$

hold true. It then turns out that the integral

$$I(\mathbf{x}, \mathbf{x}_0) = \int_{\mathbf{x}_0}^{\mathbf{x}} (g_1 dx_1 + \dots + g_n dx_n) \quad (2.9)$$

is independent of the (reasonably smooth) path connecting the points \mathbf{x} and \mathbf{x}_0 . Moreover, viewed as a function of the upper limit of the integration $\mathbf{x} = (x_1, \dots, x_n)$, this integral defines a function $F(x_1, \dots, x_n)$ such that

$$F(\mathbf{x}) - F(\mathbf{x}_0) = I(\mathbf{x}, \mathbf{x}_0), \quad (2.10)$$

and (2.6) hold true. Obviously, by fixing the value of the function F at one point \mathbf{x}_0 , we have fixed the value of F at all points in the space of the arguments \mathbf{x} that can be connected by a (smooth enough) path. In short, if the integrability conditions (2.8) hold, there exists a function F , unique up to an additive constant, such that g_i are the partial derivatives of F with respect to x_i .

This scheme is not as out-of-the blue as it may seem. For instance, we know from vector calculus as applied in mechanics or electricity and magnetism that if a vector field in three spatial dimensions satisfies

$$\nabla \times \mathbf{F}(\mathbf{x}) = 0, \quad (2.11)$$

then there exists a potential $V(\mathbf{x})$ such that

$$\mathbf{F} = -\nabla V. \quad (2.12)$$

But the condition that the vector field is irrotational (2.11) is precisely the same as requiring

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}, \quad \frac{\partial F_y}{\partial z} = \frac{\partial F_z}{\partial y}, \quad \frac{\partial F_z}{\partial x} = \frac{\partial F_x}{\partial z}, \quad (2.13)$$

which is the three-dimensional form of the integrability conditions (2.8), and the potential $V(\mathbf{x})$ is defined up to an additive constant by a path integral so that its change between two points is

$$V(\mathbf{x}) - V(\mathbf{x}_0) = - \int_{\mathbf{x}_0}^{\mathbf{x}} d\mathbf{x} \cdot \mathbf{F}(\mathbf{x}), \quad (2.14)$$

independently of the path between the points \mathbf{x}_0 and \mathbf{x} .

2.2 Laws of thermodynamics

The definitions out of the way, let us now lay down the axioms of thermodynamics.

First law

The first law could be stated as thus: There exists a state variable, called *internal energy* and denoted by U , such that in an infinitesimal transformation

$$dU = dQ - dW, \quad (2.15)$$

where dQ is the heat put in to the system, and dW is the work done by the system on the external world. If the transformation is not infinitesimal, you may imagine that it is put together from a large number of small transformations, and accordingly “integrate” the differential form of the first law.

The first law is simply a statement of the conservation of energy. In the usual case of reversible processes, we have

$$dQ = T dS. \quad (2.16)$$

The work term in the first law is of the form $dW = a_1 dA_1 + \dots + a_n dA_n$, where a_1, \dots, a_n and A_1, \dots, A_n are appropriate intensive and extensive variables, respectively. What these variables are depends on the nature of the system. For the “ pV ” system, say, gases, the work term reads

$$dW = p dV. \quad (2.17)$$

For a magnetic system one sees all kinds of dimensional hacks for the work term, but for a uniform system where the applied magnetic field \mathbf{H} and the ensuing magnetization \mathbf{M} are parallel one might accurately write

$$dW = -\mu_0 H V dM, \quad (2.18)$$

where μ_0 is the permeability of the vacuum and V is the (assumedly constant) volume of the system. If we, on the other hand, were to focus on surface tension, the work term would involve the coefficient of surface tension σ and the area of the surface A as thus,

$$dW = -\sigma dA. \quad (2.19)$$

The signs in the work term are often black magic. Many a time the best way to get the sign right is to do the calculations with both signs until one or the other sign starts producing patently wrong results. . . .

2.2.1 Second law

A concise way of stating the second law is as follows: In any differential thermodynamic process $T dS \geq dQ$.

The practical meaning is that a nonequilibrium system tends to evolve toward equilibrium on its own, and this evolution flat out generates entropy. If you bring in heat to the system, the entropy will increase; but if the process is irreversible, entropy may increase more than it would as a result of the heat brought in from the outside in a reversible manner, $dS = dQ/T$.

Huang's book explains at length how one can deduce the form $T dS \geq dQ$ from very basic observations such as heat always flowing from a higher temperature to a lower temperature, how this observation may be used to define an absolute scale of temperature, what is a Carnot's engine, and so forth. Important as these considerations are for the conceptual development of thermodynamics, however, they are of little use in applications of thermodynamics in statistical mechanics. So, a only a brief discussion of the Carnot engine is included in P. 2.7.

For our purposes the second law is mainly an intellectual curiosity, but we would be amiss not to bring up some of the profound consequences of the second law. The second law is what really gives us the notion of the "arrow of time": time runs in the direction in which entropy grows. In a classic example, if we were to put a drop of (water-soluble) dye in a beaker full of water, the dye spreads out in the water. You never see the dye spontaneously condense into the drop from the water. Given this kind of cases of growing entropy, it would usually not be difficult to perceive if someone is running a movie backwards on you even if you do not see people walking backwards.

This arrow of time is something of a problem, in that the microscopic Hamiltonians, quantum mechanical or classical, which govern everyday life, are invariant under time reversal. Inasmuch as the glass beaker is an isolated system, a drop of dye condensing and jumping out of the water is an equally plausible solution to the equations of motion as the dye dropping in and getting diluted, but the former never happens. Every theorist worth his pay knows mathematical tricks to derive irreversible behavior from a reversible Hamiltonian, but not all agree that it is entirely clear what is going on conceptually.

The second law, incidentally, explains why reversible processes are called reversible. Let us consider gas in a cylinder whose walls are kept at a fixed

temperature T , and where the gas volume may be altered by pushing and pulling a piston. Suppose that the piston is released from an initial volume V_i and charges out very fast until some kind of a stopper puts a sudden end to the expansion of the gas at volume V_f , and then the gas is allowed to equilibrate at temperature T . If this process is irreversible, entropy gets generated above and beyond what would be expected of the heat exchange between the environment. In other words, the gas took less heat from the environment than it would have taken if it had transformed reversibly from the initial state (V_i, T) to the final state (V_f, T) . Now, after the gas has equilibrated, push back the piston so slowly that the gas reversibly returns to the initial state (V_i, T) . We will then have reversed an irreversible process, and done so by means of a reversible process. What gives?

The answer is that, in the reversible compression, all of the change of the entropy incurred in the expansion part of the cycle is returned to the environment as heat. As the process is reversible, more heat is returned to the *environment* than the system took up in the expansion part. The gas may have returned to its initial state, but the entropy of the environment has increased. Given that the system plus the rest of the universe presumably makes a closed system and cannot exchange heat with anything, for the universe we always have $dS \geq 0$. As a result of the irreversible part of the expansion the entropy of the universe has increased, and will never return to its original value. It is the universe as a whole that has irreversibly changed.

Our final consequence of the present form of the second law applies to an isolated system, one that neither does work on the environment nor exchanges heat or particles with it; the universe above is an example. Since there is no heat transferred to an isolated system $dQ = 0$, and in any thermodynamic transformation $dS \geq 0$. If there is any evolution at all in the isolated system, it cannot decrease the entropy. In particular, the system may have reached an equilibrium only by either increasing the entropy, or maybe keeping it constant. Hence, in equilibrium, the entropy of an isolated system is a maximum.

2.2.2 Third law

The third law says that when $T \rightarrow 0$, $S \rightarrow 0$. Please remember that the definition of entropy as $dS = dQ/T$ only fixes entropy up to an additive constant. Besides, if the system happens to have states that cannot be connected by a reversible process, then the state space is split into separate regions that cannot be connected by a reversible process, and the constant in entropy may even be different in different regions. The third law removes the ambiguities, and gives the absolute value of entropy.

2.2.3 Zeroth law

It is customary in thermodynamics to name a zeroth law as well. Evidently, two equilibrium systems are in equilibrium with each other if no time evolution ensues when the systems are allowed to interact (exchange heat, do work on one

another, maybe even exchange particles). Without further ado, the zeroth law says that if system A is in thermal equilibrium with system B and system B is in thermal equilibrium with system C , then A is also in thermal equilibrium with C .

From a mathematical viewpoint this means that “equilibrium” between two systems is an equivalence relation that splits all system into classes such that every system in a class is in equilibrium with every other member of the same class. This may be an important element in a true axiomatic development of thermodynamics, but it is also so obvious from a physical viewpoint that we usually would use the zeroth law without even noticing that we are doing so. We will say nothing further about the zeroth law.

2.3 Building the machinery of thermodynamics

In this and the following section we will run through a number of specific aspects of, and examples about, thermodynamics. We will at the same time develop a standard bag of tools in thermodynamics, and also demonstrate the basic mathematical techniques. The mathematics is deliberately basic, and we avoid memorized rules. Differential-calculus thinking permeates our presentation; derivatives are, in a very literal sense, quotients of small quantities.

2.3.1 Thermodynamic potentials

Internal energy U

In a pV system, in a reversible process, the differential of internal energy reads

$$dU = T dS - p dV. \quad (2.20)$$

This leads us to designate S and V as the *natural variables* of internal energy, and the default representation of writing the internal energy as a function of these variables as $U(S, V)$. It should be noticed, however, that using the equations of state, other thermodynamical variables could be substituted. For instance, writing $S = S(T, p)$ and $V = V(T, p)$, we could, if we so wished, write internal energy as a function of temperature and pressure,

$$U(T, p) = U(S(T, p), V(T, p)). \quad (2.21)$$

Now, take a thermodynamic transformation in which volume is held constant. Then we have

$$dV = 0; \quad \Rightarrow \quad dU = T dS \quad \Rightarrow \quad \frac{dU}{dS} \equiv \left(\frac{\partial U}{\partial S} \right)_V = T. \quad (2.22)$$

This is a variation at constant V , and leads to a ratio of differentials in which S is varied while V is kept constant. This is the partial derivative of U with respect to S at constant V . In thermodynamics substitution of variables can lead to a

great variety of partial derivatives, and you would almost always use subscripts in a partial derivative to indicate which variables are held constant. Errors in this regard tend to lead to nasty messed-up calculations and catastrophic scores.

In exactly the same way, one finds that

$$\left(\frac{\partial U}{\partial V}\right)_S = -p. \quad (2.23)$$

In order to avoid contorted notation, let us for the time being explicitly declare that we regard S and V as the two independent variables in our differential calculus. Then, in view of (2.22) and (2.23) and taking into account that the order of derivatives may be exchanged, we have

$$\left(\frac{\partial T}{\partial V}\right)_S = \frac{\partial}{\partial V} \frac{\partial U}{\partial S} = \frac{\partial^2 U}{\partial V \partial S} = \frac{\partial^2 U}{\partial S \partial V} = \frac{\partial}{\partial S} \frac{\partial U}{\partial V} = -\left(\frac{\partial p}{\partial S}\right)_V. \quad (2.24)$$

In short, we have what is known as a *Maxwell relation* associated with the internal energy,

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V. \quad (2.25)$$

Legendre transformations

Let us next use the so-called Legendre transformations to define further thermodynamic potentials:

$$\begin{aligned} F &= U - TS && \text{Helmholtz free energy;} \\ H &= U + pV && \text{enthalpy;} \\ G &= U - TS + pV && \text{Gibbs free energy.} \end{aligned}$$

The value of these transformations should be evident from the following observations that are derived for the Helmholtz free energy just like we did in the case of internal energy:

$$dF = dU - d(TS) = T dS - dT S - T dS - p dV = -S dT - p dV; \quad (2.26)$$

$$F = F(T, V); \quad S = -\left(\frac{\partial F}{\partial T}\right)_V, \quad p = -\left(\frac{\partial F}{\partial V}\right)_T \quad (2.27)$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V. \quad (2.28)$$

The analogous expressions for the differentials, expressions of thermodynamic quantities as derivatives of thermodynamic potentials, and Maxwell relations may be found similarly for enthalpy and Gibbs free energy. Derivations are left as an exercise; the results should eventually get memorized from much use. Analogous definitions and corresponding differential calculus apply equally well to other than pV systems.

Minimum property of thermodynamic potentials

Suppose that the external conditions fix temperature and pressure of a system, but even then the system may undergo a (possibly irreversible) transformation. Then we have to write the change of the Gibbs free energy as

$$dG = dQ - p dV - d(TS) + d(pV) = dQ - T dS - S dT + V dp \quad (2.29)$$

$$= dQ - T dS \quad (2.30)$$

$$\leq 0. \quad (2.31)$$

The second line follows because we assume that temperature and pressure are held constant; the third line comes from the second law. This argument is admittedly formal; we are on a dangerous ground trying to apply thermodynamics to nonequilibrium systems without a careful characterization of the state. However, the result is valid: When T and p are held constant, in any thermodynamic process the value of the Gibbs free energy can only decrease. In other words, steady state, when there is no change in Gibbs free energy, is always reached by decreasing the Gibbs free energy. In other words, the equilibrium state is the minimum of Gibbs free energy for the given temperature and pressure.

This result is a special case of the *minimum property of thermodynamic potentials*: Suppose the external conditions fix a thermodynamic variable x_1 from the set $\{S, T\}$ and another variable x_2 from the set $\{p, V\}$, then the equilibrium is the minimum of the thermodynamic potential whose natural variables are x_1 and x_2 .

2.3.2 Response functions

Apart from the state variables per se, thermodynamic systems have various *response functions* that tell us how the state changes when the system is subject to changes in the external conditions. Consider, for example, heat capacity. If heat is brought to the system, the temperature will rise. If volume is held constant in the process, the rise of the temperature is characterized by *heat capacity at constant volume*,

$$C_V = \left(\frac{\partial Q}{\partial T} \right)_V = T \left(\frac{\partial S}{\partial T} \right)_V. \quad (2.32)$$

Now, for a process at constant V , $dV = 0$, so that

$$dU = T dS \Rightarrow \frac{dU}{dT} = T \frac{dS}{dT} \Rightarrow \left(\frac{\partial U}{\partial T} \right)_V = T \left(\frac{\partial S}{\partial T} \right)_V = C_V. \quad (2.33)$$

Similarly, we have the heat capacity at constant pressure,

$$C_p = T \left(\frac{\partial S}{\partial T} \right)_p = \left(\frac{\partial H}{\partial T} \right)_p. \quad (2.34)$$

How the volume changes in response to changing pressure is characterized by *compressibility*

$$\kappa_x = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_x . \quad (2.35)$$

Here x denotes what is held constant when the pressure is changed; $x = T$ and $x = S$ denote, respectively, *isothermal* and *adiabatic* compressibilities. Also, when the temperature increases, most substances tend to expand as characterized by the *coefficient of thermal expansion*

$$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p . \quad (2.36)$$

In both cases there is the division by the volume to make compressibility and thermal expansion coefficient intensive quantities. Besides, when pressure increases, volume decreases; hence the minus sign in the definition of compressibilities to make them positive. Other than pV systems have analogous response functions, for instance, *magnetic susceptibility*

$$\chi = \frac{1}{\mu_0} \left(\frac{\partial M}{\partial H} \right)_x . \quad (2.37)$$

The subscript x could be T or S , which rarely matters for our purposes.

2.3.3 Varying particle number

Chemical potential

Suppose now that also the particle number in the system is a variable. Adding a particle to an already existing system may change the energy for various reasons. For instance, the added particle may bring with it kinetic energy, and it may interact with the already existing particles in the system. All of this is taken into account by amending the first law to read (for pV systems)

$$dU = T dS - p dV + \mu dN , \quad (2.38)$$

where the new intensive variable μ is known as *chemical potential*.

The appearance of a work term, μdN , is not an accident. In this context, at constant entropy and hence no heat coming in from the outside, the energy coming in with the new particles acts like work, and various forms of work can be interconverted into each other (in principle) without losses. For instance, just like it is possible to store energy into a gas by compressing it and later recover the work by letting the gas expand, theoretically it is possible to convert all of the “chemical” energy μdN added to the system to work done on the external world. Though this is part of the discussion of the Carnot engine that we have mostly skipped, the same does not go with heat; one of the classic formulations of the second law says that there is no thermodynamic transformation whose sole effect is to convert an amount of heat extracted from a reservoir to work.

Chemical potential and Gibbs free energy

With the new form of the first law, the Gibbs free energy still reads

$$G = U - TS + pV; \quad (2.39)$$

its differential simply picks up the added term, as in

$$dG = -S dT + V dp + \mu dN. \quad (2.40)$$

The natural variables are obviously T , p and N , so that one is inclined to write

$$G = g(T, p, N); \quad G - g(T, p, N) = 0. \quad (2.41)$$

Here G stands for the value of Gibbs free energy, and g denotes the functional dependence of the value on the values of the variables.

But, like other thermodynamic potentials, Gibbs free energy is extensive, so that for all $\lambda > 0$ also

$$\lambda G - g(T, p, \lambda N) = 0 \quad (2.42)$$

should hold true. Putting $\lambda = 1/N$, we have

$$G = Ng(T, p, 1) \equiv N\tilde{g}(T, p). \quad (2.43)$$

From the first law on one hand, and from Eq. (2.43) on the other, we have

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{T,p} = \tilde{g}, \quad (2.44)$$

so that we have the fundamental result

$$G = N\mu. \quad (2.45)$$

More in the same vein, by combining (2.43) and (2.45) we have the Gibbs-Duhem relation

$$U = TS - pV + N\mu. \quad (2.46)$$

The division of thermodynamic variables to intensive and extensive is not idle talk, but something maybe quite unexpected follows from it.

Grand potential

Given the increased number of independent variables, the thermodynamics formalism also gets much richer. One might, for instance, introduce Legendre transformations including particle number. Let us discuss the *grand potential* defined by

$$\Omega = F - \mu N = U - TS - \mu N. \quad (2.47)$$

The differential of grand potential is

$$d\Omega = -S dT - p dV - N d\mu, \quad (2.48)$$

so that the natural variables are T , V , and μ . Some thermodynamics variables expressed in terms of derivatives of grand potential read

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V, \mu}, \quad p = - \left(\frac{\partial \Omega}{\partial V} \right)_{T, \mu}, \quad N = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V}, \quad (2.49)$$

and there are no less than three Maxwell relations:

$$\left(\frac{\partial S}{\partial \mu} \right)_{T, V} = \left(\frac{\partial N}{\partial T} \right)_{V, \mu}, \quad \left(\frac{\partial p}{\partial \mu} \right)_{T, V} = \left(\frac{\partial N}{\partial V} \right)_{T, \mu}, \quad \left(\frac{\partial p}{\partial T} \right)_{V, \mu} = \left(\frac{\partial S}{\partial V} \right)_{T, \mu}. \quad (2.50)$$

Moreover, it follows from the Gibbs-Duhem relation that

$$\Omega = -pV. \quad (2.51)$$

The example of grand potential is not as gratuitous as it may seem. Later it will turn out that the grand potential is a key link between statistical mechanics and thermodynamics.

2.4 Examples of thermodynamics

The particular aim of this section is to demonstrate how much thermodynamics and simple mathematics allow one to extract out of virtually nothing. Momentarily the examples assume that the particle number is fixed.

2.4.1 Classical ideal gas

Ideal gas is, well, an idealization that applies to most gases in the limit of either high temperature or low density. Our model is an extrapolation of two empirical observations:

- (i) Boyle's law, $pV = NkT$.
- (ii) Heat capacity of an ideal gas is $C_V = xNk$. Here x is a constant that depends on the nature of the gas. For a monatomic ideal gases $x = \frac{3}{2}$ applies, while for a diatomic gas (at high enough temperature) we have $x = \frac{5}{2}$.

By virtue of the definition of heat capacity find

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V = xNk \Rightarrow \left(\frac{\partial S}{\partial T} \right)_V = \frac{xNk}{T}. \quad (2.52)$$

Moreover, from the Maxwell relation for the Helmholtz free energy and from Boyle's law we have

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial p}{\partial T} \right)_V = \frac{\partial}{\partial T} \left(\frac{NkT}{V} \right) = \frac{Nk}{V}. \quad (2.53)$$

We therefore know the partial derivatives of entropy with respect to both of the two free variables of our theory, temperature and volume. Moreover, dispensing

with the subscripts denoting constants for this argument,

$$\frac{\partial}{\partial V} \left(\frac{\partial S}{\partial T} \right) = 0 = \frac{\partial}{\partial T} \left(\frac{\partial S}{\partial V} \right) \quad (2.54)$$

holds true. This is the integrability condition, so that the derivatives (2.53) and (2.54) are consistent and in fact define a function $S = S(T, V)$ up to an additive constant. The function is found by integrating the expression

$$dS = \left(\frac{\partial S}{\partial V} \right)_T dV + \left(\frac{\partial S}{\partial T} \right)_V dT \quad (2.55)$$

from an initial point (T_0, V_0) to the final point (T, V) along an arbitrary path in the TV plane.

In this case anything else except paths made of sections with constant T and V would be rank lunacy, so we integrate

$$S(T, V) - S(T_0, V_0) = \int_{T_0, V_0}^{T_0, V} dS + \int_{T_0, V}^{T, V} dS \quad (2.56)$$

$$= \int_{T_0, V_0}^{T_0, V} \left(\frac{\partial S}{\partial V} \right)_T dV + \int_{T_0, V}^{T, V} \left(\frac{\partial S}{\partial T} \right)_S dT \quad (2.57)$$

$$= NK[\ln V + x \ln T] + K. \quad (2.58)$$

Only the dependence of the entropy upon the variables T, V is of interest, so we have lumped everything else into an undetermined constant K . Incidentally, with $T \rightarrow 0$ it is not possible to have $S \rightarrow 0$, so the result violates the third law. Obviously, there is something seriously wrong with the classical ideal gas when the temperature approaches zero, which is pretty much as expected.

With very little work we may also find the internal energy. First note that

$$\left(\frac{\partial U}{\partial T} \right)_V = C_V = xNk. \quad (2.59)$$

Second,

$$dU = T dS - p dV; \quad \frac{dU}{dV} = T \frac{dS}{dV} - p, \quad (2.60)$$

which at constant temperature and using the previously derived expression for a relevant partial derivative (2.53) becomes

$$\left(\frac{\partial U}{\partial V} \right)_T = T \left(\frac{\partial S}{\partial V} \right)_T - p = T \frac{Nk}{V} - \frac{NkT}{V} = 0. \quad (2.61)$$

We now have the partial derivatives of U with respect to the variables T and V , which are once more compatible and can be integrated. In fact, since U only depends on the variable T , we see straight away that the result is

$$U = xNkT + K', \quad (2.62)$$

where K' is another constant that our formalism so far does not allow us to determine.

Faulty as our results obviously are at temperatures approaching zero, they nonetheless *completely* determine the thermodynamics of the model ideal gas. The unknown constants are an annoyance. Besides, the formalism also suffers from a problem known as the Gibbs paradox: Since V and N are both extensive quantities, the term in entropy $\propto N \log V \propto N \log N$ is *not* extensive. But within these limits, every observable of an ideal gas may be computed. Gibbs paradox rears its ugly head when one attempts to combine separate quantities of gas (change N) and the unknown coefficients become a problem in certain circumstances involving chemical potential. A true remedy must await until we have solid quantum statistical mechanics at our disposal.

2.4.2 Connection between heat capacities

Let us by recalling the definitions of the heat capacities at constant pressure and volume,

$$C_p = T \left(\frac{\partial S}{\partial T} \right)_p, \quad C_V = T \left(\frac{\partial S}{\partial T} \right)_V. \quad (2.63)$$

Next, using equations of state to substitute variables and applying the chain rule for derivatives we have

$$S = S(T, V) = S(T, V(T, p)); \quad \left(\frac{\partial S}{\partial T} \right)_p = \left(\frac{\partial S}{\partial T} \right)_V + \left(\frac{\partial S}{\partial V} \right)_T \left(\frac{\partial V}{\partial T} \right)_p, \quad (2.64)$$

and multiplication by T gives

$$C_p = C_V + T \left(\frac{\partial S}{\partial V} \right)_T \left(\frac{\partial V}{\partial T} \right)_p. \quad (2.65)$$

The last partial derivative is essentially the coefficient of thermal expansion, but the other factor is not related in any obvious way to known response functions.

We have to find a nonobvious relation. So, writing

$$S = S(T, p) = S(T, p(T, V)), \quad (2.66)$$

we have

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial S}{\partial p} \right)_T \left(\frac{\partial p}{\partial V} \right)_T = \frac{\left(\frac{\partial S}{\partial p} \right)_T}{\left(\frac{\partial V}{\partial p} \right)_T} = \frac{-\left(\frac{\partial V}{\partial T} \right)_p}{\left(\frac{\partial V}{\partial p} \right)_T}. \quad (2.67)$$

The third step is obvious when one considers derivatives as ratios of small quantities, and the fourth one follows from the Maxwell relation that originates from

Gibbs free energy. Another thermal expansion coefficient and isothermal compressibility are making an appearance. With this identification, Eqs. (2.65) and (2.67) and combine to give a relation between the heat capacities:

$$C_p - C_V = \frac{TV\alpha^2}{\kappa_T}. \quad (2.68)$$

One might surmise that the heat capacities are a priori unrelated, but it is not so. The response functions are closely tied together, ultimately because there are so few (2) independent thermodynamic parameters in the problem.

2.4.3 Heat capacity at low temperature

Let us begin with

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V; \quad \left(\frac{\partial S}{\partial T} \right)_V = \frac{C_V}{T}. \quad (2.69)$$

Integrating the latter expression over a contour of constant volume, the third law dictates that

$$S(T, V) - S(0, V) = \int_0^T \left(\frac{\partial S}{\partial T} \right)_V dT = \int_0^T \frac{C_V(T, V)}{T} dT \rightarrow 0 \text{ as } T \rightarrow 0. \quad (2.70)$$

This is impossible unless $C_V(T, V) \rightarrow 0$ as $T \rightarrow 0$. According to the third law, the heat capacity must vanish at zero temperature.

As esoteric the third law may seem, tangible results follow from it. Apart from another indication that something is not quite right with the classical ideal gas with the once-and-for-all constant heat capacity, we now know that all heat capacities must go to zero at low temperature. Other response functions have similar limiting properties.

2.5 Phase transitions

We all know that at atmospheric pressure water turns from liquid to solid, freezes, when cooled below 0°C , and from liquid to gas, boils, when heated above 100°C . These are examples of *phase transitions*, in which the state of the matter typically abruptly changes when some thermodynamic parameter is varied. Another phase transition frequently cited in physics is the ferromagnetic phase transition: Many materials, often but not necessarily involving iron, magnetize spontaneously when cooled below a *Curie temperature*.

In general, phase transitions abound in condensed matter systems, and classification of the phases and of the transition tends to be a major endeavor in its own right. We, however start modestly with a general thermodynamics discussion.

Condition for phase equilibrium

In an everyday phase transition such as boiling of water the environment typically sets the temperature T and pressure p of the system. We then know from the minimum property of thermodynamic potentials that in equilibrium the Gibbs free energy is a minimum. On the other hand, we also know that at the phase equilibrium point (T, p) the particles may be moved freely between the phases; at the same temperature and pressure the water may be all liquid, all gas, or anything in between. All partitions of the atoms between gas and liquid are equally equilibrium situations.

Suppose now that we move particles between the equilibrium phases I and II, so that for the individual phases $dN_I = -dN_{II} \equiv dN$. Since the phase equilibrium is a minimum of Gibbs free energy, first-order variation of the total Gibbs free energy in this transformation vanished, and we have

$$d(G_I + G_{II}) = \left[\left(\frac{\partial G_I}{\partial N} \right)_{p,T} - \left(\frac{\partial G_{II}}{\partial N} \right)_{p,T} \right] dN = (\mu_I - \mu_{II}) dN = 0, \quad (2.71)$$

so that the condition for phase equilibrium reads

$$\mu_I = \mu_{II}. \quad (2.72)$$

The chemical potentials may, and will, depend on temperature and pressure, so that in a more complete form, $\mu_I(T, p) = \mu_{II}(T, p)$. But this is an equation between temperature and pressure, from which one may solve either $p = p(T)$ or $T = T(p)$. The solution defines a line called *coexistence curve* in the Tp plane. The coexistence curve separates the regions in the pV plane containing the phases *I* and *II*. As the example of water shows, the same substance may have several phases, so that there may be several coexistence curves in, say, the Tp plane. A drawing presenting the coexistence curves and naming the phases is called a *phase diagram*. Momentarily, adding a phases diagram to this book takes more time than I have, so I do not draw here, but just add a listing of the terms that are most relevant in the solid-liquid-gas system: liquid-solid and gas-liquid coexistence curves, sublimation curve, triple point, and critical point.

Classification of phase transitions

At the transition between the two phases $\mu_I(T, p) = \mu_{II}(T, p)$, so that when T , or p , or both, are varied across a phase transition, the chemical potential, i.e., Gibbs free energy per particle, varies continuously. The same does not have to apply to the derivatives of Gibbs free energy, however. Let us define the first derivatives of the Gibbs free energy

$$s = - \left(\frac{\partial \mu}{\partial T} \right)_p = - \frac{1}{N} \left(\frac{\partial G}{\partial T} \right)_{p,N} = \frac{S}{N}, \quad (2.73)$$

$$v = \left(\frac{\partial \mu}{\partial p} \right)_T = \frac{1}{N} \left(\frac{\partial G}{\partial p} \right)_{T,N} = \frac{V}{N}; \quad (2.74)$$

in fact, entropy per particle and volume per particle. Even if the Gibbs free energy is per particle is continuous across a phase transition, it may be that one or both of these first derivatives is not. If this is the case, the phase transition is labeled *first order*; otherwise the phase transition is *continuous*.

Suppose now that particles are moved reversibly across a first-order phase transition in which entropy per particle changes. In a reversible process $dQ = T dS$, so that the corresponding change of entropy requires that the system takes up or gives off heat when particles move across the phase transition. This phenomenon is called *latent heat*. The latent absorbed in a phase transition from phase *I* to phase *II* is, per particle, is

$$\Delta q = T(s_{II} - s_I). \quad (2.75)$$

Similarly, the volume per particle changes by the amount

$$\Delta v = v_{II} - v_I. \quad (2.76)$$

In a gas-liquid-solid (pVT) system first-order phase transitions exhibit either latent heat, or change of volume per particle, or both. Obviously the ice-water phase transition requires a lot of heat, which keeps scotch on the rocks cold. Equally obviously, the volume of water changes by a lot when water boils at atmospheric pressure. Ice-water and water-steam phase transitions are first order, except that the latter actually is continuous at the critical point. On the other hand, at zero applied magnetic field, the prototypical ferromagnetic phase transition is continuous.

Clausius-Clapeyron relation

On a coexistence curve $\mu_I(T, p) = \mu_{II}(T, p)$ holds true. Let us now assume that we move a small distance, displacements dT and dp along the coexistence curve. Since we are supposed to stay on the coexistence curve, the changes of chemical potential must be equal, and we have

$$d\mu_I = \left(\frac{\partial \mu_I}{\partial T}\right)_p dT + \left(\frac{\partial \mu_I}{\partial p}\right)_T dp = d\mu_{II} = \left(\frac{\partial \mu_{II}}{\partial T}\right)_p dT + \left(\frac{\partial \mu_{II}}{\partial p}\right)_T dp. \quad (2.77)$$

In view of Eqs. (2.73), (2.74), (2.75) and (2.76) we have the *Clausius-Clapeyron equation*

$$\frac{dp}{dT} = \frac{\Delta q}{T \Delta v}. \quad (2.78)$$

The latent heat and volume change determine the slope of the coexistence curve.

2.6 Local equilibrium and stability

We begin the study of local equilibrium and stability of a system by considering an isolated container that is split in two parts *A* and *B* by a “separator” that

moves freely, and lets through both heat (energy) and particles. The separator, as specified so far, does not really separate at all, but we have introduced it as an auxiliary for our argument. We assume that both sides, per se, are in equilibrium, and ask for the conditions under which the sides are also in equilibrium with each other.

By the assumption that the box is completely isolated, the total volume, particle number, and internal energy are fixed, but the corresponding quantities on the A and B sides are allowed to vary. Let us, in fact assume that, say, internal energy is exchanged between the sides. Since $U_A + U_B = U$ is fixed, the possible variations are of the form $dU_A = -dU_B \equiv \delta U$, and likewise for V and N . U , V and N are three independent thermodynamic variables, and so the respective values on the A and B sides completely determine the equilibrium states for the sides. All told, the change of total entropy upon an allowed variations of the state up to first order in the small variations is

$$\delta S = \delta S_A + \delta S_B = \left(\frac{\partial S_A}{\partial U_A} \right)_{V_A, N_A} dU_A + \dots \left(\frac{\partial S_B}{\partial U_B} \right)_{V_B, N_B} dU_B, \quad (2.79)$$

a litany of six terms of this type.

In an isolated system the equilibrium is a maximum of the entropy. Correspondingly, when our entire system is in equilibrium, the total entropy is a maximum and does not change to first order in small variations of the state. Moreover, taking into account the constraints in the changes of energy, etc., on both sides, we have

$$0 = \left[\left(\frac{\partial S_A}{\partial U_A} \right)_{V_A, N_A} - \left(\frac{\partial S_B}{\partial U_B} \right)_{V_B, N_B} \right] \delta U + \dots, \quad (2.80)$$

where we have skipped exactly analogous term with δV and δN . Since the variations δU , δV and δN are independent, we have three separate conditions:

$$\left(\frac{\partial S_A}{\partial U_A} \right)_{V_A, N_A} - \left(\frac{\partial S_B}{\partial U_B} \right)_{V_B, N_B} = 0, \quad (2.81)$$

and likewise for the V and N derivatives.

Since the particle numbers also change, the form of the first law is

$$dU = T dS - p dV + \mu dN. \quad (2.82)$$

Looking at processes with $dN = 0$ and $dV = 0$ we have

$$dU = T dS \Rightarrow \frac{dS}{dU} = \left(\frac{\partial S}{\partial U} \right)_{V, N} = \frac{1}{T}. \quad (2.83)$$

Equation (2.81) is therefore a statement about temperatures on both sides of the separator:

$$\frac{1}{T_A} = \frac{1}{T_B} \Rightarrow T_A = T_B. \quad (2.84)$$

Likewise the conditions coming from δV and δN give

$$\frac{p_A}{T_A} = \frac{p_B}{T_B} \Rightarrow p_A = p_B, \quad (2.85)$$

$$\frac{\mu_A}{T_A} = \frac{\mu_B}{T_B} \Rightarrow \mu_A = \mu_B. \quad (2.86)$$

It is a necessary condition for the equilibrium that the temperatures, pressures and chemical potentials are the same on both sides of the system. This is intuitively as it should be. For instance, if there were a difference in temperature between the sides, the system would develop a heat current from side to side, and clearly would not be in equilibrium.

There are interesting twists to the story. Suppose, for instance that a separator, a diaphragm, really existed and would let heat and particles through, but were effectively immovable. Then the condition coming from the change-of-volume part δV , (2.85) would not apply, and the pressures on the sides of the separate could be different. This is the origin of a phenomenon called *osmotic pressure*. Likewise, if the separator does not allow particles to pass, the condition for equal chemical potentials does not apply.

So far we have studied first-order variations of the total entropy, and found necessary conditions for the entropy to be a maximum. However, while the conditions already guarantee that the entropy is stationary in the purported equilibrium, the stationary point could still be a minimum or a saddle point. To ensure that the stationary point is an extremum one needs to study the variation of entropy to second order in the variations δU , etc. . The calculation is tedious and not so interesting, but the result is. For the extremum of the entropy to be a minimum, we find the local *stability conditions* that heat capacities, compressibilities and derivatives of the form $\frac{\partial \mu}{\partial N}$ must be positive; let us not pay too much attention to what needs to be held constant in each case.

Suppose that we had a piece of matter with a negative compressibility amidst normal matter, and a small fluctuation happened that reduces the volume. For our unusual substance that would mean that the pressure would decrease. So, the outside world compresses the object in, reducing the volume further, which reduces the pressure, which reduces the volume, A runaway instability that decreases the volume (hypothetically) to zero would ensue. Analogously, if a fluctuation increased the volume, a runaway instability would again occur. Clearly, the conditions for local stability make a lot of sense.

2.7 Exercises

2.1. Find (and memorize) the natural variables, expressions of certain thermodynamic parameters as derivatives of the potentials, and the Maxwell relations, for enthalpy $H = U + pV$ and Gibbs free energy $G = U - TS + pV$.

2.2. As customary, let's seriously cut corners and write the first law for a magnetic system as $dU = T dS + B dM$, where, in fact, M is the total magnetic

moment and B means H . Show that for a magnetic substance obeying Curie's law $M = CB/T$, with C being a constant, internal energy only depends on temperature.

2.3. Let us study a material with the equation of state $p = \frac{\alpha T}{V^2}$, where α is a constant. The heat capacity of this material at constant volume is linear in temperature, $C_V = A(V)T$.

- (a) Construct from the known facts an expression for $\left(\frac{\partial S}{\partial V}\right)_T$.
- (b) In fact, the coefficient $A(V)$ must not depend on the volume V . Prove this.
- (b) Find $S(V, T)$ assuming the value $S(T_0, V_0) = S_0$.
- (c) Find the heat capacity at constant pressure $C_p = T \left(\frac{\partial S}{\partial T}\right)_p$.

2.4. Consider a sample of gas in a cylinder that fixes the total volume and the temperature. The cylinder is divided into two parts by a moving piston. By exerting a force somebody acting outside of the system may move the piston, but, let's assume, so slowly that the gases on both sides are in equilibrium and at all times satisfy the relevant equations of state. Show that the state that is sustained without an external force, the obvious equilibrium in the absence of the external agent, is an extremum (actually, minimum) of one of the thermodynamic potentials.

2.5. How hot does air get in the pump when you try to put twice the atmospheric pressure in a bicycle tube at the ambient temperature of 20°C? Hints for beginners: Assume that the compression is reversible and takes place without exchange of heat with the environment. Air is predominantly diatomic molecules.

2.6.

(a) Electromagnetic radiation in a cavity, in equilibrium with the walls at the temperature T , is called black body radiation. It is known that (i) energy density (energy per unit volume) is some function $e(T)$ of temperature alone, and that (ii) the pressure that the radiation exerts on the walls is $p = \frac{1}{3} e(T)$. Find the temperature dependence of U , p and S .

(b) Either as part of the solution to part (a) or otherwise, show that $\mu = 0$ holds true.

2.7. Show that $C_p/C_V = \kappa_T/\kappa_S$.

2.8. Find the change in temperature during the adiabatic compression starting at 150°C of a solid from 10^5 Pa to 10^8 Pa (1 atm to 1000 atm). Assume that the specific heat $2 \text{ J cm}^{-3} \text{ K}^{-1}$ and the coefficient of thermal expansion $8 \times 10^{-5} \text{ K}^{-1}$ remain constants over the relevant parameter range. Use your head, watch out for units!

2.9. Two bodies with fixed volumes are otherwise thermally, chemically and mechanically isolated except that there is a channel for conduction of heat between them. The bodies start out at slightly different temperatures. Because of the heat current, the temperatures equilibrate. Find how much entropy is generated if everything but the heat transfer is reversible.

2.10. Consider a thin wire whose length as a function of the pulling force F and temperature T is $l = l_0[1 + \alpha(T - T_0) + YF]$. Here l_0 is the length of the wire at the reference temperature T_0 , and α and Y are constants. Moreover, let us assume that the heat capacity of the wire at constant length, C_l , may be regarded as a constant. How much does the temperature of the wire change when it is stretched a small fraction β of its original length? The first law in this system reads, of course, $dU = T dS + F dl$.

2.11. Consider a cycle (same initial and final state) consisting of four parts in which an arbitrary thermodynamic system (i) absorbs heat from a reservoir at the temperature T_2 [with, possibly, some incidental work done]; (ii) does work on the environment [or environment does work on system] without any exchange of heat; (iii) delivers some amount of heat to a reservoir at the temperature $T_1 < T_2$ [incidental work possible...]; and (iv) does work on the environment [\pm sign...] without heat exchange. If all of this happens reversibly, the cycle is called a *Carnot engine*. “Reservoir” means that your source/sink of heat is regarded as large enough so that the temperature does not change with addition or removal of heat.

(a) Show that the total work done on the environment and the heat taken up at temperature T_2 satisfy

$$\eta = \frac{\Delta W}{Q_2} = 1 - \frac{T_1}{T_2}. \quad (2.87)$$

From a thermodynamical viewpoint reversibility of a Carnot engine means that it can be, ahem, reversed; run in the opposite direction so that it absorbs work from the environment to transfer heat from a lower temperature T_1 to a higher temperature T_2 . This is basically how a refrigerator works.

(b) Now suppose that there exists an engine that operates between two temperatures T_2 and T_1 with a better efficiency than a Carnot engine. Show that, by plugging the work from this engine to a suitable Carnot engine run in reverse, it is possible to achieve a thermodynamic transformation whose sole effect is to deliver an amount of heat from a lower to a higher temperature.

(c) The impossibility of this feat is exactly the *Clausius* statement of the second law of thermodynamics. But show that it will not go starting from the class version of the second law, $T dS \geq dQ$.

Carnot engine is therefore the most efficient engine. Given that $T_1 > 0$ always holds true, it is impossible to convert an amount of heat entirely to work. This is the essence of the *Kelvin* statement of the second law of thermodynamics.

2.12.

(a) Electromagnetic radiation in a cavity, in equilibrium with the walls at the temperature T , is called black body radiation. It is known that (i) energy density (energy per unit volume) is some function $e(T)$ of temperature alone, and that (ii) the pressure that the radiation exerts on the walls is $p = \frac{1}{3} e(T)$. Find the temperature dependence of U , p and S .

(b) Either as part of the solution to part (a) or otherwise, show that $\mu = 0$ holds true.

2.13. Consider an adsorption site on a surface in contact with an assumedly ideal monatomic gas. An atom may be trapped in such a site of energy $-D$; i.e., when the atom is trapped, the net energy D is released to the rest of the ideal gas. Find the probability that the adsorption site is occupied. In this case, unfortunately, the constants that come from quantum statistical mechanics do matter. For the time being let us take it as given that the chemical potential of an ideal gas is of the form $\mu = kT \ln[n/n_Q(T)]$, where n is density and $n_Q(T)$ is a certain quantum unit of density for the given temperature T . In the current BEC parlance the ratio $\zeta = n/n_Q(T)$ is called “phase space density.”

2.14.

(a) Suppose that in a gas-liquid phase transition the latent heat is a once-and-for all constant, that the volume of the liquid is negligible compared to the volume of the gas, and that the gas may be assumed ideal. Find the liquid-gas coexistence curve.

(b) Find vapor pressure data for water as a function of temperature from, say, CRC Handbook of Physics and Chemistry. How good is the functional form found in part (a)?

2.15. Let us write the first law for a magnetic material as $dU = T dS + H dM$, where H is the applied magnetic field and M is the total magnetic moment. In practice a normal metal does not magnetize, while a type I superconductor expels the magnetic field completely, $B = \mu_0 H + M/V = 0$, until the metal turns normal at the critical field $H_c(T)$. Calculate the difference of Gibbs free energy between the superconductor and the normal metal as a function of H and T .

2.16. *Local-density approximation.* When a gas is subject to an external potential energy (per particle) $\Phi(\mathbf{r})$, it may happen that the potential varies little over the length scale in which the gas can find an equilibrium, and so the gas may be taken to behave locally as if there were no confining potential at all. The gas simply has some equilibrium properties that vary as a function of position.

(a) Sketch a local equilibrium argument showing that at fixed temperature the equilibrium condition is $\mu(\mathbf{r}) + \Phi(\mathbf{r}) = \text{constant}$. HINT: The external potential creates forces on the molecules, so this really is a question of mechanical equilibrium between the forces from pressure and from the potential.

- (b) Find the chemical potential for an ideal gas. In a form of what is known as Gibbs paradox, you may meet with thermodynamic quantities that are neither extensive nor intensive. As discussed later, the cure is to subtract a term $k N \ln N$ from entropy; do it, but do not get too hung up on the reason now.
- (c) Suppose the atmosphere is at constant temperature T , and is made of molecules of mass m . How does the pressure vary as a function of height?

Chapter 3

Statistical mechanics: preliminaries

3.1 Quantum versus classical

Thermodynamics produces relations between the properties of macroscopic bodies, but will not predict the value of any quantity from the microscopic Hamiltonian of the many-particle system. The latter task is the purview of statistical mechanics, especially in the case of thermal equilibrium.

In the old days the curriculum went from thermodynamics to classical statistical mechanics to quantum statistical mechanics, but this sequence is contorted to the extreme. The reason why statistical mechanics works is that the states of a quantum system are discrete and can be counted, a property that does not hold in classical mechanics. Besides state counting or the absence thereof, there is a difference in the structure of classical and quantum theory of identical particles: classically exchanging two identical particles gives rise to a new distinct state, not so quantum mechanically. This leads to another major difference in the counting of the states, and under certain conditions, to drastically different physics results from classical and quantum mechanics. This is not a minor point. For instance, naive classical statistical mechanics presents the same Gibbs paradox for the entropy of an ideal gas that we already pointed out in thermodynamics, but the Gibbs paradox goes away when one property takes into account the indistinguishability of identical particles in quantum mechanics.

Here the strategy is to go to quantum statistical mechanics first, classical statistical mechanics comes second. There are rigorous methods to take the limit from quantum theory to classical physics, but we will proceed heuristically. However, while in our development classical statistical mechanics is sort of an afterthought, its practical importance should not be underestimated. Many useful calculations are possible classically using simple analytical methods, whereas the corresponding quantum analysis would be *much* more difficult.

3.2 Demonstrating state counting

We demonstrate state counting with a system of N spins in a lattice. Our spins are classical, so that they only have the possible values \uparrow and \downarrow ; no superpositions are allowed. Furthermore, the spins are distinguishable; a spin at each site is regarded as different and distinct from the spins at other sites. This scheme is very general. Another guise is *lattice gas*, in which space is divided into cells and a cell is either occupied by a molecule or empty. By the same token you can go to a parking lot and do a statistical-mechanics analysis of the cars in the slots.

Since each of the distinguishable spins can be in either of the two states \uparrow and \downarrow , the total number of possible configurations, *microstates*, of the lattice is 2^N . However, if this were a microscopic description of, say, a magnet, ordinarily one could not resolve the microstates, but is instead interested in macroscopic properties such as energy, magnetization or total magnetic moment, *macrostates*, which are characterized by the numbers N_\uparrow and N_\downarrow of spins up and down.

For simplicity, let us take the total number of lattice sites N to be even, and define an integer-valued quantity s , call it spin excess, as follows:

$$\begin{cases} N_\uparrow + N_\downarrow = N \\ N_\uparrow - N_\downarrow \equiv 2s \end{cases} \Leftrightarrow \begin{cases} N_\uparrow = \frac{1}{2}N + s \\ N_\downarrow = \frac{1}{2}N - s \end{cases} . \quad (3.1)$$

If the magnetic moments for the \uparrow and \downarrow spins are $+\mu$ and $-\mu$, the total magnetic moment of the lattice thereby equals $M = \mu(N_\uparrow - N_\downarrow) = \frac{1}{2}\mu s$. The possible values of, say, N_\downarrow run from 0 to N , and equivalently, s runs from $-N/2$ to $N/2$ in steps of one. The number of possible macrostates is therefore $N + 1$, for a large number of sites N much smaller than the number of microstates 2^N .

Let us now recall the result from elementary combinatorics that one can select k out of n objects in

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (3.2)$$

different ways. To get a macrostate characterized by N_\uparrow spins up, one has to select N_\uparrow sites out of N and put the spin up in the selected sites, down in the rest. This can be done in

$$g(N, s) = \frac{N!}{N_\uparrow! N_\downarrow!} = \frac{N!}{(\frac{1}{2}N + s)!(\frac{1}{2}N - s)!} \quad (3.3)$$

different ways, which specifies the *multiplicity* of the state with spin excess s . The sum of the multiplicities should equal the total number of microstates, which in fact is the case:

$$\sum_{s=-\frac{N}{2}}^{\frac{N}{2}} g(N, s) = \sum_{N_\uparrow=0}^N \binom{N}{N_\uparrow} = \sum_{N_\uparrow=0}^N \binom{N}{N_\uparrow} 1^{N_\uparrow} 1^{N-N_\uparrow} = (1+1)^N = 2^N . \quad (3.4)$$

Another frequently used trick of the trade is the *Stirling approximation*, according to which in the limit of large n the expression

$$\ln n! = n \ln n - n + \mathcal{O}(n^0) \quad (3.5)$$

holds true. The integers one deals with in statistical mechanics are indeed large, the Avogadro number $\sim 10^{23}$ is conventionally regarded as representative. The constant beyond $\ln n! \simeq n \ln n - n$ is occasionally relevant, but if needed it can usually be deduced from other considerations.

Consider now the multiplicity $g(N, s)$ as a function of the spin excess s . By symmetry it is a maximum for $s = 0$, and $N \gg 1$ the Stirling approximation is appropriate for $|s| \ll N$:

$$\begin{aligned} f(s) &\equiv \ln g(N, s) = \ln(N!) - \ln\left[\left(\frac{1}{2}N + s\right)!\right] - \ln\left[\left(\frac{1}{2}N - s\right)!\right] \\ &\simeq N \ln N - \left(\frac{1}{2}N + s\right) \ln\left(\frac{1}{2}N + s\right) - \left(\frac{1}{2}N - s\right) \ln\left(\frac{1}{2}N - s\right), \end{aligned} \quad (3.6)$$

the next term of a series in N being a constant. We have

$$f'(s) \simeq \ln\left(\frac{1}{2}N - s\right) - \ln\left(\frac{1}{2}N + s\right), \quad f'(0) = 0; \quad (3.7)$$

$$f''(s) \simeq -\frac{1}{\frac{1}{2}N + s} - \frac{1}{\frac{1}{2}N - s}, \quad f''(0) = -\frac{4}{N}, \quad (3.8)$$

and so

$$f(s) \simeq f(0) + sf'(0) + \frac{1}{2}s^2 f''(0) \simeq K - \frac{2}{N}s^2; \quad (3.9)$$

$$g(N, s) = e^{f(s)} \simeq C e^{-\frac{2}{N}s^2}. \quad (3.10)$$

Here K and C are constants that we do not know up front as we terminated the Stirling approximation at the order we did. As a mathematical fine detail it may be noted that the scale of variation of $f(s)$ with s , and therefore the range over which the expansion (3.9) is useful, scales with N as $\Delta s \sim N$, while the multiplicity $g(N, s)$ effectively becomes zero over the scale $\Delta s \sim \sqrt{N}$. The expression (3.10) therefore becomes arbitrarily accurate as $N \rightarrow \infty$.

Although the value is not important for our argument, to illustrate a couple of common mathematical techniques let us also figure out the coefficient C in the multiplicity. We have

$$2^N = \sum_s g(N, s) \simeq \int_{-\frac{1}{2}N}^{\frac{1}{2}N} g(N, s) ds \simeq C \int_{-\infty}^{\infty} ds e^{-\frac{2}{N}s^2} = \sqrt{\frac{\pi N}{2}} C. \quad (3.11)$$

The first approximate inequality is because for large N $g(N, s)$ varies slowly as a function of s and the sum may be replaced by an integral. Even if the integration limits technically are $\pm \frac{1}{2}N$, at the limits the integrand equals $e^{-N/2}$ and is exponentially small for large N . As in the second approximate equality, we may just as well pull the integration limits to $\pm\infty$. Overall, we have an

equation for the coefficient C , and the final approximate expression for the multiplicity,

$$g(N, s) \simeq 2^N \sqrt{\frac{2}{\pi N}} e^{-\frac{2s^2}{N}}. \quad (3.12)$$

This is a Gaussian function of s centered at $s = 0$, and has the room-mean square width $\Delta s = \sqrt{N}/2$.

3.3 Simple statistical mechanics for spins

The purpose of the present section is to illustrate a few basic ideas of the physics of macroscopic system in thermodynamic equilibrium. Our statements and the demonstrations thereof are rather far apart. The notions we are discussing will be met over and over again in the rest of this book; at this point the reader is advised not to engage in attempts at rigor, or in philosophical ponderations.

While any macroscopic system under study naturally has a microstate, it is not known and in practice usually not even knowable. Instead, the microstate is regarded as random. The probabilities for the microstates are then given by the fundamental postulate of statistical mechanics: every microstate compatible with whatever given specifications of the macroscopic state is equally likely.

Suppose we a priori specify no macroscopic properties for our example spin system (other than the number of spins). This means that every microscopic state is compatible with the specifications, and is equally likely. One might think that, as a result, all possible values of macroscopic properties are equally likely as well. The maybe surprising lesson from our state counting argument is that it is not so. Since every microscopic state is equally likely, the value of spin excess s will occur at a probability proportional to the number of microscopic states with the given value of s , i.e., at a probability proportional to the multiplicity $g(N, s)$. Accordingly, the spin excess comes in a band with the width $\Delta s \sim \sqrt{N}$ around $s = 0$. This may not seem small in the limit $N \rightarrow \infty$, but the possible range of s is from $-\frac{1}{2}N$ to $\frac{1}{2}N$, and so only the fraction $\sim \Delta s/N \sim 1/\sqrt{N}$ of the possible values of s is in practice found in this statistical mechanics system. This fraction tends to zero in the limit $N \rightarrow \infty$.

The mere fact that the number of microscopic states is large tends to establish well-defined values for macroscopic variables. Depending on the microstate, and hence, from experiment to experiment, the value of the spin excess is different. But, the larger is the number of spins, the smaller are the fluctuations, i.e., the range in relative terms of the values that in practice occur in the spin system. This is the origin of thermodynamic equilibrium, and of the fact that in the simplest form of thermodynamics the fluctuations of macroscopic variables may be ignored altogether.

Before we may turn the example about the spins into a full-blown exercise in statistical mechanics, a important element is still needed. In thermodynamics entropy plays a crucial role, but as of yet we do not know what is the microscopic meaning of entropy. Let us therefore state another postulate: Thermodynamical

entropy is given by

$$S = k \ln W ,$$

where W is the number of microstates compatible with any given specifications of the macroscopic state and k is the usual Boltzmann constant.

Armed with these tools, we return to the spin system. Suppose the spin states have different energies, $\epsilon_{\uparrow} = \frac{1}{2}\epsilon$ and $\epsilon_{\downarrow} = -\frac{1}{2}\epsilon$ for the states \uparrow and \downarrow . If the spins do not interact with each other, the total energy is

$$E = \frac{\epsilon}{2}(N_{\uparrow} - N_{\downarrow}) = s\epsilon . \quad (3.13)$$

Let us now plausibly equate the total energy E with the thermodynamical internal energy U . Therefore we may express the spin excess as a function of internal energy simply as $s = E/\epsilon = U/\epsilon$. The number of microstates compatible with the specification of the system that the internal energy has a given value U equals the multiplicity for the spin excess $s = U/\epsilon$, so that the entropy corresponding to the internal energy U is written with the aid of the Stirling approximation as

$$\begin{aligned} S(U) &= k \ln W = k \ln g(N, U/\epsilon) \\ &\simeq k \left[N \ln N - \left(\frac{N}{2} + \frac{U}{\epsilon} \right) \ln \left(\frac{N}{2} + \frac{U}{\epsilon} \right) - \left(\frac{N}{2} - \frac{U}{\epsilon} \right) \ln \left(\frac{N}{2} - \frac{U}{\epsilon} \right) \right] . \end{aligned} \quad (3.14)$$

But for the spin system with no coupling to any external provider of work, the first law says

$$dU = T dS; \quad \frac{1}{T} = \frac{dS}{dU} . \quad (3.15)$$

Taking the derivative of (3.14) we have

$$\frac{1}{T} = -\frac{k}{\epsilon} \ln \frac{\frac{N}{2} + \frac{U}{\epsilon}}{\frac{N}{2} - \frac{U}{\epsilon}}, \quad \frac{\frac{N}{2} + \frac{U}{\epsilon}}{\frac{N}{2} - \frac{U}{\epsilon}} = e^{-\frac{\epsilon}{kT}} = e^{-\beta\epsilon} . \quad (3.16)$$

Here $\beta = \frac{1}{kT}$ is a notation so standard in statistical mechanics that you very seldom see it explained. There are numerous ways of putting the result, among them the following:

$$\begin{aligned} U &= -\frac{N\epsilon}{2} \frac{1 - e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} = -\frac{N\epsilon}{2} \frac{e^{\beta\epsilon/2} - e^{-\beta\epsilon/2}}{e^{\beta\epsilon/2} + e^{-\beta\epsilon/2}} = -\frac{N\epsilon}{2} \tanh \frac{\beta\epsilon}{2} , \\ N_{\uparrow} &= \frac{N}{2} + s = \frac{N}{2} + \frac{U}{\epsilon} = N \frac{e^{-\beta\epsilon/2}}{e^{\beta\epsilon/2} + e^{-\beta\epsilon/2}} , \\ N_{\downarrow} &= N \frac{e^{\beta\epsilon/2}}{e^{\beta\epsilon/2} + e^{-\beta\epsilon/2}} , \\ \frac{N_{\uparrow}}{N_{\downarrow}} &= e^{-\beta\epsilon} . \end{aligned} \quad (3.17)$$

The meat of the argument is in the last form (3.17). The ratio of the numbers of spin up and spin down is, of course, equal to the ratio of the probabilities that any given spin is up or down, and the ratio equals the usual *Boltzmann factor* $e^{-\beta\epsilon_{\uparrow}}/e^{-\beta\epsilon_{\downarrow}} = e^{-\beta\epsilon}$ as in undergraduate thermal physics.

The questions of why it pays to make the fundamental postulate and why entropy should be defined microscopically as we have defined it may be approached from a multitude of mathematical and philosophical angles, but we again warn against attempts at being “fundamental.” The ultimate arbiter here is empirical evidence that has accumulated over a hundred and thirty years. As the connection to the well-known Boltzmann factor we have developed above demonstrates, it all hangs together nicely.

3.4 Exercises

3.1. *A variation of the infamous lattice gas.* Suppose we have a big volume V divided into cells of volume v . Each cell can accommodate at most one diatomic molecule. There are N ($N \gg 1$, $N \ll V/v$) molecules, and by assumption there are no other degrees of freedom for a molecule except that it is in one of the cells. However, if supplied the dissociation energy D , a molecule will break up into two atoms, each of which can similarly occupy one of the cells. There are no constraints for an atom and a molecule occupying the same cell.

- (a) Find the free energy (can’t say if it is Helmholtz or Gibbs, as there are no volume/pressure type thermodynamic parameters in this problem) at temperature T , assuming that the gas is either all atoms, or all molecules.
- (b) Take it as given that there is a phase transition in which the molecules dissociate. What is the transition temperature?

Chapter 4

Quantum mechanics primer

All over the world elementary quantum mechanics courses are still formulated in terms of state vectors. However, there is a more general concept of a state of a system in quantum mechanics than state vector, called density operator, and quantum statistical mechanics is naturally formulated in terms of density operators. The same applies to fields such as quantum information, quantum measurement theory, and theory of open systems that interact with an environment. We start our foray into quantum statistical mechanics with an extensive broadening of the standard quantum mechanics wisdom; really, more extensive than would be needed for the purposes of quantum statistical mechanics alone.

We will assume that the student is familiar with the standard “Quantum I” material, in other words understands what the following formulation of the axioms of quantum mechanics means:

1. A Hilbert space represent a physical system, and a vector in the Hilbert space $|\psi\rangle$ represent the state of the system.
2. Hermitian operators A represent observables.
3. Suppose $\{|n\rangle\}$ is an orthonormal basis of the eigenstates of an observable A , and a_n are the corresponding eigenvalues.
 - (a) A measurement of the observable A gives as a result one of the eigenvalues a_n .
 - (b) The probability for the measurement value a_n is $p_n = |\langle n|\psi\rangle|^2$.
 - (c) Immediately after a measurement that gave the observed value a_n , the state of of the system is $|\psi'\rangle = |n\rangle$.
4. Time evolution of the state is generated by a Hermitian Hamiltonian H according to

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (4.1)$$

The elementary form of the measurement postulate 3 contains an ambiguity if it so happens that two or more of the eigenvalues a_n are the same; for the time being assume that the eigenvalues are nondegenerate. The postulate 3.c spells out the infamous “collapse (reduction) of the wave packet.” It is there to ensure the validity of the obviously reasonable condition that if a measurement is immediately repeated, the result will also repeat. The fourth postulates states the time dependent Schrödinger equation. It is valid in a closed system that does not interact with an environment.

The fact that the postulates specify two different types of evolution, Schrödinger equation and the reduction of the wave packet upon measurement, has been the source of much consternation since the first days of quantum mechanics. After all, if one enlarges the system to include the measurement apparatus, the result is seemingly a closed system and should be described by the Schrödinger equation, so that there seems to be no need for an extra postulate. Unfortunately, it turns out that simply adding the measurement apparatus to the quantum mechanical description does not remove the problem. This is one manifestation of the issue of quantum-classical interface, which remains just as unsolved as it was eighty years ago.

According to the measurement postulates, the expectation value of the observable A in the state $|\psi\rangle$ is

$$\begin{aligned}\langle A \rangle &= \sum_n p_n a_n = \sum_n a_n |\langle n|\psi\rangle|^2 = \sum_n a_n \langle\psi|n\rangle\langle n|\psi\rangle \\ &= \langle\psi| \left(\sum_n a_n |n\rangle\langle n| \right) |\psi\rangle = \langle\psi|A|\psi\rangle.\end{aligned}\tag{4.2}$$

This calculation involves the spectral representation of the observable A that will be explained below, plus some trickery with the Dirac notation, so that a slight unfamiliarity with the technical details is not cause for alarm at this point. There are two points we want to make here, though. First, Dirac notation is a bag of tricks that hides a lot of subtle mathematics. Learn to use it fluently, but, unless you are extremely mathematically inclined, do *not* worry about exactly what mathematics. Second, the final result is what is called quantum mechanical expectation value of the operator A , but the chain of equalities starts with the classical average of the numbers a_n that occur with the probabilities p_n . Now, according to the measurement postulates, the values a_n occur with the probabilities p_n ! So, the expectation value is not just some number that happens to be called expectation value: According to the usual interpretation of classical probabilities and by virtue of the measurement postulates, the average value obtained in a large number of measurements on (copies of) the system in the state $|\psi\rangle$ equals the quantum mechanical expectation value. The same applies to the standard deviation ΔA defined through

$$(\Delta A)^2 = \langle A^2 \rangle - \langle A \rangle^2 = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2.\tag{4.3}$$

4.1 Topical review

We now review, and elaborate on, some basic quantum mechanics. In quantum mechanics you meet with both continuum states and state spaces with both finite and numerably infinite dimensions. We will say nothing about continuum states. At any rate, continuum states turn discrete if the quantum system is confined to a finite volume, although the limit of a large confining volume can be tricky to analyze. In our discussions we state all the results as if the state space of the quantum system were a finite-dimensional Hilbert space \mathcal{H} . There are mathematical pathological cases in which certain results are different in infinite-dimensional Hilbert spaces. However, such pathologies come up rarely in practical quantum mechanics, and are henceforth mostly ignored. If it makes sense and if we specifically do not say differently, all states we will be dealing with are implicitly normalized to unity, so that $(\psi, \psi) = \langle \psi | \psi \rangle = 1$. Finally, we make extensive use of the magic of the Dirac notation. We will not attempt to state any full set of rules for the Dirac notation, but present many more explicit calculations than we really have to in order to provide examples to emulate.

We begin with a handy auxiliary concept of *dyadic product* $|\phi_1\rangle\langle\phi_2|$ of two states. It is a linear operator defined so that

$$(|\phi_1\rangle\langle\phi_2|)|\psi\rangle = |\phi_1\rangle\langle\phi_2|\psi\rangle \quad (4.4)$$

holds for an arbitrary vector $|\psi\rangle$. If the (normalized) states in the dyadic product are the same, we have a *projector* or *projection operator* onto the state: $P_\psi = |\psi\rangle\langle\psi|$. As any projection operator, this one too is its own square:

$$P_\psi P_\psi = |\psi\rangle\langle\psi||\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = P_\psi. \quad (4.5)$$

A subspace of a Hilbert space \mathcal{S} is a set of vectors that is closed with respect to addition of vectors and multiplication with scalars. It is a Hilbert space in its own right and may, in fact, coincide with the Hilbert space \mathcal{H} itself. A frequent source of subspaces in quantum mechanics is the eigenvalue problem: Given any linear operator, its eigenvectors belonging to the same eigenvalue amended with the zero vector (which by definition is not an eigenvector) is a subspace.

A set of vectors $\{|n\rangle\}$ is orthonormal if the vectors are orthogonal and normalized to unity, so that $\langle m|m\rangle = \delta_{nm}$ where δ_{nm} is the Kronecker delta symbol. Given any set of vectors, with the aid of the well-known Gram-Schmidt construction it is possible to construct a set of orthonormal vectors that spans the same subspace as the initial set of vectors.

A set of vector $\{|\psi_n\rangle\}$ is a basis if every vector in the linear space can be represented as a linear combination of the vectors $|\psi_n\rangle$. A Hilbert space admits an orthonormal basis, a set of vectors such that

$$\langle n|m\rangle = \delta_{nm}, \quad \sum_n |n\rangle\langle n| = I, \quad (4.6)$$

where the latter extremely useful expression says that the sum of the projectors onto the members of an orthonormal basis is the unit operator of the Hilbert

space; the operator that maps every vector onto itself. By the same token, it is possible to construct an orthonormal basis for every subspace \mathcal{S} of the Hilbert space. The ensuing unit operator in the subspace \mathcal{S} is called (orthogonal) *projector* to the subspace \mathcal{S} when viewed as an operator in the underlying Hilbert space:

$$P_{\mathcal{S}} = \sum_{n \in \mathcal{S}} |n\rangle\langle n|. \quad (4.7)$$

This is a proper projection operator, with $P_{\mathcal{S}}^2 = P_{\mathcal{S}}$. The action of the projector on any Hilbert space vector is to pick out the unique component of the vector belonging to the subspace, as if to pick, say, the x -component of a three-dimensional vector. The projectors onto two subspaces \mathcal{R} and \mathcal{S} that have only the zero vector in common satisfy $P_{\mathcal{S}}P_{\mathcal{R}} = 0$. In particular, given a projector P onto some subspace \mathcal{S} , $Q = I - P$ is a projector onto the complement \mathcal{S}^{\perp} of \mathcal{S} in the Hilbert space; the set of vectors that do not belong to \mathcal{S} , plus the zero vector. The spaces \mathcal{S} and \mathcal{S}^{\perp} are orthogonal, i.e., for any $|\psi\rangle \in \mathcal{S}$ and $|\phi\rangle \in \mathcal{S}^{\perp}$ we have $\langle\psi|\phi\rangle = 0$.

Define the Hermitian conjugate A^{\dagger} of an operator A in such a way that $(\psi, A\phi) = (A^{\dagger}\psi, \phi)$ for any two vectors ψ and ϕ . An operator A that is its own Hermitian conjugate, $A = A^{\dagger}$, is *Hermitian* or *self-adjoint*. As is well known, the great virtues of a Hermitian operator in quantum mechanics are that its eigenvalues are real, and that the eigenvectors belonging to two different eigenvalues are orthogonal. Ultimately, that is why Hermitian operators are used to represent observables in elementary quantum mechanics.

If an operator U satisfies¹

$$UU^{\dagger} = U^{\dagger}U = I, \quad (4.8)$$

the operator U is *unitary*. Given any two vectors ψ and ϕ , we have $(U\psi, U\phi) = (U^{\dagger}U\psi, \phi) = (\psi, \phi)$, or

$$(U\psi, U\phi) = (\psi, \phi). \quad (4.9)$$

In this sense, unitary operators preserve the values of inner products. Given an arbitrary orthonormal basis $\{|n\rangle\}$, the set of vectors $\{|u_n\rangle\}$ with $|u_n\rangle = U|n\rangle$ is therefore also an orthonormal basis. Moreover, we have for an arbitrary $|\psi\rangle$

$$U|\psi\rangle = \sum_n U(|n\rangle\langle n|)|\psi\rangle = \left(\sum_n |u_n\rangle\langle n|\right)|\psi\rangle, \quad (4.10)$$

so that

$$U = \sum_n |u_n\rangle\langle n| \quad (4.11)$$

holds true as an operator equation. Finally, any operator of this form where $\{|n\rangle\}$ and $\{|u_n\rangle\}$ are arbitrary orthonormal bases satisfy

$$UU^{\dagger} = \sum_{n,m} |u_n\rangle\langle n|m\rangle\langle u_m| = \sum_n |u_n\rangle\langle u_n| = I, \quad (4.12)$$

¹This is one of the rare instances when finite- and infinite-dimensional spaces differ even in practical terms. In a finite-dimensional space it is sufficient to require, say, that $UU^{\dagger} = I$, but in general omitting the other equality may get you in trouble.

and likewise $U^\dagger U = I$. All three properties we have listed in (4.8), (4.9) and (4.11) can equivalently serve as the definition of a unitary operator.

Unitary operators have multiple uses in quantum mechanics. The property (4.8) is the basis of *unitary transformations*. For instance, if for some fixed unitary U one defines a transformation of an arbitrary state and an operator as

$$|\tilde{\psi}\rangle = U|\psi\rangle, \quad \tilde{A} = UAU^\dagger, \quad (4.13)$$

then by (4.8)

$$\langle\tilde{\psi}|\tilde{A}|\tilde{\phi}\rangle = \langle U\psi, UAU^\dagger U\phi\rangle = \langle U^\dagger U\psi, AU^\dagger U\phi\rangle = \langle\psi|A|\phi\rangle. \quad (4.14)$$

Moreover

$$\begin{aligned} [\tilde{A}, \tilde{B}] &= \tilde{A}\tilde{B} - \tilde{B}\tilde{A} = UAU^\dagger UBU^\dagger - UBU^\dagger UAU^\dagger = U[A, B]U^\dagger \\ &= \widetilde{[A, B]}. \end{aligned} \quad (4.15)$$

A unitary transformation preserve the structure of quantum mechanics. There is a subtlety with the Hamiltonian that occur if the unitary transformation happens to depend explicitly on time, but otherwise the unitarily transformed version of quantum mechanics is isomorphic to the original one.

Another noteworthy development follows from (4.9). Namely, take two vectors evolving according to the time dependent Schrödinger equation and its Hermitian conjugate, according to

$$\frac{\partial}{\partial t}|\phi(t)\rangle = -\frac{i}{\hbar}H|\phi(t)\rangle, \quad \frac{\partial}{\partial t}\langle\psi(t)| = \frac{i}{\hbar}\langle\psi(t)|H. \quad (4.16)$$

Calculus applies to vectors and operators just like it does to real and complex numbers as long as you do not improperly change the order of the objects in the expressions, so that we have from the above

$$\frac{\partial}{\partial t}\langle\psi(t)|\phi(t)\rangle = \left(\frac{\partial}{\partial t}\langle\psi(t)|\right)|\phi(t)\rangle + \langle\psi(t)|\left(\frac{\partial}{\partial t}|\phi(t)\rangle\right) = 0. \quad (4.17)$$

Schrödinger equation time evolution preserves inner products between vectors. Now, given the state $|\psi(t_0)\rangle$ at t_0 , it uniquely determines the state $|\psi(t)\rangle$ at time t ; and besides, since the Schrödinger equation is linear, so is the mapping $U(t, t_0)$ that tells how $|\psi(t_0)\rangle$ gets transformed into $|\psi(t)\rangle$. The operator $U(t, t_0)$ is therefore unitary by virtue of (4.9). The time evolution as described by the time evolution postulate of quantum mechanics is a unitary transformation of the state vector.

A *normal* operator N satisfies $NN^\dagger = N^\dagger N$. In fact, both Hermitian and unitary operators are normal. The importance of normal operators lies in the following theorem:

Theorem. For any normal operator there exists complex numbers c_n and an orthonormal basis $\{|n\rangle\}$ such that

$$N = \sum_n c_n |n\rangle\langle n|. \quad (4.18)$$

Proof. Let us take for granted that a normal operator has at least one eigenvalue, and for precisely one eigenvalue the corresponding eigenvectors span the whole Hilbert space.² The proof is then by induction with respect to the number of discrete eigenvalues. If there is only one discrete eigenvalue λ , by the assertions we took for granted, (4.18) is valid with $c_n \equiv \lambda$ and for an arbitrary orthonormal basis $\{|n\rangle\}$. Suppose the statement is valid for any normal operator with up to k discrete eigenvalues, and then consider a normal operator N with $k+1$ discrete eigenvalues. Denote the subspace of the $k+1^{\text{th}}$ eigenvalue by \mathcal{S} , the orthogonal projection onto \mathcal{S} by P , the orthogonal complement of \mathcal{S} by \mathcal{S}^\perp , and the projection onto \mathcal{S}^\perp by $Q = I - P$, or $P + Q = I$. We may therefore identically write

$$N = INI = PNP + PNQ + QNP + QNQ. \quad (4.19)$$

Now, for any vector $|\psi\rangle$ in \mathcal{S} , $N|\psi\rangle = \lambda_{k+1}|\psi\rangle$ also is in \mathcal{S} and is orthogonal to any vector in \mathcal{S}^\perp . Hence $QNP = 0$. On the other hand, for any such $|\psi\rangle$ also $NN^\dagger|\psi\rangle = N^\dagger N|\psi\rangle = \lambda_{k+1}N^\dagger|\psi\rangle$ holds true by virtue of normality, so that $N^\dagger|\psi\rangle$ likewise belongs to the subspace \mathcal{S} . Therefore $QN^\dagger P = 0$, and taking the Hermitian conjugate also $PNQ = 0$. Overall, we have $N = PNP + QNQ$. Since $PQ = QP = 0$, we have from the normality of N that $(PN^\dagger P)(PNP) + (QN^\dagger Q)(QNQ) = (PNP)(PN^\dagger P) + (QNQ)(QN^\dagger Q)$, and multiplying this equation from the left by P , $(PN^\dagger P)(PNP) = (PNP)(PN^\dagger P)$. Therefore PNP is a normal operator, and in the same way so is QNQ . By the induction hypothesis we may now write both PNP and QNQ in the form (4.18), and in such a way that the orthonormal vectors in the representations of the operators PNP and QNQ , belonging to the respective subspaces \mathcal{S} and \mathcal{S}^\perp , are orthonormal. Therefore we immediately have the decomposition (4.18) for $N = PNP + QNQ$. \square

The *spectral representation* (4.18) tells us that a normal operator can be diagonalized. In fact, the numbers c_n are the eigenvalues (not all necessarily distinct) and $|n\rangle$ are the corresponding orthonormal eigenstates. The spectral representation is an extremely valuable in the use of quantum mechanics. It therefore bears noting that if an operator N has a spectral representation as in (4.18), then one finds that

$$N^\dagger N = \left(\sum_n c_n^* |n\rangle \langle n| \right) \left(\sum_m c_m |m\rangle \langle m| \right) = \sum_n |c_n|^2 |n\rangle \langle n| = N^\dagger N. \quad (4.20)$$

Every operator with a spectral representation is also normal; do not try to argue with a spectral representation unless your operator is normal.

As already noted, the eigenvalues of a Hermitian operator are real. Therefore the expectation value of a Hermitian operator in an arbitrary state is also real. Namely, let a_n be the eigenvalues and $\{|n\rangle\}$ be the corresponding orthonormal

²In a finite-dimensional linear space both of these statements are true of any linear operator; the case of an infinite-dimensional Hilbert space is one of the subtleties we do not address here.

eigenbasis so that $A|n\rangle = a_n|n\rangle$, then

$$\langle\psi|A|\psi\rangle = \langle\psi|\left(\sum_n a_n|n\rangle\langle n|\right)|\psi\rangle = \sum_n a_n\langle\psi|n\rangle\langle n|\psi\rangle = \sum_n a_n|\langle n|\psi\rangle|^2 \quad (4.21)$$

is clearly real. By picking $|\psi\rangle = |n\rangle$ for each n , we may see that the only Hermitian operator whose expectation value is zero in all states $|\psi\rangle$ has all eigenvalues a_n equal to zero, i.e., by virtue of the spectral representation, such an operator is the zero operator. Let us next write an arbitrary operator A , Hermitian or not, identically as

$$A = \frac{1}{2}(A + A^\dagger) + i\frac{1}{2}[-i(A - A^\dagger)] \equiv \mathcal{A} + i\mathcal{B}, \quad (4.22)$$

where the operators \mathcal{A} and \mathcal{B} with the obvious definitions are both Hermitian. Now, if $\langle\psi|A|\psi\rangle$ is real for all vectors $|\psi\rangle$, then obviously $\langle\psi|\mathcal{B}|\psi\rangle = 0$ for all $|\psi\rangle$, so that $\mathcal{B} = -i[A - A^\dagger] = 0$ and A is in fact Hermitian. Hermitian operators are therefore precisely those operators whose expectation value in every state of the Hilbert space is real.

A *positive* operator A by definition has $\langle\psi|A|\psi\rangle \geq 0$ for all vectors $|\psi\rangle$. A positive operator is obviously Hermitian. The eigenvalues of a positive operator are non-negative. In fact, if there were an eigenvector $|\lambda\rangle$ corresponding to an eigenvalue $\lambda < 0$, then $\langle\lambda|A|\lambda\rangle = (\lambda, A\lambda) = \lambda < 0$, and A were not positive. The orthogonal projectors P as in (4.7) are positive operators. Also, given any linear operator A , $A^\dagger A$ is positive.

As another demonstration of the power of spectral representations, we can take over any function of a complex variable $f(z)$ and make out of it an operator-valued function of a normal operator as follows: Given that (4.18) holds, we define

$$f(N) = \sum_n f(c_n)|n\rangle\langle n|. \quad (4.23)$$

Another way of defining functions of operators, actually for an arbitrary operator, runs as follows. Suppose a Taylor expansion (the same could be done with the Laurent expansion, or, say, with the kind-of Maclaurin expansion around the identity operator) of the function $f(z)$ is

$$f(z) = \sum_k b_k z^k, \quad (4.24)$$

then define

$$f(A) = \sum_k b_k A^k. \quad (4.25)$$

When both expressions (4.23) and (4.25) apply, they agree. This is the case for some operators, such as $f(z) = e^z$. On the other hand, two functions common in formal developments of quantum mechanics, $\ln z$ and \sqrt{z} , do not have a power

series expansion around $z = 0$, in which case (4.23) could be understood.³

The remaining element in our discussion of elementary quantum mechanics is the trace of an operator, a function that maps an operator into a complex number:

$$\text{Tr} A = \sum_n \langle n | A | n \rangle. \quad (4.26)$$

Trace is a linear operator in its argument operators. The sum is over an arbitrary orthonormal basis, and it does not matter which basis. Namely, let $\{|m\rangle\}$ be another orthonormal basis, then

$$\begin{aligned} \text{Tr} A &= \sum_n \langle n | A | n \rangle = \sum_n \langle n | A I | n \rangle = \sum_{m,n} \langle n | A | m \rangle \langle m | n \rangle = \sum_{m,n} \langle m | n \rangle \langle n | A | m \rangle \\ &= \sum_m \langle m | I A | m \rangle = \sum_m \langle m | A | m \rangle. \end{aligned} \quad (4.27)$$

In particular, for a normal operator one may use an orthonormal basis that diagonalizes it, whereupon it is evident that for a normal operator the trace equals the sum of the eigenvalues. For the product of two operators we have

$$\begin{aligned} \text{Tr}(AB) &= \sum_n \langle n | AB | n \rangle = \sum_n \langle n | A I B | n \rangle = \sum_{m,n} \langle n | A | m \rangle \langle m | B | n \rangle \\ &= \sum_{m,n} \langle m | B | n \rangle \langle n | A | m \rangle = \sum_m \langle m | B I A | m \rangle = \sum_m \langle m | B A | m \rangle \\ &= \text{Tr}(BA). \end{aligned} \quad (4.28)$$

Since A and B are arbitrary, it follows that the trace is more generally unchanged when the argument is a product of operators and the operators in the product are rotated cyclically,

$$\text{Tr}(A_1 A_2 \dots A_{p-1} A_p) = \text{Tr}(A_p A_1 A_2 \dots A_{p-1}). \quad (4.29)$$

The final useful property of the trace we will mention here has to do with the variation of operators. Let us vary the operator A by a “small” amount δA , then any operator function $f(A)$ changes as well, by $\delta f(A) = f(A + \delta A) - f(A)$. This change is easiest to analyze from the form (4.25) for operator functions. To see the problem with the operator functions, consider as an example the third-order product

$$(A + \delta A)^3 = AAA + (\delta A)AA + A(\delta A)A + AA(\delta A) + \mathcal{O}[(\delta A)^2]. \quad (4.30)$$

If A and δA commute, we have $(A + \delta A)^3 \simeq A^3 + 3A^2 \delta A$, and similarly in all orders in the representation of the operator function (4.25). In this case *only*,

$$\delta f(A) = \sum_k k b_k A^{k-1} \delta A = f'(A) \delta A, \quad (4.31)$$

³Inner product gives a convenient distance between vectors, $d(\psi, \phi) = \sqrt{(\psi - \phi, \psi - \phi)}$, and thereby easy definitions of limit and convergence. In contrast, what it means for an operator series to converge may be a complicated question. This is another piece of mathematical fine print that we bypass.

where f' is the derivative of the function f . But A and δA do not commute in general. Taking the trace fixes the problem, because by the cyclic invariance $\text{Tr}(A + \delta A)^3 \simeq \text{Tr}A^3 + 3\text{Tr}(A^2 \delta A)$ holds true. This applies to all powers of $A + \delta A$ in the expansion as in (4.25), so that we have the valid rule

$$\text{Tr}[\delta f(A)] = \text{Tr}[f'(A) \delta A]. \quad (4.32)$$

4.2 Combining quantum systems

Every system except (possibly) the entire universe has an environment, also a system in its own right, and every system ultimately interacts with its environment. This brings in the question of combining two quantum systems. The same item comes up also in much smaller settings. Say, we have two spin- $\frac{1}{2}$ particles, how to represent the state? And technically, one may even think every degree of freedom of a particle as an independent quantum system, so that the three-dimensional motion is already a combination of three quantum systems.

Let's consider two system, call them S and E for “system” and “environment.” If the system is in the state $|\psi_S\rangle$ and the environment in the state $|\psi_E\rangle$, one may simply say so and write the total state

$$|\psi\rangle = \{|\psi_S\rangle, |\psi_E\rangle\} = |\psi_S\rangle \otimes |\psi_E\rangle = |\psi_S\rangle |\psi_E\rangle = |\psi_S\rangle_S |\psi_E\rangle_E = |\psi_S \psi_E\rangle. \quad (4.33)$$

We are really talking about a pair of states, but it is frequently termed tensor product and denoted by the symbol \otimes ; other possible notations are indicated. If now we have an operator that acts solely on the system A_S and another one that acts on the environment A_E , the action of the joint operator on the joint state is

$$A_S \otimes A_E |\psi_S\rangle |\psi_E\rangle = A_S |\psi_S\rangle \otimes A_E |\psi_E\rangle. \quad (4.34)$$

If we have an operator acting only on one of the systems, say, the one we call system, it is understood that we have an operator on the joint system $A_S \otimes I_E$, with the identity operator in the degrees of freedom of the environment.

However, this simple scheme is not enough. Even though it is not always recognized as such, quantum mechanics adds another postulate. Namely, take orthonormal bases for the system and environment $\{|n\rangle\}$ and $\{|\alpha\rangle\}$ and the natural inner product

$$(\langle \alpha' | \langle n' |) |n\rangle |\alpha\rangle = \delta_{nn'} \delta_{\alpha\alpha'}, \quad (4.35)$$

then the joint states $\{|n\rangle |\alpha\rangle\}_{n,\alpha}$ span a Hilbert space \mathcal{H}_{S+E} in their own right. \mathcal{H}_{S+E} is taken to be the state space for the joint system $S + E$. Equations (4.33) and (4.34) are independent of the bases used to express the vectors and the operators. The scheme for the states of $S + E$ encompasses the product expressions (4.33) and (4.34) and simply extends them so that arbitrary linear combinations of the product states $|\psi_S\rangle |\psi_E\rangle$ and product operators $A_S \otimes A_E$ can be dealt with. It should then come as no surprise that physical predictions, as per the measurement postulates of quantum mechanics, are independent of the

particular choice of orthonormal bases for the system and the environment used in the construction of the states for $S + E$.

We will not enter any prolonged discussions of the mathematics involved in combining two (or three, or ...) quantum systems, but simply note that reasonable application of the Dirac notation gets all calculations right. We are in the habit of writing the components in the tensor product of states in opposite order for bras and kets; be advised, not all authors do so.

Take now two two-state systems A and B with the orthonormal bases $\{|0\rangle, |1\rangle\}$. In elementary physics these could be spins; in quantum information theory, such systems are called qubits. By virtue of the construction of the state space for $A + B$, the state

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B) \quad (4.36)$$

is perfectly legitimate. However, it cannot be constructed in the form (4.33). Namely, take the most general states of system A and B ,

$$|\psi_A\rangle_A = \alpha_A |0\rangle_A + \beta_A |1\rangle_A, \quad |\psi_B\rangle_B = \alpha_B |0\rangle_B + \beta_B |1\rangle_B, \quad (4.37)$$

then the requirement that

$$|\psi_A\rangle_A |\psi_B\rangle_B = |\Psi^+\rangle \quad (4.38)$$

leads to the equations

$$\alpha_A \alpha_B = \beta_A \beta_B = \frac{1}{\sqrt{2}}, \quad \alpha_A \beta_B = \alpha_B \beta_A = 0, \quad (4.39)$$

which are contradictory. There is no way to express the state $|\Psi^+\rangle$ as a tensor product of two states, one from system A and one from system B . Likewise, it is not possible in general to express a linear operator on $A + B$ as a tensor product of two operators, one acting on A and one on B .

Given two systems S and E , the state $|\psi\rangle$ in $S + E$ is a *direct-product* state if there exists vectors $|\psi_S\rangle$ and $|\psi_E\rangle$ in S and E such that

$$|\psi\rangle = |\psi_S\rangle |\psi_E\rangle, \quad (4.40)$$

otherwise the state $|\psi\rangle$ is *entangled*. Entanglement is the other peculiarity of quantum mechanics, besides superposition, that has no analog in classical mechanics. Much of quantum information theory is about the analysis of entanglement. In particular, our state $|\Psi^+\rangle$ is one of the four *Bell states* that are used in *teleportation* of a quantum state. However, so far this line will not be pursued further.

The final ingredient of our general discussion of combined systems is the concept of *partial trace*. Given a tensor product of two operators in $S + E$, partial trace with respect to the system E is a linear mapping from the operators

on the system $S + E$ to the operators on the system S . It is defined as the result that obtains when one traces away the E part of the tensor product,

$$\text{Tr}_E(A_S \otimes A_E) = A_S \text{Tr}_E(A_E). \quad (4.41)$$

Linearity extends this form to an arbitrary operator on $S + E$. We write in the Dirac notation for an arbitrary operator in A in $S + E$

$$\text{Tr}_E A = \sum_{\alpha} \langle \alpha | A | \alpha \rangle, \quad (4.42)$$

with the understanding that $|\alpha\rangle$ and $\langle\alpha|$ only act on the environment part of the operator A , whatever that may mean in practice.

As an example, take a matrix representation of the operator A in the basis $\{|n\rangle|\alpha\rangle\}$, so that

$$A = I A I = \sum_{n,\alpha,m,\beta} |n\rangle|\alpha\rangle\langle\alpha|\langle n|A|m\rangle|\beta\rangle\langle\beta|\langle m| = \sum_{n,\alpha,m,\beta} |n\rangle|\alpha\rangle A_{m\beta}^{n\alpha} \langle\beta|\langle m|. \quad (4.43)$$

Then, doing the trace in the same basis which α and β already refer to, we have

$$\text{Tr}_E A = \sum_{\gamma} \langle \gamma | A | \gamma \rangle = \sum_{n,\alpha,m,\beta,\gamma} |n\rangle\langle\gamma|\alpha\rangle A_{m\beta}^{n\alpha} \langle\gamma|\beta\rangle\langle m| = \sum_{n,m} \left(\sum_{\alpha} A_{m\alpha}^{n\alpha} \right) |n\rangle\langle m|. \quad (4.44)$$

What is happening here is for every fixed n and m we have the operator on the E part of the form

$$A_E^{nm} = \sum_{\alpha,\beta} A_{m\beta}^{n\alpha} |\alpha\rangle\langle\beta|, \quad (4.45)$$

and the total operator A may be written

$$A = \sum_{m,n} |n\rangle\langle m| \otimes A_E^{nm}. \quad (4.46)$$

The partial trace takes the trace of each E operator in this sum, and adds the results. Since the trace of A_E^{nm} in the system E is independent of the orthonormal basis used to take the trace, so is also the operator resulting from the partial trace via the expression (4.42). Besides, if one first takes the partial trace and then traces the result over the remaining degrees of freedom, a combination of (4.43) and (4.44) gives the outcome

$$\text{Tr}_S(\text{Tr}_E A) = \sum_{n,\alpha} \langle \alpha | \langle n | A | n \rangle | \alpha \rangle = \text{Tr}_{A+B} A, \quad (4.47)$$

trace over the entire system $S + E$.

4.3 Density operator

We are now ready to tackle the extension of the concept of state of a quantum system from state vector to *density operator* or *density matrix*.

Density operator of statistical ensemble

Suppose that we may think that the physical system is in some well-defined state, but we do not know which state. Instead, one way or the other, we only know that a state $|\psi_i\rangle$ occurs with the probability p_i . When we calculate the expectation value for an arbitrary operator, we must first calculate the expectation value for each state $|\psi_i\rangle$, and then average the results using the probabilities for each state. Introducing an arbitrary orthonormal basis $\{|n\rangle\}$, we have

$$\begin{aligned}\langle A \rangle &= \sum_i p_i \langle \psi_i | A | \psi_i \rangle = \sum_i p_i \langle \psi_i | A I | \psi_i \rangle = \sum_{i,n} p_i \langle \psi_i | A | n \rangle \langle n | \psi_i \rangle \\ &= \sum_{i,n} p_i \langle n | \psi_i \rangle \langle \psi_i | A | n \rangle = \sum_n \langle n | \left(\sum_i p_i |\psi_i\rangle \langle \psi_i| \right) A | n \rangle \\ &= \text{Tr}(\rho A),\end{aligned}\tag{4.48}$$

where the density operator is defined as

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|. \tag{4.49}$$

Now, the expectation value of the density operator itself in an arbitrary state is

$$\langle \psi | \rho | \psi \rangle = \sum_i p_i \langle \psi | \psi_i \rangle \langle \psi_i | \psi \rangle = \sum_i p_i |\langle \psi | \psi_i \rangle|^2 \geq 0 \tag{4.50}$$

since the probabilities p_i must be nonnegative. Also, taking the trace of the density operator, we have

$$\text{Tr} \rho = \sum_{i,n} p_i \langle n | \psi_i \rangle \langle \psi_i | n \rangle = \sum_{i,n} p_i \langle \psi_i | n \rangle \langle n | \psi_i \rangle = \sum_i p_i \langle \psi_i | \psi_i \rangle = \sum_i p_i = 1, \tag{4.51}$$

since by convention all of our states are normalized to unity, and the probabilities must sum up to one. The density operator of a statistical ensemble is positive (hence, Hermitian), and has unit trace. Moreover, the expectation value of an arbitrary operator may be written $\langle A \rangle = \text{Tr}(\rho A)$.

Reduced density operator

Suppose next that we have combined two quantum systems, call them again, and suggestively, system and environment. Let the total state of the system be

$$|\psi\rangle = \sum_{n,\alpha} c_{n\alpha} |n\rangle |\alpha\rangle. \tag{4.52}$$

Now take an arbitrary operator A that acts only on the system degrees of freedom, i.e., is an identity operator on the environment. Then a somewhat

formal Dirac notation trickery gives the expectation value

$$\begin{aligned}
\langle A \rangle &= \langle \psi | A \otimes I_E | \psi \rangle \\
&= \sum_{n,m,\alpha,\beta} c_{n\alpha}^* c_{m\beta} \langle \alpha | \langle n | A \otimes I_E | m \rangle | \beta \rangle = \sum_{n,m,\alpha,\beta} c_{n\alpha}^* c_{m\beta} \langle \alpha | \beta \rangle \langle n | A | m \rangle = \\
&= \sum_{n,m,\alpha} c_{n\alpha}^* c_{m\alpha} \langle n | A | m \rangle = \sum_{n,m,\alpha} c_{n\alpha}^* c_{m\alpha} \langle n | A I_S | m \rangle \\
&= \sum_{n,m,p,\alpha} c_{n\alpha}^* c_{m\alpha} \langle n | A | p \rangle \langle p | m \rangle \\
&= \sum_p \langle p | \left\{ \left[\sum_{n,m} \left(\sum_{\alpha} c_{m\alpha} c_{n\alpha}^* \right) | m \rangle \langle n | \right] A \right\} | p \rangle \\
&= \text{Tr}_S(\rho_S A). \tag{4.53}
\end{aligned}$$

The *reduced density operator* is defined, obviously, as

$$\rho_S = \sum_{n,m} \left(\sum_{\alpha} c_{m\alpha} c_{n\alpha}^* \right) | m \rangle \langle n |. \tag{4.54}$$

It is an operator on the system degrees of freedom only. In fact, it equals the partial trace

$$\rho_S = \text{Tr}_E(|\psi\rangle\langle\psi|). \tag{4.55}$$

Now, simple calculations show that, for an arbitrary state vector $|\psi\rangle$ is S we have

$$\langle \psi | \rho_S | \psi \rangle = \sum_{\alpha} \left| \sum_n c_{n\alpha} \langle n | \psi \rangle \right|^2 \geq 0, \tag{4.56}$$

and

$$\text{Tr}_S \rho_S = \sum_{n,\alpha} |c_{n\alpha}|^2 = 1, \tag{4.57}$$

so that the reduced density operator, too, is a positive operator with unit trace. Moreover it, too, can be used to calculate the expectation value of any system operator as $\langle A_S \rangle = \text{Tr}_S(\rho_S A_S)$. The reduced density operator automatically emerges when we are interested in part of the system only.

Density operator: general

The general scheme is now obvious. We may call any positive (hence, Hermitian) operator with unit trace a *density operator*. Density operator will be our new and improved concept of the state of a quantum system. The way it works is that the expectation value of any observable A in the state ρ will be $\langle A \rangle = \text{Tr}(\rho A)$.

Specifically, the density operator ρ defines, or simply is, a *pure state* if there exists a state vector on the system $|\psi\rangle$ such that $\rho = |\psi\rangle\langle\psi|$, and a *mixed state* otherwise. For a pure state $\rho = |\psi\rangle\langle\psi|$ obviously $\text{Tr}(\rho A) = \langle \psi | A | \psi \rangle$, so that for

a pure state the density operator is just a trick way of writing the ket state. An easy practical rule reveals when a state is pure:

Theorem. The following statements about a density operator ρ are equivalent: (a) ρ is a pure state. (b) $\rho^2 = \rho$. (c) $\text{Tr}\rho^2 = \text{Tr}\rho = 1$.

Proof. Trivially, (a) \Rightarrow (b) \Rightarrow (c). Now assume (c), $\text{Tr}\rho^2 = 1$. ρ is a density operator, and so it has a spectral representation $\rho = \sum_n p_n |n\rangle\langle n|$ with nonnegative eigenvalues p_n that add up to unity, $\sum_n p_n = 1$. But by our assumption, also $\text{Tr}\rho^2 = \sum_n p_n^2 = 1$. These two sums can be simultaneously valid for the numbers p_n with $0 \leq p_n \leq 1$ only if one of the numbers equals one and all of the others are zero, so that $p_n = \delta_{n,n_0}$ for some n_0 . Hence the density operator is of the form $\rho = |n_0\rangle\langle n_0|$, and (c) \Rightarrow (a). \square

The reason why we have introduced density operators is that almost all states you encounter in nature are really mixed, and a discussion in terms of pure states, while possible, is contorted.

The set of density operators is *convex*: Take any density operators ρ_k and nonnegative numbers p_k such that $\sum_k p_k = 1$, $\rho = \sum_k p_k \rho_k$ is also a density operator. To see this, note that $\langle\psi|\rho|\psi\rangle = \sum_k p_k \langle\psi|\rho_k|\psi\rangle \geq 0$ holds for any $|\psi\rangle$, and also $\text{Tr}\rho = \sum_k p_k \text{Tr}\rho_k = \sum_k p_k = 1$. A combinations of density operators of the form $\sum_k p_k \rho_k$ is called a *mixture* of the states ρ_k . A mixture should be carefully distinguished from a superposition of pure states.

One possible origin for density operators is statistical mixture as in (4.49). By virtue of the spectral representation, every density operator can, in fact, be written in this form, and even in such a way that the states that are being mixed are orthonormal. Given a mixture, we may think of it as a system being in some pure states with some known probabilities. The question you should be asking yourself is, is the representation of a density operator as a mixture of pure states unique? The general answer is no. We do not study this matter any further, but just add the cryptic note that, as opposed to quantum mechanics, in classical mechanics the representation of any state as a mixture of pure states is unique.

Given a system in a mixed state, in a technique widely utilized in quantum information arguments, one can always construct a fictitious environment such that the combined system is in a pure state. This technique is called *purification*. For instance, take a system in the state with the density operator given by the spectral representation

$$\rho = \sum p_k |n\rangle\langle n|. \quad (4.58)$$

Now imagine duplicating the system, so that there is also an environment that is in this case identical with the system itself, and take the new system+environment in the pure state

$$|\psi\rangle = \sum_n \sqrt{p_n} |n\rangle |n\rangle. \quad (4.59)$$

Tracing away the made-up environment shows that the state of the system is (4.58). Obviously, this is not the only way to achieve purification.

Another trick of the trade is *Schmidt decomposition*:

Theorem For any pure state $|\psi\rangle$ of the total system $S + E$, it is always possible to find an orthonormal set of states of the system $\{|n\rangle_S\}$ and of the environment $\{|n\rangle_E\}$ and positive numbers λ_n satisfying $\sum_n \lambda_n^2 = 1$ such that

$$|\psi\rangle = \sum_i \lambda_n |n\rangle_S |n\rangle_E. \quad (4.60)$$

Proof. Suppose the dimension of the Hilbert space of the system S is equal or smaller than the dimension of the environment E ; if not, swap S and E in the following argument. Take as the vectors $\{|n\rangle_S\}$ the eigenvectors of the reduced density operator of the system ρ_S with the corresponding probabilities p_n , and take an arbitrary basis $\{|\alpha\rangle_E\}$ for the environment. We thus have the decomposition

$$|\psi\rangle = \sum_{n,\alpha} c_{n\alpha} |n\rangle_S |\alpha\rangle_E \quad (4.61)$$

where the coefficients $c_{n\alpha}$ must satisfy

$$\sum_{\alpha} c_{n\alpha} c_{m\alpha}^* = p_n \delta_{nm} \quad (4.62)$$

by Eq. (4.49). Therefore the vectors $|n\rangle'_E = \sum_{\alpha} c_{n\alpha} |\alpha\rangle_E$ are orthogonal, and we have the decomposition

$$|\psi\rangle = \sum_n |n\rangle_S |n\rangle'_E. \quad (4.63)$$

Dropping from the sum those terms n for which $p_n = 0$ and correspondingly $|n\rangle'_E = 0$ and defining the orthonormal vectors $|n\rangle_E = |n\rangle'_E / \sqrt{p_n}$ for the rest of the n , we have the decomposition (4.60) with $\lambda_n = \sqrt{p_n}$. \square

For an arbitrary combination of qubits one may prove that it is always possible to realize an arbitrary unitary transformation by combining a few very simple operations on individual qubits and pairs of qubits. This leads to another hallmark notion of quantum information science that it is possible in principle to realize an arbitrary unitary transformation on a state. This idea is occasionally implicit in our discussions.

4.4 Postulates of quantum mechanics revisited

If density operator is to be the new concept of state, then the postulates of quantum mechanics at the very least need to be restated. We are about to get to such a restatement, and while at that, also to a significant extension. Obviously, whatever we are going to say has to be compatible with what we knew earlier about quantum mechanics. There is another side to the coin, too: If we are extending what we have already said, then we should also ask for the conditions under which the extension is valid and unique. For the most part we are going to keep quiet about the latter issue.

Let us get the easy part, time evolution of the density operator in a closed system, out of the way first. For a pure state $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$ we know how the ket and the bra should evolve, c.f. (4.16), and so we know what should happen to the pure-state density operator as well,

$$i\hbar \frac{d}{dt}(|\psi(t)\rangle\langle\psi(t)|) = H|\psi(t)\rangle\langle\psi(t)| - |\psi(t)\rangle\langle\psi(t)|H = [H, |\psi(t)\rangle\langle\psi(t)|]. \quad (4.64)$$

Suppose next that at some time t_0 we have established a representation of the density operator as a mixture of pure states,

$$\rho(t_0) = \sum_n p_n |\psi_n(t_0)\rangle\langle\psi_n(t_0)|. \quad (4.65)$$

Spectral representation works here just as well as any other decomposition of a mixed state. If we require that the evolution of the density operator is linear, the evolution up to some time t should produce the same mixture of the projectors that have evolved from the projectors at time t_0 ,

$$\rho(t) = \sum_n p_n |\psi_n(t)\rangle\langle\psi_n(t)|. \quad (4.66)$$

Combining the above arguments we then have

$$i\hbar \frac{d}{dt}\rho(t) = \sum_n p_n [H, |\psi_n(t)\rangle\langle\psi_n(t)|] = [H, \sum_n p_n |\psi_n(t)\rangle\langle\psi_n(t)|] = [H, \rho(t)]. \quad (4.67)$$

In short, we have the *Liouville-von Neumann* equation for the density operator of a closed system,

$$i\hbar \frac{d}{dt}\rho = [H, \rho]. \quad (4.68)$$

This is the counterpart of the time dependent Schrödinger equation in the version of quantum mechanics that uses the density operator to express the state.

We have argued earlier that, by virtue of the Schrödinger equation, there is a unitary time evolution operator U so that the state vector at time t is given in terms of the state vector at t_0 by $|\psi(t)\rangle = U|\psi(t_0)\rangle$. For a pure state $\rho = |\psi\rangle\langle\psi|$ the same evolution reads $\rho(t) = U\rho(t_0)U^\dagger$, and since by assumption the evolution of the density operator is linear, the same applies to an arbitrary, even mixed-state, density operator. Liouville-von Neumann equation is a first-order differential equation in time so that $\rho(t_0)$ uniquely determines $\rho(t)$, and by our construction we must therefore have the same result $\rho(t) = U\rho(t_0)U^\dagger$ also from the Liouville-von Neumann equation. The time evolution according to Liouville-von Neumann equation is a unitary transformation on the density operator.

Next comes the question of measurement postulates in quantum mechanics. First, it is customary to alter the approach slightly. Instead of asking what is the measured result, we formulate everything in terms of yes-no questions. So, instead of asking which n in $\hbar(n + \frac{1}{2})\omega$ we got when we measured the energy of

a harmonic oscillator, we say: We got $n = 0$, true or false; $n = 1$, true or false; and so on.

Our initial reformulation of the quantum measurement postulates corresponds to what is called *projective measurements*. We associate with each such true-or-false question a subspace \mathcal{S}_m in the Hilbert space, and an orthogonal projection onto this subspace P_m . Typically, corresponding to some observable (Hermitian operator) A , the subspace \mathcal{S}_m would be the subspace associated with the eigenvalue a_m , and the true-false question is whether the eigenvalue a_m was found. We insist that all possible measurement results are covered, and therefore require

$$\sum_m P_m = I. \quad (4.69)$$

Moreover, in order to ensure that an immediate repetition of the measurement will also repeat the result, we insist that the projections are orthogonal, $P_m P_n = \delta_{mn} P_n$. The measurement postulates could now be stated in the following way:

- (a) A measurement gives the value true to precisely one assertion m , whereupon it is said that the measurement result was m .
- (b) The probability for the measurement result m equals $P(m) = \text{Tr}(P_m \rho)$.
- (c) Immediately after a measurement that gave the result m , the state of the system is

$$\rho_m = \frac{P_m \rho P_m}{P(m)} = \frac{P_m \rho P_m}{\text{Tr}(P_m \rho P_m)}. \quad (4.70)$$

At this point we pause for some practicalities and common terminology. If one wants numbers instead of formal results, operator equations per se rarely suffice. Instead, one represents, say, the density operator as a matrix in some orthonormal basis $\{|n\rangle\}$, as a set of numbers $\rho_{nm} = \langle n|\rho|m\rangle$, and does tangible matrix calculations, say, numerically. Given the basis, the projector $|n\rangle\langle n|$ is the one corresponding to the question if the system is in the state $|n\rangle$, and the probability for finding the system in this state is $P(n) = \text{Tr}(|n\rangle\langle n|\rho) = \rho_{nn}$. The diagonal elements of the density matrix ρ_{nn} are correspondingly called *populations*. Off-diagonal elements ρ_{nm} with $m \neq n$ are similarly called *coherences*. It should be kept in mind that coherences and populations are always relative to a given basis.

This is about where the buck stops in statistical mechanics. However, for the sheer intrinsic interest of it, we will broaden the concept of measurement all the way to the scope that appears in modern literature on quantum mechanics and its applications. Suppose we initially have our system in the state ρ and an environment in some prepared state $|e_0\rangle$. Then some unitary transformation on the system plus environment U is applied, and finally a projective measurement is applied to see if the environment is found in some state $|e_k\rangle$. This is in fact how most measurements of microscopic quantum phenomena work; just think of the “environment” as the “apparatus.” Incidentally, it should not be forgotten that the problem of the quantum-classical interface does not go away with this

scheme, it is just gets moved elsewhere. At some point we just have to say that there was a measurement that produced a definite result.

At any rate, introducing a system basis $\{|n\rangle\}$, at the point of the measurement the state and the projectors for the measurement are

$$\rho_{S+E} = U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger, \quad P_k = \sum_n |n\rangle\langle e_k| \langle e_k| \langle n|. \quad (4.71)$$

Hence the probability for the outcome k is

$$\begin{aligned} P(k) &= \text{Tr}(\rho_{S+E} P_k) = \sum_n \langle e_k| \langle n| [U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger] |n\rangle |e_k\rangle \\ &= \sum_{npq\alpha\beta} \langle e_k| \langle n| U|p\rangle\langle\alpha| \langle\alpha| \langle p| (\rho \otimes |e_0\rangle\langle e_0|) |q\rangle\langle\beta| \langle\beta| \langle q| U^\dagger |n\rangle |e_k\rangle \\ &= \sum_{npq} \langle e_k| \langle n| U|p\rangle\langle e_0| \langle p| \rho |q\rangle \langle e_0| \langle q| U^\dagger |n\rangle |e_k\rangle \\ &= \sum_n \langle n| M_k \rho M_k^\dagger |n\rangle = \text{Tr}_S(M_k \rho M_k^\dagger) = \text{Tr}_S(M_k^\dagger M_k \rho). \end{aligned} \quad (4.72)$$

Here the definition and a frequently used shorthand for the *measurement operator* M_k , an operator in the system degrees of freedom, is

$$M_k = \sum_{n,m} (\langle e_k| \langle n| U |m\rangle |e_0\rangle) |n\rangle \langle m| \equiv \langle e_k| U |e_0\rangle. \quad (4.73)$$

Using the symbolic shorthand, we also easily see that

$$\begin{aligned} P_k \rho_{S+E} P_k &= \sum_{nm} |n\rangle |e_k\rangle \langle e_k| \langle n| U (\rho \otimes |e_0\rangle\langle e_0|) U^\dagger |m\rangle |e_k\rangle \langle e_k| \langle m| \\ &= \sum_{nm} |n\rangle |e_k\rangle \langle n| M_k \rho M_k^\dagger |m\rangle \langle e_k| \langle m|. \end{aligned} \quad (4.74)$$

Inasmuch as only the state of the system after the measurement is of interest, we find by taking the partial trace and normalizing properly that

$$\rho_k = \frac{\text{Tr}_E(P_k \rho_{S+E} P_k)}{P(k)} = \frac{M_k \rho M_k^\dagger}{\text{Tr}_S(M_k \rho M_k^\dagger)}. \quad (4.75)$$

Note that

$$\langle \psi | M_k \rho M_k^\dagger | \psi \rangle = \langle \psi | M_k^\dagger \rho M_k | \psi \rangle^* = (M_k \psi, \rho M_k \psi)^* \geq 0, \quad (4.76)$$

so that ρ_k is a valid density operator [as long as $P(k) > 0$].

The operator $M_k^\dagger M_k$ is, of course positive. Moreover, by using the shorthand notation again for the case that k enumerates all possible measurement outcomes, we find somewhat symbolically⁴ that

$$\sum_k M_k^\dagger M_k = \sum_k \langle e_0| U^\dagger |e_k\rangle \langle e_k| U |e_0\rangle = \langle e_0| U^\dagger U |e_0\rangle = I_S. \quad (4.77)$$

⁴If you are not satisfied, just do the calculation by means of full expansions also in the system basis.

This result confirms that the probabilities $P(k)$ from (4.72) sum up to one, as they should.

In fact, our introduction to the measurement operators also works in reverse: Given operators M_k satisfying nothing more than $\sum_k M_k^\dagger M_k = I_S$, it is always possible to construct a fictitious environment and a unitary evolution for the system+environment such that the operators M_k are the measurement operators according to (4.73). This construction is the subject of an exercise.

Let us summarize the reformulation of the axioms of quantum mechanics thus far:

1. A Hilbert space represent a physical system, and a density operator ρ on the Hilbert space represents a state.
2. A measurement on a system is characterized by a set of the outcomes $\{k\}$ associated with operators M_k on the system such that $\sum_k M_k^\dagger M_k = I$.
3. (a) Each measurement produces one of the outcomes k .
 (b) The probability for the outcome k is $P(k) = \text{Tr}(M_k \rho M_k^\dagger)$.
 (c) Immediately after a measurement that gave the outcome k , the state of the system is

$$\rho_k = \frac{M_k \rho M_k^\dagger}{P(k)}.$$

4. Evolution of the state $\rho(t)$ is governed by a trace-preserving completely positive map.

It may happen that once the measurement is done, the system is discarded and the state after the measurement is irrelevant. Then only the probabilities of the measurement results are relevant. Now, we trivially have

$$\text{Tr}[M_k \rho M_k^\dagger] = \text{Tr}[M_k^\dagger M_k \rho] = \text{Tr}[E_k \rho]; \quad E_k \equiv M_k^\dagger M_k. \quad (4.78)$$

The operators E_k make a “positive operator valued measure” (POVM), a set of positive operators E_k satisfying $\sum_k \text{Tr}(E_k A) \leq \text{Tr} A$ for any positive A . A POVM completely characterizes the probabilities of the outcomes of a measurement. In fact, our POVM satisfies $\sum_k E_k = I$ and therefore $\sum_k \text{Tr}(E_k A) = \text{Tr} A$, but the genuine inequality < 1 might be formally allowed to account for missing measurement results. It should be noted that a POVM alone will not suffice to specify the post-measurement state: Given a positive E_k , one may always define an operator $M_k = \sqrt{E_k}$ that satisfies $E_k = M_k^\dagger M_k$, but the same also holds for $M_k = \sqrt{E_k} U_k$ for arbitrary unitary operators U_k .

This is pretty much the end of the road at this time. Straightforward variations of the theme, for instance, alternative measurements occurring at certain probabilities, may be considered easily. On the other hand, one can envisage schemes that outright break the present framework, but as far as we know, they have not proven relevant enough that anyone would had worked out the required extensions of the postulates of quantum mechanics.

4.5 Quantum operations

Suppose that we do not read out the measurement result, or discard it, the state of the system after the measurement is obviously

$$\rho' = \sum_k P(k) \rho_k = \sum_k M_k \rho M_k^\dagger. \quad (4.79)$$

On the immediate level, this is a linear operation that converts a linear operator on a Hilbert space to another linear operator. Below we tend to call this type of operator-to-operator operations “maps”, or “operations”, and denote them as, say $\rho' = \mathcal{L}(\rho)$, or even $\rho' = \mathcal{L}\rho$. Just like linear operators inherit the structure of linear vector space from vectors, operations inherit the structure of linear vector space from operators: $(\lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2)A \equiv \lambda_1(\mathcal{L}_1 A) + \lambda_2(\mathcal{L}_2 A)$.

The particular map (4.79) produces a convex combination of density operators, hence a valid density operator in itself. An operation that transform a positive operator into a positive operator is called a *positive map*, so that the mapping $\rho \rightarrow \rho'$ is a positive map.

In fact, this map has the even stronger property that it is *completely positive* (CP). First, imagine adding an arbitrary environment E to the system, and take an arbitrary positive operator O on $S + E$. Then map the system part of O with \mathcal{L} and leave the environment part untouched, and finally take the expectation value of the ensuing operator in the state $|\psi\rangle$ of $S + E$. We find as a result of these operations the expression

$$\langle \psi | \sum_k (M_k \otimes I_E) O (M_k^\dagger \otimes I_E) | \psi \rangle = \sum_k \langle \psi_k | O | \psi_k \rangle \geq 0 \quad (4.80)$$

with

$$|\psi_k\rangle = M_k^\dagger \otimes I_E |\psi\rangle. \quad (4.81)$$

This means that while we apply our operation \mathcal{L} on the S part only, every positive operator on $S + E$ nonetheless gets mapped to a positive operator. This is what we should require of valid transformations of the density operators for the combined system $S + E$; if we do something that keeps the reduced density matrix ρ_S valid and does not touch the environment at all, the joint density operator ρ_{S+E} should also remain valid. Besides, there are positive maps \mathcal{L} that are not completely positive, so that the distinction between positive and completely positive is nontrivial.

Completely positive maps are an extremely useful concept in quantum mechanics. This is mostly because of the following observation, which is in a sense the inverse of our arguments above:

Theorem. Every CP map \mathcal{L} of operators on a Hilbert space into operators on the same system can be written in the form

$$\mathcal{L}(A) = \sum_k K_k A K_k^\dagger \quad (4.82)$$

for some set of system operators K_k .

Proof. Let us pick an environment E with the same dimension as the system, and denote by \mathcal{I} the identity map on environment operators. Let us furthermore define an auxiliary state on $S + E$

$$|\Psi\rangle = \sum_n |n\rangle_S |n\rangle_E, \quad (4.83)$$

where the system and environment states both make an orthonormal basis. As a final preparatory item we define the operator on $S + E$,

$$O = (\mathcal{L} \otimes \mathcal{I})(|\Psi\rangle\langle\Psi|). \quad (4.84)$$

Let us now pick arbitrary system states and the corresponding “shadow” states of the environment as

$$|\psi\rangle_S = \sum_n \psi_n |n\rangle_S, \quad |\tilde{\psi}\rangle_E = \sum_n \psi_n^* |n\rangle_E; \quad |\phi\rangle_S = \sum_n \phi_n |n\rangle_S, \quad |\tilde{\phi}\rangle_E = \sum_n \phi_n^* |n\rangle_E. \quad (4.85)$$

We have

$${}_E\langle\tilde{\psi}|O|\tilde{\phi}\rangle_E = {}_E\langle\tilde{\psi}|\left[\sum_{nm} \mathcal{L}(|m\rangle_{SS}\langle n|) \otimes (|m\rangle_{EE}\langle n|)\right]|\tilde{\phi}\rangle_E \quad (4.86)$$

$$= \sum_{nm} \psi_m \phi_n^* \mathcal{L}(|m\rangle_{SS}\langle n|) \quad (4.87)$$

$$= \mathcal{L}(|\psi\rangle_{SS}\langle\phi|). \quad (4.88)$$

Since the map \mathcal{L} is completely positive, the operator O is positive and can be written in the form

$$O = \sum_k {}_{S+E}|k\rangle\langle k|_{S+E}, \quad (4.89)$$

where $|k\rangle_{S+E}$ are orthogonal though not necessarily normalized vectors in $S+E$. Also define operators on the system degrees of freedom whose actions on the arbitrary system state $|\psi\rangle_S$ are

$$K_k(|\psi\rangle_S) = {}_E\langle\tilde{\psi}|k\rangle_{S+E}. \quad (4.90)$$

These are obviously linear operators. Moreover, we have

$$\begin{aligned} \sum_k K_k(|\psi\rangle_{SS}\langle\phi|)K_k^\dagger &= \sum_k {}_E\langle\tilde{\psi}|k\rangle_{S+ES+E}\langle k|\tilde{\phi}\rangle_E \\ &= {}_E\langle\tilde{\psi}|O|\tilde{\phi}\rangle_E \\ &= \mathcal{L}(|\psi\rangle_{SS}\langle\phi|). \end{aligned} \quad (4.91)$$

The action of the combination of operator products in the first form of this chain of equalities on an arbitrary dyad of system states is the same as the action of the completely positive map \mathcal{L} , hence so is the action on all system operators. We have found system operators K_k such that equality (4.82) holds true. \square

In this general context the operators K_k are often called Kraus operators. If they satisfy $\sum_k K_k^\dagger K_k = I$, the map also preserves the trace. Suppose now that we have the Kraus operators K_k and a unitary matrix with the elements u_{lk} , and we define the operators

$$L_l = \sum_k u_{lk} K_k. \quad (4.92)$$

Then we have the CP mapping of the operator A

$$A' = \sum_l L_l A L_l^\dagger = \sum_{l,p,q} u_{lp} K_p A u_{lq}^* K_q = \sum_{l,p,q} u_{ql}^\dagger u_{lp} K_p A K_q = \sum_k K_k A K_k^\dagger. \quad (4.93)$$

The Kraus operators K_k and L_l therefore define the same completely positive map, so that the map does not uniquely determine the Kraus operators. Though we do not go into the details, such unitary ambiguity completely exhausts the nonuniqueness: If two sets of Kraus operators $\{K_k\}$ and $\{L_l\}$ define the same CP map, then there exists a unitary matrix u_{lk} such that Eq. (4.92) holds true. If the sets $\{K_k\}$ and $\{L_l\}$ are different sizes, one needs to pad the shorter set with zero operators so that the number of Kraus operators is the same.

Completely positive maps that preserve the trace are the most general possible transformations of a density operator. In formal descriptions, however, it is customary to relax the condition of conservation of trace, and define a *quantum operation* \mathcal{E} as any CP map such that for any positive operator A , $\text{Tr}[\mathcal{E}(A)] \leq \text{Tr} A$. The inequality qualifies the kind of operator products such as $\mathcal{E}_k(\rho) = M_k \rho M_k^\dagger$ that occur in measurement theory as quantum operations. A quantum operation is trace preserving if the equality holds, i.e., $\text{Tr}[\mathcal{E}(A)] = \text{Tr} A$ for every positive A .

All legitimate linear transformations of the density operator must be trace preserving quantum operations. Quantum operations can always be characterized by Kraus operators, and obviously, for a trace-preserving quantum operation the Kraus operators must satisfy $\sum_k K_k^\dagger K_k = I$. A composition of two, or arbitrary many, trace preserving quantum operations is obviously also a trace-preserving quantum operation; in fact, if

$$\mathcal{E}(\rho) = \sum_k K_k \rho K_k^\dagger, \quad \mathcal{F}(\rho) = \sum_l L_l \rho L_l^\dagger, \quad (4.94)$$

with

$$\sum_k K_k^\dagger K_k = I, \quad \sum_l L_l^\dagger L_l = I, \quad (4.95)$$

then

$$\mathcal{F}(\mathcal{E}(\rho)) = \sum_{kl} L_l K_k \rho K_k^\dagger L_l^\dagger = \sum_{kl} (L_l K_k) \rho (L_l K_k)^\dagger, \quad (4.96)$$

with

$$\sum_{kl} (L_l K_k)^\dagger (L_l K_k) = \sum_{kl} K_k^\dagger L_l^\dagger L_l K_k = \sum_k K_k^\dagger K_k = I. \quad (4.97)$$

The transformation from pre-measurement to post-measurement state in the case when the measurement result is discarded, Eq. (4.79), is a trace preserving completely positive map, and so is unitary time evolution. But the scheme may be stretched even further. For instance, photon detectors have a finite efficiency and miss photons. No problem: There is a quantum operation to describe the effect of the dropped counts. The power and beauty of quantum operations and Kraus operators is in the generality of these concepts.

In fact, the generality goes even further. One may define a quantum operation as a mapping between two different Hilbert spaces, and most theorems about quantum operations remain intact. As an example, trace is a quantum operator that maps the operators of the Hilbert space \mathcal{H} to operators on one-dimensional Hilbert space of the complex numbers, in other words, to complex numbers. Moreover, positive operators get mapped to positive operators, i.e., to nonnegative real numbers. Let us introduce a trivial orthonormal basis $|b\rangle$ to the space of complex numbers and an orthonormal basis $\{|k\rangle\}_k$ for the Hilbert space \mathcal{H} , then there is, for instance, a set of Kraus operators corresponding to trace:

$$(\text{Tr } A) |b\rangle\langle b| = \sum_k K_k A K_k^\dagger; \quad K_k = |b\rangle\langle k|. \quad (4.98)$$

Similarly, partial trace qualifies as a quantum operation. This kind of arguments may seem gimmicky, but they are extremely useful in the mathematics of quantum information.

4.6 Relaxation terms in time evolution

As we know, for a closed system the time evolution of the density operator is given by the Liouville-von Neumann (LvN) equation. However, there really is no such thing as a closed system (except possibly the universe). As an example of the power of our extended scheme of quantum mechanics, we study here briefly how the environment affects the evolution of the density operator of the system.

One could think of taking the LvN equation for the system plus environment, solving, and tracing out the environment to see what happens. Unfortunately, such a process is usually impossible in practice. Fortunately, it turns out that under conditions that are quite common the effect of the environment can be described to an excellent approximation by the evolution of the system only, which has the following properties: (i) The evolution of the state is linear. (ii) The evolution has the Markov property: the present state determines the future state. (iii) The evolution of the state is continuous enough that it is described by a first-order differential equation in time. So, the evolution of the system *alone* is described by a *master equation* of the form

$$\dot{\rho} = -\frac{i}{\hbar} [H, \rho] + \mathcal{L}\rho. \quad (4.99)$$

We first have the usual LvN part, as if we had an isolated system. Next come *relaxation terms* that characterize the effect of the environment. \mathcal{L} is some map on system operators.

Let us momentarily forget about the LvN term, then we could write the formal solution for the evolution of the density operator due to the relaxation terms as

$$\rho(t) = e^{\mathcal{L}(t-t_0)} \rho(t_0). \quad (4.100)$$

From our freshly extended postulates of quantum mechanics we require that this evolution of ρ is a trace preserving quantum operation. The key result in this direction, and this remains unchanged if we also incorporate the LvN part, is the following: The evolution in (4.100) is a completely positive trace preserving map of the density operator if and only if the relaxations terms are of what is known as the *Lindblad form*,

$$\mathcal{L}\rho = \sum_k [2L_k \rho L_k^\dagger - L_k^\dagger L_k \rho - \rho L_k^\dagger L_k], \quad (4.101)$$

where L_k are some system operators. To recap, relaxation terms of the Lindblad form, and only relaxation terms of the Lindblad form, keep any density operator a legitimate density operator as time goes on.

That the form (4.101) is sufficient for a valid system density operator is easy to see. Namely, starting from $\rho(t) \equiv \rho$, we have to first order in the time differential dt

$$\rho(t+dt) = \rho + dt \sum_k [2L_k \rho L_k^\dagger - L_k^\dagger L_k \rho - \rho L_k^\dagger L_k]. \quad (4.102)$$

First, $\rho(t+dt)$ is hermitian, and the trace is preserved. Second, given an arbitrary vector $|\psi\rangle$, for a small enough dt we can have $\langle\psi|\rho(t+dt)|\psi\rangle < 0$ only if $\langle\psi|\rho|\psi\rangle = 0$. But since ρ is positive, $\langle\psi|\rho|\psi\rangle = 0$ implies that $\rho|\psi\rangle = 0$, so that we have

$$\langle\psi|\rho(t+dt)|\psi\rangle = 2dt \sum_k \langle\psi|L_k \rho L_k^\dagger|\psi\rangle = 2dt \sum_k \langle L_k^\dagger \psi|\rho|L_k^\dagger \psi\rangle \geq 0 \quad (4.103)$$

by the positivity of ρ . Thus, evolution according to Eq. (4.101) preserves positivity of the density operator. This argument can be amended easily to cover complete positivity, so that the Lindblad form is sufficient for a valid evolution of the density operator. Unfortunately, proving the converse, namely, that the Lindblad form is also necessary, seems to be an excruciating exercise, and requires added assumptions about the time evolution.

4.7 Exercises

4.1. Linear operators of a linear space make a vector space in their own right when endowed with the natural sum and multiplication with a scalar. For instance, given the operators A and B , the operator $A+B$ is defined in such a way that, when acting on an arbitrary vector $|\psi\rangle$, the result is $(A+B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle$. Show that the complex function of two operators defined as $(A,B) = \text{Tr}(A^\dagger B)$ is an inner product. This is a situation in which infinite-dimensional spaces make trouble, ignore it quietly.

4.2.

- (a) Show that all eigenvalues of a unitary operator have unit modulus.
- (b) Show that an operator U is unitary if and only if there is a Hermitian operator A such that $U = e^{iA}$.
- (b) Suppose the evolution of a closed quantum system is generated by a time-independent Hamiltonian H . What is the significance of the operator $e^{-iH(t-t_0)/\hbar}$?

4.3. Take a subspace \mathcal{S} of a Hilbert space \mathcal{H} . Suppose we have defined an operator $U : \mathcal{S} \rightarrow \mathcal{H}$ with the property that $(U\psi, U\phi) = (\psi, \phi)$ for all vectors ψ and ϕ in the subspace \mathcal{S} . Show that the operator U can be extended from \mathcal{S} to the whole Hilbert space \mathcal{H} in such a way that the result is a unitary operator on \mathcal{H} . HINTS: Gram-Schmidt construction and property (iii) of unitary operators.

4.4. The tensor product state $|\psi_S\rangle|\psi_E\rangle$ for two systems simply says that S is in the state $|\psi_S\rangle \in \mathcal{H}$ and E in the state $|\psi_E\rangle \in \mathcal{R}$. The idea behind the Hilbert space for the joint system $S + E$, call it $\mathcal{H} \otimes \mathcal{R}$, is to define a multiplication of a tensor product state by a scalar and addition of two tensor product states in such a way that the ensuing linear-space operations satisfy

$$(\lambda|\psi\rangle)|\phi\rangle = |\psi\rangle(\lambda|\phi\rangle) = \lambda|\psi\rangle|\phi\rangle \quad (4.104)$$

$$(|\psi_1\rangle + |\psi_2\rangle)|\phi\rangle = |\psi_1\rangle|\phi\rangle + |\psi_2\rangle|\phi\rangle \quad (4.105)$$

$$|\psi\rangle(|\phi_1\rangle + |\phi_2\rangle) = |\psi\rangle|\phi_1\rangle + |\psi\rangle|\phi_2\rangle \quad (4.106)$$

- (a) By assuming (4.104)-(4.106), show how the tensor product $|\psi\rangle|\phi\rangle$ is expanded out in the basis $\{|n\rangle|\alpha\rangle\}$, given that we know the expansions of $|\psi\rangle$ and $|\phi\rangle$ in the bases $\{|n\rangle\}$ and $\{|\alpha\rangle\}$.
- (b) Conversely, given the expansion of the tensor product $|\psi\rangle|\phi\rangle$ as in part (a) and the usual linear-space structure of the product space $\mathcal{H} \otimes \mathcal{R}$, show that (4.104)-(4.106) hold true.

4.5. Take a qubit with the two states $|\pm\rangle$. Write down the density operators corresponding to the superposition state $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ and to the 50/50 mixture of the states $|+\rangle$ and $|-\rangle$.

4.6. Given the Bell state $|\Psi^+\rangle$ of two qubits, what is the state of either one of its individual subsystems?

4.7.

- (a) Show that for any density operator ρ and normalized $|\psi\rangle$ the inequalities $0 \leq \langle\psi|\rho|\psi\rangle \leq 1$ hold true.
- (b) Given two density operators ρ_1 and ρ_2 , show that $0 \leq \text{Tr}(\rho_1\rho_2) \leq 1$.
- (c) Suppose that at least one of the density operators ρ_1 and ρ_2 is not a pure state. Show that no mixture (convex combination) of the density operators ρ_1 and ρ_2 other than the trivial ones ($\rho = \rho_1$ and ρ_2) can be a pure state.

(d) Take two pure states ρ_1 and ρ_2 . So that a mixture of the two states is pure only if either the mixture is trivial, or $\rho_1 = \rho_2$

This exercise tells you that the pure states are precisely the “corners” of the convex set of density operators; the only way to get a pure state as a mixture of two states is the trivial way, i.e., mix the state with itself.

4.8.

(a) A density operator can always be written in the form $\rho = \sum_k |\psi_k\rangle\langle\psi_k|$, where the vectors $|\psi_k\rangle$ need neither be normalized, nor orthogonal. Suppose we have two representations of this kind with the vectors $|\psi_k\rangle$ and $|\phi_k\rangle$. For convenience, assume that there are equally many vectors $|\psi_k\rangle$ and $|\phi_k\rangle$; if not, you can always pad the smaller set with zero vectors to achieve the same size. Show that the corresponding density operators ρ_ψ and ρ_ϕ are the same if there exists a unitary matrix u_{pq} such that $|\psi_p\rangle = \sum_q u_{pq}|\phi_q\rangle$. The decomposition of a density operator into a mixture of pure states is in general not unique.

(b) [A tough one.] Show that the reverse also holds: ρ_ψ and ρ_ϕ are the same density operator only if a unitary matrix u_{pq} as in part (a) exists.

4.9. Imagine two spins prepared in a singlet state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle).$$

The spins are then transported far apart, but so that the entangled state is preserved, and the value of the first spin is measured ($|\downarrow\rangle$ or $|\uparrow\rangle$). Assume that the measurement result is not known at the site of the second spin, for instance because at the time the site is outside of the light cone of the measurement. What is the effect of the measurement on the state of the other spin?

4.10. Starting from the Liouville-von Neumann equation, show that the property “ ρ is a pure state” or “ ρ is a mixed state” is preserved in the time evolution of a closed system. In particular, if the system is in a mixed state at some point in time, it will always be in a mixed state and can never be described by a state vector.

4.11. Take a collection of measurement operators M_k in a system S . Now imagine a measurement system (apparatus) E with a reference state $|e_0\rangle$ and an orthonormal basis $\{|e_k\rangle\}$, one state k for each measurement operator M_k . Define a linear operator whose action on a product state $|\psi\rangle \otimes |e_0\rangle$ is

$$U(|\psi\rangle \otimes |e_0\rangle) = \sum_k M_k |\psi\rangle \otimes |e_k\rangle. \quad (4.107)$$

(a) Show that this operator preserves inner products of vectors of the form $|\psi\rangle \otimes |e_0\rangle$:

$$(\langle e_0| \otimes \langle \psi|) U^\dagger U (|\phi\rangle \otimes |e_0\rangle) = \langle \psi|\phi\rangle. \quad (4.108)$$

- (b) Show that the operator U may be extended so that it is a unitary on the whole state space for $S + E$.
- (c) Show that, once the operator has been extended as in part (b), it satisfies

$$\langle e_k | U | e_0 \rangle = M_k .$$

In class it was essentially shown that, given projective measurements P_k for the environment and a unitary U to evolve $S + E$, operators of the form $M_k = \text{Tr}_E(P_k U P_k)$ are legitimate measurement operators. The present result shows that the reverse also holds: Every set of measurement operators corresponds to projective measurements in a suitable piece of apparatus.

4.12. Let us consider qubits living in a two-dimensional state space spanned by the orthonormal basis $\{|0\rangle, |1\rangle\}$. For purposes of quantum encryption, the ever-busy Alice sends Bob one or the other of the states $|\psi_1\rangle = |0\rangle$ or $|\psi_2\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Now, it transpires that no single measurement can reliably distinguish between two nonorthogonal states, so that Bob cannot tell for sure which state Alice sent. However, there are schemes in which he will not *misidentify* the state:

- (a) Define the operators

$$E_1 = \frac{\sqrt{2}}{1 + \sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}, \quad E_2 = \frac{\sqrt{2}}{1 + \sqrt{2}} |1\rangle\langle 1|, \quad E_3 = I - E_1 - E_2 . \quad (4.109)$$

Show that these are positive and satisfy $\sum_k E_k = I$. Therefore, they make a POVM.

- (b) Show that if the measurement E_1 or E_2 gives an affirmative answer, the incoming state must have been, respectively, $|\psi_1\rangle$ or $|\psi_2\rangle$.
- (c) However, show that the measurement result E_3 does not unambiguously distinguish between the possible incoming states.
- (d) Suppose that Alice sends one of the states $|\psi_1\rangle$ or $|\psi_2\rangle$ at random with equal probabilities $1/2$. What is the probability that this measurement scheme cannot reliably distinguish between the states?

4.13. In the previous problem a POVM was specified that never misidentifies one of the two quantum states coming in, but may give an ambiguous result. In the example, when Alice sent both of the states with the probability $1/2$, the probability for an ambiguous result was $1/\sqrt{2}$, i.e., you can identify the state with certainty only about 30% of the time. While this sounds rather measly, it is the best one can do: The POVM of the problem has the smallest possible probability for an ambiguous result under the given conditions (incoming states and their emission probabilities). Prove this.

4.14. Consider again a two-state system spanned by the vectors $\{|1\rangle, |2\rangle\}$, ignoring the unitary (Hamiltonian) time evolution that has no bearing on the

present argument. Only relaxation terms remain, which we write as

$$\dot{\rho}_{11} = \Gamma \rho_{22}, \quad \dot{\rho}_{22} = -\Gamma \rho_{22}, \quad \dot{\rho}_{21} = -\gamma \rho_{21}, \quad \dot{\rho}_{12} = -\gamma \rho_{12}.$$

Show that for $\gamma < \Gamma/2$ it is always possible to construct an example in which the density operator will not remain positive during the time evolution. The physics lesson is that you cannot write any old relaxation terms and just wish that the evolution keeps the density operator positive. Technical hint: Study the evolution of the determinant of the density matrix ρ_{ij} , i.e., the evolution of the product of the eigenvalues, when the system starts in a pure state.

Chapter 5

Quantum statistical mechanics

We are now (more than) ready to proceed to quantum statistical mechanics (QSM). The goal ultimately is to predict thermal-equilibrium properties of a large system starting from microscopic arguments. The theory is probabilistic, and a few concepts have to be thoroughly understood in order to understand what is going on in statistical mechanics:

- An *ensemble* is an imaginary collection of copies of a system. An experiment carried out on a system is assumedly carried out on a randomly picked member of the ensemble.
- The states of the systems in the ensemble need not be the same. *Density operator* is a device to characterize the states. In particular, the average of the measurements of an observable A over the randomly picked members of the ensemble will come out as $\text{Tr}(\rho A)$.

The bottom line is, the result that you get when you measure a QSM system is random. The above statements recite the standard operational meaning of probabilities for the case of QSM. QSM predicts *ensemble averages*, nothing more or nothing less. In contrast, you might, for instance measure some value in one given system as a function of time over a long time, and average the result. This is the *time average*. Time average does not have to equal ensemble average, although in an *ergodic* system the two averages are the same.

A traditional exposition of statistical mechanics starts from two postulates, which we recite for QSM as follows:

- P.1 *Random phases.* The density operator is diagonal in an eigenbasis of the system Hamiltonian.
- P.2 *No a priori probabilities.* Every state compatible with the specifications of the ensemble is equally likely.

The system may interact with an environment, in which case one somehow has to ascribe a Hamiltonian H to the system only, as if the interactions with the environment were severed. Because of the random-phase postulate, any orthonormal basis that we will write down from now on is by default an eigenbasis of the system Hamiltonian: $H|n\rangle = E_n|n\rangle$. There is an interesting ambiguity in the above random-phases postulate in that energy eigenvalues might be degenerate, and the eigenbasis is therefore not unique. It turns out, though, that the standard density operators in statistical mechanics are diagonal in *any* eigenbasis of the Hamiltonian. What the specifications as in the second postulate would be, we will find out shortly when we discuss the three standard ensembles in statistical mechanics.

There is one more cornerstone in QSM that merits a mention already at the outset. Namely, in practice one often does *not* calculate the density operator and then proceed to expectation values. Rather, there is a connection between some quantity particular to the QSM ensemble that furnishes a link to thermodynamics. Using this link, in practical calculations one then freely mixes statistical mechanics and thermodynamics. In our development such a link always goes back to entropy, and to the microscopic expression for the entropy. From now on the QSM counterpart of the thermodynamic entropy is the usual *von Neumann entropy*,

$$S = -k \operatorname{Tr}(\rho \ln \rho). \quad (5.1)$$

The reader should realize that while much shorter, concise, and elegant developments are possible, in our discussion of the QSM ensembles the aim is to demonstrate a multitude of angles and approaches. Moreover, we do not really prove most of our assertions, but at best demonstrate that they fit in. It may even remain unclear which is the assumption and which is the result. In physics there are basically two ultimate arbiters for an argument: consistency and empirical evidence. The reader should rest assured that what we are going to say has successfully confronted a hundred years' worth of experimental evidence.

5.1 Ensembles in quantum statistical mechanics

5.1.1 Microcanonical ensemble

A *microcanonical* ensemble is set up to describe an isolated system, one in which both energy and particle number are fixed.

The exact energy of a macroscopic system, of course, is not knowable, so we more specifically state that the energy E belongs to an interval $[\mathcal{E}, \mathcal{E} + \Delta)$. Here the energy resolution Δ is typically assumed to satisfy $\Delta \ll \mathcal{E}$. Now, by the random-phases postulate the density operator satisfies

$$\langle m | \rho | n \rangle = p_n \delta_{mn}, \quad (5.2)$$

with the default assumption that

$$H|n\rangle = E_n|n\rangle. \quad (5.3)$$

By the no-a-priori-probabilities assumption we have

$$p_n = \begin{cases} p, & E_n \in [\mathcal{E}, \mathcal{E} + \Delta) \\ 0, & \text{otherwise,} \end{cases} \quad (5.4)$$

where p is a positive constant. The unit trace of a density operator gives

$$\sum p_n = \sum_{\{n: E_n \in [\mathcal{E}, \mathcal{E} + \Delta)\}} p = p \Gamma(\mathcal{E}) = 1, \quad \Rightarrow p = \frac{1}{\Gamma(\mathcal{E})}, \quad (5.5)$$

where $\Gamma(\mathcal{E})$ is the number of energy eigenstates satisfying $E_n \in [\mathcal{E}, \mathcal{E} + \Delta)$. Introducing the characteristic function for the set of points S

$$\chi_S(x) = \begin{cases} 1, & x \in S \\ 0, & x \notin S, \end{cases} \quad (5.6)$$

we have an explicit expression for the occupation probabilities of the states, and indeed for the density operator itself,

$$p_n = \frac{1}{\Gamma(\mathcal{E})} \chi_{[\mathcal{E}, \mathcal{E} + \Delta)}(E_n); \quad \rho = \frac{\chi_{[\mathcal{E}, \mathcal{E} + \Delta)}(H)}{\text{Tr}[\chi_{[\mathcal{E}, \mathcal{E} + \Delta)}(H)]}. \quad (5.7)$$

We have found the density operator, and in principle the entire statistical mechanics problem is solved. In practice, however, one would usually proceed differently.

Connection to thermodynamics

The von Neumann entropy of the microcanonical ensemble equals

$$\begin{aligned} S &= -k \text{Tr}(\rho \ln \rho) = -k \sum_n p_n \ln p_n = -k \sum_n p_n \ln p \\ &= -k \ln p \sum_n p_n = -k \ln p = k \ln \Gamma(\mathcal{E}) = k \ln W, \end{aligned} \quad (5.8)$$

k times the logarithm of the number of accessible states, just like on the grave stone of Ludwig Boltzmann. The standard procedure in the microcanonical ensemble is to calculate the von Neumann entropy S as a function of energy \mathcal{E} , then interpret S as the thermodynamic entropy and \mathcal{E} as the thermodynamic internal energy, and continue with thermodynamics from there on.

For instance, the inverse temperature may be obtained as

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_V \simeq \frac{dS}{d\mathcal{E}}. \quad (5.9)$$

The second step is the connection between thermodynamics and statistical mechanics. Notice that the counterpart in thermodynamics of the statistical-mechanics energy derivative takes place at constant volume. The idea is that,

while the energy of the system is varied, the energy eigenstates themselves remain unchanged. As the eigenstates depend on the volume (quantization volume), the assumption of unchanged energy eigenstates strongly suggests that the process takes place at constant volume.

There is another subtlety in the argument, namely, that entropy seemingly depends on the energy resolution Δ . In the thermodynamic limit that cannot be, and is not, the case. To see this, let us also introduce the number of energy eigenstates under the energy shell Σ and the density of states Ω as

$$\Sigma(\mathcal{E}) = \sum_{\{n: E_n < \mathcal{E}\}} 1, \quad \Omega(\mathcal{E}) = \frac{\partial \Sigma}{\partial \mathcal{E}}. \quad (5.10)$$

In the limit with $\Delta \ll \mathcal{E}$, the number of accessible states in the microcanonical ensemble is then

$$\Gamma(\mathcal{E}) = \Omega(\mathcal{E})\Delta. \quad (5.11)$$

Let us now suppose that we can scale the size of the system, keeping the intensive quantities such as energy per particle $\epsilon = \mathcal{E}/N$ constant. In a statistical-mechanics system one can rarely resolve a single-particle energy and an energy resolution larger than the energy itself is not of much use, so that we have reasonable bounds for the energy resolution Δ ,

$$\epsilon \leq \Delta \leq N\epsilon. \quad (5.12)$$

Likewise we have an approximate bound for the number of accessible state, and for the logarithm thereof, as

$$\epsilon\Omega \leq \Gamma \leq N\epsilon\Omega; \quad \ln \epsilon + \ln \Omega \leq \ln \Gamma \leq \ln N\epsilon + \ln \Omega. \quad (5.13)$$

Empirically, the energy density $\Omega(N\epsilon)$ scales with the particle number in such a way that $\ln \Omega \sim N$, and we have the estimate

$$\frac{\ln \epsilon}{N} + \frac{\ln \Omega}{N} \leq \frac{\ln \Gamma}{N} \leq \frac{\ln N\epsilon}{N} + \frac{\ln \Omega}{N}, \quad (5.14)$$

which in the limit $N \rightarrow \infty$ turns into

$$\frac{\ln \Omega}{N} = \frac{\ln \Gamma}{N}. \quad (5.15)$$

Along these lines, we see that to the leading order in the particle number $N \gg 1$ we have

$$\ln \Sigma(\mathcal{E}) = \ln \Gamma(\mathcal{E}) = \ln \Omega(\mathcal{E}). \quad (5.16)$$

The definition of entropy is independent of energy resolution for any reasonable value of Δ . It should also be noted that, because $\ln \Gamma$, and therefore also $\ln \Omega$ and $\ln \Sigma$ scale with the particle number as $\sim N$, the von Neumann entropy is an extensive quantity.

5.1.2 Canonical ensemble

A *canonical ensemble* is set up to model a system that exchanges energy, but not particles, with the environment.

Let us take the system $S \equiv 1$ and the environment $E \equiv 2$, and consider the joint system $S + E$. The Hamiltonian in general is

$$H = H_1 + H_2 + H_{12}, \quad (5.17)$$

where H_1 (H_2) is the Hamiltonian of the system (environment), and H_{12} is the interaction between the system and the environment. But now, in the thermodynamic limit when the systems are made larger keeping the intensive quantities constant, the Hamiltonians H_1 and H_2 scale with the volume V , while the interaction Hamiltonian H_{12} typically scales with the area of the surface between the system and the environment, $V^{2/3}$. In the thermodynamic limit the interaction Hamiltonian may therefore be neglected, and the energies of the system and the environment simply add up:

$$H = H_1 + H_2; \quad E = E_1 + E_2. \quad (5.18)$$

The system-environment interaction is needed to establish thermal equilibrium, but we know that the larger the systems, the longer it takes to equilibrate. Ignoring H_{12} , in a sense, is a matter of time scale, and is valid at least over a time scale short compared to the time it takes the system and environment to equilibrate.

Moving forward, let us treat the joint system $S + E$ as microcanonical with the energy resolution Δ . Energy may be exchanged between S and E so that while $E = E_1 + E_2$ is fixed, E_1 and E_2 individually are not. The particle numbers N_1 and N_2 , however, are assumed fixed. Moreover, the system is taken to be much smaller than the environment, so that $E_1 \ll E, E_2$.

Let us now ask, what is the ratio of the probabilities that the system is in a state with energy E_1 and energy E'_1 ? Fixing the system energy at E_1 still leaves a choice of the energy of the environment over a range of Δ around $E_2 = E - E_1$, so that the number of states of $S + E$ such that the system is in a state with energy E_1 equals $\Gamma_2(E - E_1)$; and likewise for E'_1 . By the postulate of no a priori probabilities the probability for the system to have any given energy is proportional to the number of states of $S + E$ compatible with that energy, so that for the two probabilities we have the ratio

$$\frac{P(E_1)}{P(E'_1)} = \frac{\Gamma_2(E - E_1)}{\Gamma_2(E - E'_1)}. \quad (5.19)$$

But if the number of states in the environment with a given range of energy is Γ_2 , we can express the same formally in terms of the entropy of the environment using $S_2 = k \ln \Gamma_2$, or $\Gamma_2 = e^{S_2/k}$. So, we have

$$\frac{P(E_1)}{P(E'_1)} = \exp \left\{ \frac{1}{k} [S_2(E - E_1) - S_2(E - E'_1)] \right\}$$

$$\begin{aligned}
&= \exp \left\{ \frac{1}{k} \left[-(E_1 - E'_1) \frac{\partial S_2}{\partial E} \right] \right\} = \exp \left\{ \frac{1}{k} \left[-\frac{E_1 - E'_1}{T} \right] \right\} \\
&= \frac{e^{-E_1/kT}}{e^{-E'_1/kT}} = \frac{e^{-\beta E_1}}{e^{-\beta E'_1}}.
\end{aligned} \tag{5.20}$$

In the second step we have expanded in E_1 and E'_1 . In the third step we have identified

$$\frac{\partial S_2}{\partial E} = \frac{1}{T_2} \equiv \frac{1}{T}.$$

This firstly defines the temperature of the environment T_2 , but by the idea that in equilibrium the large “heat bath” of the environment determines the temperature of the system as well, also the joint system-environment temperature T .

Let us finally forget about the environment, and consider the (reduced) density operator of the system only. Obviously we have from the first postulate

$$\langle n | \rho | m \rangle = \delta_{nm} p_n, \tag{5.21}$$

with the result that we deduced from the second postulate

$$p_n = C e^{-\beta E_n}; \tag{5.22}$$

C is a so far undetermined constant, but in fact we have

$$\text{Tr } \rho = \sum_n p_n = C \sum_n e^{-\beta E_n} = 1 \quad \Rightarrow \quad C = Z^{-1}; \quad Z = \sum_n e^{-\beta E_n}. \tag{5.23}$$

Putting all of this together, we have the density operator for the canonical ensemble

$$\rho = \frac{1}{Z} e^{-\beta H}, \tag{5.24}$$

where the constant

$$Z = \sum_n e^{-\beta E_n} = \text{Tr } e^{-\beta H} \tag{5.25}$$

is called *canonical partition function*.

Once more we have the density operator, and may in principle proceed to expectation values. In practice, however, one again often resorts to a connection to thermodynamics.

Connection to thermodynamics

Given the canonical density operator, the von Neumann entropy is given by

$$\begin{aligned}
S &= -k \text{Tr}(\rho \ln \rho) = -k \sum_n p_n \ln p_n = k \sum_n p_n (\ln Z + \beta E_n) \\
&= k \ln Z + \frac{1}{T} \sum_n p_n E_n = k \ln Z + \frac{\langle H \rangle}{T},
\end{aligned} \tag{5.26}$$

or

$$-kT \ln Z = \langle H \rangle - TS. \quad (5.27)$$

The thermodynamics counterpart of the expectation value of the Hamiltonian obviously is the internal energy U , so that the right-hand side of (5.27) has the appearance of the thermodynamical definition of Helmholtz free energy, $F = U - TS$. We thus establish the rule that connect thermodynamics and the canonical partition function:

$$F = -kT \ln Z. \quad (5.28)$$

From now on, again, we may mix thermodynamics and statistical mechanics freely as it suits our purposes.

Examples

Statistical mechanics and thermodynamics of a single spin. Let us consider a single spin with the spin-up and spin-down states \uparrow and \downarrow and the corresponding energies $\epsilon_{\uparrow} = -\epsilon/2$ and $\epsilon_{\downarrow} = \epsilon/2$. This may not seem like a macroscopic system, but, because of the coupling to a big environment, statistical mechanics applies nevertheless. The partition function is

$$Z = e^{-\beta\epsilon_{\uparrow}} + e^{-\beta\epsilon_{\downarrow}} = e^{\frac{\beta\epsilon}{2}} + e^{-\frac{\beta\epsilon}{2}} = 2 \cosh \frac{\beta\epsilon}{2}. \quad (5.29)$$

The probabilities for the spin-up and spin-down states are

$$p_{\uparrow,\downarrow} = \frac{e^{-\beta\epsilon_{\uparrow,\downarrow}}}{Z} = \frac{e^{\pm\frac{\beta\epsilon}{2}}}{e^{\frac{\beta\epsilon}{2}} + e^{-\frac{\beta\epsilon}{2}}}, \quad (5.30)$$

and the expectation value of the energy therefore is

$$\langle H \rangle = U = \epsilon_{\uparrow} p_{\uparrow} + \epsilon_{\downarrow} p_{\downarrow} = -\frac{\epsilon}{2} \frac{e^{\frac{\beta\epsilon}{2}} - e^{-\frac{\beta\epsilon}{2}}}{e^{\frac{\beta\epsilon}{2}} + e^{-\frac{\beta\epsilon}{2}}} = -\frac{\epsilon}{2} \tanh \frac{\beta\epsilon}{2}. \quad (5.31)$$

On the other hand, one may quickly switch to thermodynamics. The free energy is

$$F = -kT \ln Z = -kT \ln \left(2 \cosh \frac{\beta\epsilon}{2} \right); \quad (5.32)$$

let us call it Helmholtz free energy, although in this simple problem there are no analogs of volume, pressure, particle number, and so on. Entropy, however, may still be obtained as

$$S = -\frac{\partial F}{\partial T}. \quad (5.33)$$

Internal energy is therefore

$$\begin{aligned} U &= F + TS = F - T \frac{\partial F}{\partial T} = -kT \ln Z - T \frac{\partial}{\partial T} (-kT \ln Z) = kT^2 \frac{\partial}{\partial T} \ln Z \\ &= -kT \beta \frac{\partial}{\partial \beta} \ln Z = -\frac{\partial}{\partial \beta} \ln Z. \end{aligned} \quad (5.34)$$

This result, obtained using $\beta = (kT)^{-1}$ to transform the derivatives, is still completely general. In our special case it gives

$$U = -\frac{\epsilon}{2} \tanh \frac{\beta\epsilon}{2} \quad (5.35)$$

as before. The heat capacity can also be calculated in several ways, giving the same result:

$$C = \frac{\partial U}{\partial T} = T \frac{\partial S}{\partial T} = k \frac{\left(\frac{\epsilon}{2kT}\right)^2}{\cosh^2\left(\frac{\epsilon}{2kT}\right)}. \quad (5.36)$$

This tends to zero both in the limits $T \rightarrow 0$ and $T \rightarrow \infty$, and obviously peaks at a value $\sim k$ when $kT \sim \epsilon$.

Energy fluctuations in the canonical ensemble. In a calculation that might not come across as intuitively obvious at first sight but which is utterly standard in statistical mechanics, we proceed as follows:

$$\begin{aligned} \frac{\partial}{\partial \beta} \langle H \rangle &= \frac{\partial}{\partial \beta} \text{Tr}(\rho H) = \frac{\partial}{\partial \beta} \left[\frac{1}{Z(\beta)} \text{Tr}(e^{-\beta H} H) \right] \\ &= -\frac{1}{Z^2} \frac{\partial Z}{\partial \beta} \text{Tr}(e^{-\beta H} H) + \frac{1}{Z} \frac{\partial}{\partial \beta} \text{Tr}(e^{-\beta H} H) \\ &= -\frac{1}{Z^2} \left[\frac{\partial}{\partial \beta} \text{Tr}(e^{-\beta H}) \right] \text{Tr}(e^{-\beta H} H) + \frac{1}{Z} \frac{\partial}{\partial \beta} \text{Tr}(e^{-\beta H} H) \\ &= \left[\frac{1}{Z} \text{Tr}(e^{-\beta H} H) \right]^2 - \frac{1}{Z} \text{Tr}(e^{-\beta H} H^2) = \langle H \rangle^2 - \langle H^2 \rangle \\ &= -(\Delta E)^2. \end{aligned} \quad (5.37)$$

The fourth line follows because differential calculus works for operators just like for numbers, as long as the order of noncommuting objects is not unduly changed. So, we have an estimate for the fluctuations of energy in the canonical ensemble. On the other hand, by comparing with thermodynamics we also find

$$\frac{\partial}{\partial \beta} \langle H \rangle \cong -kT^2 \frac{\partial}{\partial T} U = -kT^2 \left(\frac{\partial U}{\partial T} \right)_V = kT^2 C_V. \quad (5.38)$$

The thermodynamical derivative is at constant volume, again, since we correspondingly regard the energy spectrum of the Hamiltonian as fixed. Comparing the two results for $\partial \langle H \rangle / \partial \beta$, we have

$$\Delta E = \sqrt{kT^2 C_V}. \quad (5.39)$$

Standard thermodynamics says nothing about fluctuations, so our reasoning here goes beyond thermodynamics. Second, since the heat capacity as we have defined it is an extensive quantity, in the thermodynamical limit $N \rightarrow \infty$ we have

$$\frac{\Delta E}{E} \propto \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \rightarrow 0. \quad (5.40)$$

The difference between microcanonical and canonical ensembles is that in the latter energy is allowed to fluctuate; but in practice the fluctuations are negligible in the thermodynamical limit anyway. Therefore, in the thermodynamical limit, the canonical and the microcanonical ensemble are *equivalent*. Which one to use is purely a matter of convenience.

5.1.3 Grand canonical ensemble

A *grand canonical ensemble* is set up to model a system that exchanges energy and particles with the environment.

Our approach to the grand canonical ensemble is much different from, and complementary to, what we have said before. Thus, given the system S that we are considering, let us suppose that we can define a Hamiltonian and a particle number operators \hat{H} and \hat{N} . The Hamiltonian is obviously what would result if we were to ignore the interactions with the environment, a usual assumption when we are discussing the thermal-equilibrium state. How to construct and handle the particle number operator is a standard issue in many-body quantum mechanics, but is beyond the scope of our discussion. Suffice it to say that such an \hat{N} can be neatly defined.

We assume that, in the absence of interactions and particle exchange with the environment, the particle number is conserved: $[\hat{H}, \hat{N}] = 0$. This means that it is principle possible to construct for the system an orthonormal basis of simultaneous eigenstates of particle number and energy, $\{|Nn\rangle\}$:

$$\hat{N}|Nn\rangle = N|Nn\rangle, \quad \hat{H}|Nn\rangle = E_n(N)|Nn\rangle. \quad (5.41)$$

Given these preliminaries, we now simply state the density operator of the grand canonical ensemble:

$$\rho = \frac{1}{\mathcal{Z}} e^{-\beta(\hat{H} - \mu\hat{N})}, \quad (5.42)$$

where

$$\mathcal{Z} = \text{Tr} e^{-\beta(\hat{H} - \mu\hat{N})} \quad (5.43)$$

is called the *grand partition function*, and μ is the chemical potential set by the (large) environment that acts as a source and sink of both energy and particles.

The grand partition function may be developed as

$$\begin{aligned} \mathcal{Z} &= \text{Tr} e^{-\beta(\hat{H} - \mu\hat{N})} = \sum_{N,n} e^{-\beta(E_n(N) - \mu N)} = \sum_N e^{\beta\mu N} \left(\sum_n e^{-\beta E_n(N)} \right) \\ &= \sum_N z^N Z_N, \end{aligned} \quad (5.44)$$

where $z = e^{\beta\mu}$ is called *fugacity*, and Z_N is the N -particle *canonical* partition function.

Incidentally, here the question arises how to define the canonical partition function for *zero* particles, Z_0 . We define $Z_0 \equiv 1$. This convention will assure

that in the connection to thermodynamics we will discuss shortly the pressure in the situation with no particles present turns out to be zero.

Further insights may be gleaned by considering the expectation value of an arbitrary function of particle number operator $\langle f(\hat{N}) \rangle$. The result is

$$\langle f(\hat{N}) \rangle = \frac{1}{\mathcal{Z}} \text{Tr} \left(e^{-\beta(\hat{H} - \mu \hat{N})} f(\hat{N}) \right) = \sum_N \frac{z^N Z_N}{\mathcal{Z}} f(N). \quad (5.45)$$

Now, the ratio inside the sum satisfies

$$\frac{z^N Z_N}{\mathcal{Z}} \geq 0, \quad \sum_N \frac{z^N Z_N}{\mathcal{Z}} = \frac{\mathcal{Z}}{\mathcal{Z}} = 1. \quad (5.46)$$

We have

$$p_N = \frac{z^N Z_N}{\mathcal{Z}}; \quad \langle f(\hat{N}) \rangle = \sum_N p_N f(N), \quad (5.47)$$

where p_N may obviously be interpreted as the *probability that the system has N particles*. In fact, the projector corresponding to the question of whether the particle number is N obviously is

$$P_N = \sum_n |Nn\rangle \langle Nn|, \quad (5.48)$$

so that

$$P(N) = \text{Tr}(\rho P_N) = \frac{z^N Z_N}{\mathcal{Z}} = p_N. \quad (5.49)$$

Connection to thermodynamics

Experience has shown that particle number fluctuations in the grand canonical ensemble are usually small, so that the probabilities p_N are peaked around some characteristic particle number \bar{N} and terms with $N \simeq \bar{N}$ also dominate in the grand partition functions. We therefore have

$$\begin{aligned} -kT \ln \mathcal{Z} &= -kT \ln \left(\sum_N z^N Z_N \right) \simeq -kT \ln \left(z^{\bar{N}} Z_{\bar{N}} \right) = -kT \bar{N} \beta \mu - kT \ln Z_{\bar{N}} \\ &\cong F(\bar{N}, V, T) - \bar{N} \mu = \Omega, \end{aligned} \quad (5.50)$$

where we have used the previously established connection between canonical partition function and Helmholtz free energy. The key connection between thermodynamics and grand canonical ensemble goes via the grand potential,

$$\Omega = -kT \ln \mathcal{Z}. \quad (5.51)$$

From now on, it is again legitimate to mix freely statistical mechanics and thermodynamics. In particular, it is often useful to recall the connection between grand potential and pressure, $\Omega = -pV$.

Much along the lines of our discussion of energy fluctuations in the canonical ensemble, we could also estimate the fluctuations of particle number in the grand canonical ensemble. Consistently with what we have already said, the relative fluctuations of particle number scale in the thermodynamical limit as $1/\sqrt{N}$, and vanish with $N \rightarrow \infty$ (except in a pathological special case). Since the difference between canonical and grand canonical ensembles is particle number fluctuations, in the thermodynamical limit the ensembles are equivalent. Which one to use is once more a matter of convenience.

Why the particular density operator?

We have in effect postulated the density operator (5.42) for the grand canonical ensemble. We conclude by sketching a post facto justification that also illustrates another possible approach to statistical mechanics in general.

Thus, we take as the starting point the von Neumann entropy

$$S = k \operatorname{Tr}(\rho \ln \rho), \quad (5.52)$$

and require that the thermal-equilibrium density operator maximizes the entropy under the constraints of the ensemble.

For the grand canonical ensemble the proper conditions read

$$\begin{aligned} \operatorname{Tr} \rho &= 1, \\ \langle \hat{H} \rangle = \operatorname{Tr}(\rho \hat{H}) &= E, \\ \langle \hat{N} \rangle = \operatorname{Tr}(\rho \hat{N}) &= N, \end{aligned} \quad (5.53)$$

where the first one is the trace condition for the density operator, and we insist on fixed *expectation values* of energy and particle number. By varying the density operator, taking into account (4.31), we find

$$\begin{aligned} -\frac{1}{k} \delta S &= \operatorname{Tr}[\delta \rho (\ln \rho + 1)] = 0, \\ \delta \operatorname{Tr} \rho &= \operatorname{Tr}[\delta \rho] = 0, \\ \delta \operatorname{Tr}(\rho \hat{H}) &= \operatorname{Tr}[\delta \rho \hat{H}] = 0, \\ \delta \operatorname{Tr}(\rho \hat{N}) &= \operatorname{Tr}[\delta \rho \hat{N}] = 0. \end{aligned} \quad (5.54)$$

The first one of these expresses the observation that the maximum of entropy is insensitive to small variations of the density operator. The variations of the density operator are not free, but must satisfy three conditions. The standard trick to handle this kind of situations is to introduce Lagrange's undetermined multipliers, call them $k_0 - 1$, k_1 and k_2 , to the constraint equations and add everything, so that we have

$$\operatorname{Tr}[\delta \rho (\ln \rho + k_0 + k_1 \hat{H} + k_2 \hat{N})] = 0. \quad (5.55)$$

Given the undetermined multipliers, we may now treat $\delta \rho$ as a free variation, which leads to

$$\ln \rho + k_0 + k_1 \hat{H} + k_2 \hat{N} = 0; \quad \rho = e^{-k_0 - k_1 \hat{H} - k_2 \hat{N}}. \quad (5.56)$$

With a suitable identification of the undetermined multipliers, we have indeed found the density operator for the grand canonical ensemble, (5.42).

5.2 Exercises

5.1. Find the energy eigenstates and eigenvalues for a single free massive particle under the periodic boundary conditions in a quantization volume $V = L^3$. Periodic boundary conditions mean that the wave function is periodic in each Cartesian coordinate with the period L .

5.2. Consider a closed system (all macroscopic systems may be regarded as closed over sufficiently short times), so that the time evolution is given by the Liouville-von Neumann equation. Show that the postulate of random phases is equivalent to the statement that the expectation value of any observable that does not depend explicitly on time (in the Schrödinger picture) is independent of time. The postulate of random phases simply states that all time evolution in the system has ceased; i.e., that the system is in thermodynamic equilibrium!

5.3. In the limit of high energy or low atom density, the number of energy eigenstates with energy $\leq E$ for an ideal gas with $N \gg 1$ atoms in a volume V is

$$\Sigma(E) = \frac{(2\pi m E)^{\frac{3N}{2}} V^N}{(2\pi\hbar)^{3N} N! (3N/2)!};$$

if $(3N/2)!$ gives you pause, assume that N is even.

(a) Demonstrate that in the thermodynamic limit $\ln \Sigma \propto N$ in fact holds true.

(b) What is the entropy S as a function of temperature T and atom density $n = N/V$? The entropy is extensive, and the additive constant that cannot be deduced from the thermodynamics of an ideal gas also comes out right if you do the math correctly.

5.4. Consider $N \gg 1$ noninteracting three-dimensional harmonic oscillators with (angular) frequency ω . It turns out that in the limit when the energy per oscillator is much larger than a harmonic-oscillator quantum $\hbar\omega$, the number of many-particle states with a total energy $\leq E$ is

$$\Sigma(E) = \frac{1}{N! (3N)!} \left(\frac{E}{\hbar\omega} \right)^{3N}.$$

Calculate the heat capacity of the system as a function of temperature.

5.5. Show that for a joint system 1+2 made of two completely independent subsystems 1 and 2 the canonical partition function factorizes:

$$Z_{1+2} = Z_1 Z_2.$$

The same, of course, applies if an *arbitrary* number of independent systems are combined.

5.6. In the basic model for a diatomic molecule the good quantum numbers may be picked to be the center-of-mass wave vector \mathbf{k} for the translational degree of freedom, the total angular momentum $J = 0, 1, \dots$ and the z component of the angular momentum $m = -J, J+1, \dots, J$ for the rotation of the molecule, and the vibrational quantum number $n = 0, 1, \dots$ for the vibrations of the distance between the atoms, so that the energy of the molecule for the given quantum numbers is

$$E(\mathbf{k}, J, m, n) = \frac{\hbar^2 \mathbf{k}^2}{2m} + BJ(J+1) + \hbar\omega n,$$

where $B > 0$ is the rotational constant and $\omega > 0$ is the (angular) frequency of the vibration mode.

- (a) Find the high- and low-temperature limits of the heat capacities in the center-of-mass motion, rotation, and vibration of a single molecule.
- (b) It is known that at room temperature the heat capacity per molecule of the atmospheric gases such as N_2 and O_2 is $\frac{5}{2}k$. What does this tell you about the rotational and vibrational motion?

MATHEMATICAL REMARKS. As discussed in more detail later, the sum over the wave vectors \mathbf{k} may be calculated in the *continuum approximation* as

$$\sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3k.$$

In the high-temperature limit the summation index J (and n) may also be regarded as continuous.

5.7. According to the canonical ensemble, the probability that an individual state is occupied is proportional to the Boltzmann factor $e^{-\beta E}$, but there is also the density of states $\Omega(E)$ to consider. Therefore, the probability density for energy $f(E)$, such that the probability for the system to be in an energy interval $[E, E + dE]$ equals $dP = f(E) dE$, must be of the form $f(E) = Ce^{-\beta E}\Omega(E) = Ce^{-\beta E + \ln \Omega(E)}$, where C is a normalization constant. Now, it is known empirically that in the thermodynamic limit we have $E = N\epsilon$ and $\ln \Omega(E) = N\sigma(\epsilon)$ where ϵ and σ are intensive, so that $f(E) = Ce^{N[-\beta\epsilon + \sigma(\epsilon)]}$. Ad hoc mixing of SM and TD is permitted here.

(a) By studying the expression of $f(E)$ in the neighborhood of its maximum at the most probable energy $\bar{E} = N\bar{\epsilon}$, in the limit of large N , argue that the probability of a fluctuation from equilibrium with the *second-order* change of entropy $\Delta^{(2)}S \propto (E - \bar{E})^2$ is proportional to $e^{\Delta^{(2)}S/k}$. Of course, the most probable energy corresponds to the thermodynamical equilibrium.

(b) Deduce from the above that the fluctuations of energy $(\Delta E)^2 = \langle (E - \bar{E})^2 \rangle$ are $(\Delta E)^2 = kT^2 C_V$.

5.8. In the limit of high temperature or low density, the canonical partition function for an ideal gas of N atoms with mass m in a volume V is known to be

$$Z_N = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N, \quad \lambda = \left(\frac{2\pi\hbar^2}{mkT} \right)^{1/2}.$$

- (a) Calculate the grand partition function \mathcal{Z} .
 (b) Using the known connections of the grand-canonical ensemble to thermodynamics [in terms of $\Omega = F - N\mu$], find the equation of state $pV = NkT$.

5.9. Given the grand partition function \mathcal{Z} , show that the expectation value of particle number and the root-mean-square fluctuations of the particle number satisfy

$$N = z \frac{\partial}{\partial z} \ln \mathcal{Z}, \quad (\Delta N)^2 = z \frac{\partial}{\partial z} z \frac{\partial}{\partial z} \ln \mathcal{Z}.$$

How do the relative fluctuations, $\Delta N/N$, scale in the thermodynamic limit?

5.10. Take a system of n adsorption sites, each of which may bind a molecule; putting a molecule to a site lowers the energy by ϵ compared to the energy of the molecule floating around in the environment.

- (a) Given an environment where the molecules have temperature T and chemical potential μ , the grand partition function for the molecules of this lattice gas is

$$\mathcal{Z} = [1 + e^{\beta(\epsilon+\mu)}]^n.$$

Why?

- (b) Show that the probability that a site is occupied by a molecule is $p = [1 + e^{-\beta(\epsilon+\mu)}]^{-1}$.

Chapter 6

Ideal quantum gases

6.1 State counting and statistics

An ideal gas is characterized by the property that the Hamiltonian is a straight sum of the Hamiltonians of the individual particles, with no interaction terms between the particles:

$$H = \sum_n h_n. \quad (6.1)$$

Here h_n is the *one-particle Hamiltonian* for particle $n = 1, \dots, N$.

For example, for free massive particles without any internal structure, we could write

$$h_n = \frac{\mathbf{p}_n^2}{2m} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial \mathbf{r}_n^2}. \quad (6.2)$$

The corresponding one-particle energy eigenstates and energies are denoted by $|i\rangle_n$ and ϵ_i :

$$h_n |i\rangle_n = \epsilon_i |i\rangle_n. \quad (6.3)$$

Quantizing in a three-dimensional box of size L and volume V , with periodic boundary condition, the index i and the energy ϵ_i could stand for

$$i \leftrightarrow (n_1, n_2, n_3), \quad n_1, n_2, n_3 = 0, \pm 1, \pm 2, \dots; \quad \epsilon_i = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L} \right)^2 (n_1^2 + n_2^2 + n_3^2). \quad (6.4)$$

One *might* think that the energy eigenstates of the ideal gas are simply the tensor products

$$|\psi\rangle = |i_1\rangle_1 |i_2\rangle_2 \dots |i_N\rangle_N. \quad (6.5)$$

In fact, one finds that

$$\begin{aligned} H|\psi\rangle &= h_1|\psi\rangle + h_2|\psi\rangle + \dots = \epsilon_{i_1}|\psi\rangle + \epsilon_{i_2}|\psi\rangle + \dots \\ &= \left(\sum_n \epsilon_{i_n} \right) |\psi\rangle, \end{aligned} \quad (6.6)$$

so that the energy of the state $|\psi\rangle$ would simply be the sum of the corresponding one-particle energies.

However, for *indistinguishable particles*, states of the form (6.5) are in general not realized in Nature. Basically, you cannot say which particle n is in which eigenstate i_n . The complete characterization of a many-body energy eigenstate (of an ideal gas) is to tell *how many* particles n_i are in each one-particle energy eigenstate i .

Besides, Nature also imposes constraints on the available *occupation numbers* n_i . All “particles” fall into one of two classes:

- fermions: $n = 0, 1$
- bosons: $n = 0, 1, 2, 3 \dots$

corresponding to *Fermi-Dirac* and *Bose-Einstein statistics*. In a Fermi gas there may only be one particle in any given one-particle state, which is commonly referred to as Pauli exclusion principle; in a Bose gas there are no restrictions on the particle number. Cases of *parastatistics*, say, that the occupation numbers could be precisely 0, 1 and 2, do not occur, although there is, of course, nothing to prevent one from studying a parastatistics formally. The statistics is a property of the particle species. In fact, by the *spin-statistics theorem*, particles with a half-integer angular momentum are fermions, and particles with an integer angular momentum are bosons. The particles do not have to be elementary. For instance, as long as one does not try to break apart or combine ^4He atoms, empirically they behave as a bosons. Different internal states of a particle (elementary or composite) may be regarded as different species of indistinguishable particles.

The question of how one should write down the fermion and boson states with given one-particle occupation numbers $\{n_i\}$ belongs to the realm of many-body physics, and is beyond the scope of our discussion. Fortunately, in order to do statistical mechanics for ideal gases, one does not have to be able to write down the energy eigenstates. It suffices to know that a state with the occupation numbers $\{n_i\}$ is the eigenstate of both particle number and total energy, the eigenvalues being

$$N = \sum_i n_i, \quad E = \sum_i n_i \epsilon_i. \quad (6.7)$$

6.2 Grand partition function

The grand canonical ensemble conveniently solves the problem of quantum statistical ideal gases. To wit, we develop as follows,

$$\begin{aligned} \mathcal{Z} &= \sum_N z^N Z_N = \sum_N z^N \sum_{\{\{n_i\}: \sum_i n_i = N\}} e^{-\beta E(\{n_i\})} \\ &= \sum_N \sum_{\{\{n_i\}: \sum_i n_i = N\}} e^{-\beta[E(\{n_i\}) - \mu N]} = \sum_N \sum_{\{\{n_i\}: \sum_i n_i = N\}} e^{-\sum_i \beta n_i (\epsilon_i - \mu)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{\{n_i\}} e^{-\sum_i \beta n_i (\epsilon_i - \mu)} = \sum_{\{n_i\}} \left(\prod_i e^{-\beta n_i (\epsilon_i - \mu)} \right) \\
&= \prod_i \left(\sum_{n_i} e^{-\beta n_i (\epsilon_i - \mu)} \right). \tag{6.8}
\end{aligned}$$

The step to the third line means that if you first sum over the occupation numbers n_i under the constraint that $\sum_i n_i = N$ and then sum over all values of N , you have in effect carried out an unconstrained sum over all occupation numbers n_i . A moment's reflection will show that the step to the fourth line is an analog of the simple identity $a_1 b_1 + a_1 b_2 + a_2 b_1 + a_2 b_2 = (a_1 + a_2)(b_1 + b_2)$.

For fermions the possible occupation number are $n_i = 0, 1$, so that we have

$$\begin{aligned}
\mathcal{Z} &= \prod_i \left(1 + e^{-\beta(\epsilon_i - \mu)} \right) = \prod_i (1 + z e^{-\beta \epsilon_i}); \\
\Omega &= -kT \ln \mathcal{Z} = -kT \sum_i \ln(1 + z e^{-\beta \epsilon_i}). \tag{6.9}
\end{aligned}$$

For bosons the sum over $n_i = 0, 1, \dots$ is a geometric series, and the corresponding results are

$$\begin{aligned}
\mathcal{Z} &= \prod_i \left(1 - e^{-\beta(\epsilon_i - \mu)} \right)^{-1} = \prod_i \frac{1}{1 - z e^{-\beta \epsilon_i}}; \\
\Omega &= -kT \ln \mathcal{Z} = kT \sum_i \ln(1 - z e^{-\beta \epsilon_i}). \tag{6.10}
\end{aligned}$$

As one more item in the development, let us calculate the total number of atoms for, say, fermions; in the sense of QSM, the expectation value thereof:

$$\langle \hat{N} \rangle = N = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{V, T} = kT \frac{\partial}{\partial \mu} \ln \mathcal{Z} = z \frac{\partial}{\partial z} \ln \mathcal{Z} \equiv \sum_i \langle n_i \rangle. \tag{6.11}$$

First notice the liberal mixing of statistical mechanics and thermodynamics. Second, the substitution $\mu \rightarrow z$ often helps with derivatives when dealing with the grand partition function. Third, the expectation value of the number of particles comes out as a sum over the one-particle states. It is therefore natural, and in fact correct in the sense of many-body physics, to identify the quantity $\langle n_i \rangle$ in the last sum as the expectation value of the number of particles in the state i . We have for *fermions*

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1}. \tag{6.12}$$

A similar calculation gives for bosons

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1}, \tag{6.13}$$

the difference being a change in sign.

We have found the thermal occupation numbers as quoted in elementary textbooks. From the partition functions and occupation numbers, one can in principle calculate anything one ever needs to know about ideal Fermi and Bose gases. In the process the observations that

$$N = \langle \hat{N} \rangle = \sum_i \langle n_i \rangle, \quad E = \langle \hat{H} \rangle = \sum_i \langle n_i \rangle \epsilon_i \quad (6.14)$$

frequently come in handy.

6.3 Ideal Fermi gas

Numerous variation of the theme of the ideal Fermi gas are possible. For instance, nowadays it is possible to carry out experiments in situation in which a nearly ideal Fermi gas (or, for that matter, Bose gas) is effectively one- or two-dimensional; the one-particle energy as a function of momentum, call it dispersion relation, is different for massive and massless particles; and so on. We will take up a few examples with the understanding that a student should be able to analyze further variants of the problem by him/herself.

We discuss massive particles, m , with the angular momentum or “spin” s . The one-body energy eigenstates are obtained by quantizing in a cubic box of size L with periodic boundary conditions, so that the energy eigenstates are completely characterized by the value of momentum \mathbf{p} or wave vector $\mathbf{k} = \mathbf{p}/\hbar$. Specifically,

$$\mathbf{k} = \frac{2\pi}{L}(q_1\hat{\mathbf{e}}_1 + q_2\hat{\mathbf{e}}_2 + q_3\hat{\mathbf{e}}_3), \quad q_1, q_2, q_3 = 0, \pm 1, \dots; \quad \epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m}. \quad (6.15)$$

Since fermions necessarily have a (half-integer) nonzero spin, we usually have to contend with the quantum number of the projection of the angular momentum to a given quantization axis,

$$m_s = -s, -s+1, \dots, s-1, s \quad (6.16)$$

as well. The complete one-particle state is therefore specified by a wave vector and a magnetic quantum number,

$$|i\rangle \equiv |\mathbf{k}, m_s\rangle. \quad (6.17)$$

For the present purposes we assume that all the

$$g = 2s + 1 \quad (6.18)$$

spin states are degenerate. More often than not this is not the case in the presence of a magnetic field, in a polarized gas it may happen that only one m_s component is present, and these are just a few of the possible alternatives.

The equation for the particle number reads

$$N = \sum_i \langle n_i \rangle = \sum_{\mathbf{k}, m_s} \langle n_i(\epsilon_{\mathbf{k}}) \rangle = g \sum_{\mathbf{k}} \langle n_i(\epsilon_{\mathbf{k}}) \rangle. \quad (6.19)$$

At this point we note that the possible values of the wave vector \mathbf{k} dwell in 3D lattice with the lattice constant $(2\pi/L)^3$. In the thermodynamic limit when the quantization volume tends to infinity we have $L \rightarrow \infty$, the lattice becomes infinitely dense, and so it is plausible to replace the sum with an integral. To effect this we have to introduce the density of states in \mathbf{k} space, the inverse of the volume per state, to ensure that we correctly count the number of states inside a given volume in \mathbf{k} space. We have thus come to the continuum approximation, a central mathematical method in QM and QSM, which we write here in the general d -dimensional form:

$$\sum_{\mathbf{k}} f(\mathbf{k}) = \frac{L^d}{(2\pi)^d} \int d^d k f(\mathbf{k}). \quad (6.20)$$

This is valid as long as the function $f(\mathbf{k})$ varies little from site to site in the lattice of quantum states in \mathbf{k} space.

As far as we can tell, for fermions the continuum approximation is always valid in the thermodynamic limit. We may thus develop the particle number in three dimensions as follows,

$$\begin{aligned} N &= \frac{gV}{(2\pi)^3} \int d^3 k \frac{1}{z^{-1} e^{\beta \epsilon_{\mathbf{k}}} + 1} = \frac{gV}{(2\pi)^3} \int d^3 k \frac{z e^{-\beta \epsilon_{\mathbf{k}}}}{1 + z e^{-\beta \epsilon_{\mathbf{k}}}} \\ &= \frac{gV}{2\pi^2} \int k^2 dk \frac{z e^{-\beta \epsilon_k}}{1 + z e^{-\beta \epsilon_k}} = \frac{gV}{2\pi^2} \left(\frac{2mkT}{\hbar^2} \right)^{3/2} \int dx x^2 \frac{z e^{-x^2}}{1 + z e^{-x^2}} \\ &= \frac{gV}{2\pi^2} \left(\frac{2mkT}{\hbar^2} \right)^{3/2} \int dx x^2 z e^{-x^2} (1 - z e^{-x^2} + z^2 (e^{-x^2})^2 - \dots) \\ &= \frac{gV}{2\pi^2} \left(\frac{2mkT}{\hbar^2} \right)^{3/2} \int dx x^2 (e^{-x^2} - z^2 e^{-2x^2} + z^3 e^{-3x^2} - \dots) \\ &= \frac{gV}{\lambda^3} f_{3/2}(z). \end{aligned} \quad (6.21)$$

Here the standard Gaussian integral

$$\int_0^\infty dx x^2 e^{-\lambda x^2} = \frac{\sqrt{\pi}}{4\lambda^{3/2}} \quad (6.22)$$

has been used to define the *thermal de Broglie wavelength*

$$\lambda = \sqrt{\frac{2\pi\hbar^2}{mkT}}, \quad (6.23)$$

and an instance of a class of functions $f_\alpha(z)$,

$$f_\alpha(z) = \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1} z^\ell}{\ell^\alpha}. \quad (6.24)$$

Of course, the expansion as in the third line in Eq. (6.21) is only valid for $|z| < 1$.

We have thereby expressed the density $n = N/V$ of the ideal Fermi gas in the form

$$n = \frac{g}{\lambda^3} f_{3/2}(z). \quad (6.25)$$

In the same way the internal energy becomes

$$U = \sum_i \langle n_i \rangle \epsilon_i = \frac{3}{2} \frac{gV kT}{\lambda^3} f_{5/2}(z). \quad (6.26)$$

As to grand partition function, we might write

$$\begin{aligned} \ln \mathcal{Z} &= \sum_i \ln(1 + z e^{-\beta \epsilon_i}) = \frac{gV}{(2\pi)^3} \int d^3k \ln(1 + z e^{-\beta \epsilon_k}) \\ &= \frac{gV}{(2\pi)^3} \int d^3k (z e^{-\beta \epsilon_k} - z^2 e^{-2\beta \epsilon_k} + z^3 e^{-3\beta \epsilon_k} - \dots), \end{aligned} \quad (6.27)$$

to obtain

$$\Omega = -kT \ln \mathcal{Z} = -\frac{gV kT}{\lambda^3} f_{5/2}(z). \quad (6.28)$$

We already know that $\Omega = -pV$, so we have, for instance, the following observations:

$$\frac{p}{kT} = \frac{g}{\lambda^3} f_{5/2}(z); \quad U = \frac{3}{2} pV. \quad (6.29)$$

6.3.1 High temperature/low density limit

Usually one wishes to eliminate the fugacity z from the theory to obtain expressions such as $U = U(n, T)$, $p = p(n, T)$, and so on. How this might work may be seen from the following. First take the chain of equalities

$$\frac{n\lambda^3}{g} \equiv z_0 = f_{3/2}(z) = z - \frac{z^2}{2\sqrt{2}} + \frac{z^3}{3\sqrt{3}} - \dots \quad (6.30)$$

In the limit of high T or low n , obviously $z_0 \simeq z \ll 1$. We thus *invert the power series* by inserting the ansatz

$$z = z_0 + \alpha_2 z_0^2 + \alpha_3 z_0^3 + \dots, \quad (6.31)$$

into the power series in (6.30), and attempt to satisfy the ensuing equation for z_0 order by order in z_0 . Thus we obtain

$$z_0 \equiv z_0 + \left(\alpha_2 - \frac{1}{2\sqrt{2}} \right) z_0^2 + \left(\alpha_3 - \frac{\alpha_2}{\sqrt{2}} - \frac{1}{3\sqrt{3}} \right) z_0^3 + \dots, \quad (6.32)$$

which gives

$$\alpha_2 = \frac{1}{2\sqrt{2}}, \quad \alpha_3 = \frac{9 - 4\sqrt{3}}{36}, \quad \dots, \quad (6.33)$$

and so we have the inverted series

$$z = z_0 + \frac{1}{2\sqrt{2}} z_0^2 + \frac{9 - 4\sqrt{3}}{36} z_0^3 + \dots \quad (6.34)$$

Inserting now the power series (6.34) into the power series (6.26) for the internal energy and developing the result into a power series in z_0 , we find

$$U = \frac{3}{2} kTN \left[1 + \frac{1}{4\sqrt{2}} \left(\frac{n\lambda^3}{g} \right) + \dots \right]. \quad (6.35)$$

The leading term here is the energy of the classical ideal gas. The quantum corrections, *degeneracy effects*, become important when the particle density is high enough or the temperature is low enough to satisfy

$$\frac{n\lambda^3}{g} = \frac{N\lambda^3}{gV} = \left(\frac{\lambda}{\ell} \right)^3 \sim 1; \quad \ell = \left(\frac{Vg}{N} \right)^{1/3}, \quad (6.36)$$

i.e., when the characteristic distance ℓ between the particles of the same species (here, each m_s) is less than the thermal de Broglie wavelength.

It should be observed that, for instance in the limit of $n \rightarrow 0$ with a fixed T , we have $z \simeq z_0 \rightarrow 0$, so that $\mu = kT \ln z \rightarrow -\infty$. Moreover, the thermal occupation numbers are small and satisfy

$$\langle n_i \rangle = \frac{1}{z^{-1} e^{\beta \epsilon_i} + 1} \simeq z e^{-\beta \epsilon_i}. \quad (6.37)$$

The occupation numbers of one-particle states will become proportional to the Maxwell-Boltzmann factor $e^{-\beta \epsilon_i}$, a circumstance known as the *Maxwell-Boltzmann* statistics. Maxwell-Boltzmann statistics readily emerges if each particle is regarded as a QSM system independent of the other particles, but that is permissible with indistinguishable particles only in suitable limiting cases. Basically the particles have to be far enough apart, further than the thermal de Broglie wavelength from one another, so that one can effectively distinguish between them...

6.3.2 Low temperature/high density

The case $T = 0$

At exact zero temperature, or at “low enough” temperature, we may take $\beta \rightarrow \infty$, and the occupation numbers become

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1} = \theta(\mu - \epsilon_i), \quad (6.38)$$

where θ stands for the Heaviside unit step function:

$$\theta(x) = \begin{cases} 0 & x < 0, \\ 1 & x > 0. \end{cases} \quad (6.39)$$

All one-particle states with the energy up to the chemical potential are fully occupied, $\langle n_i \rangle = 1$, and the states with the energy above the chemical potential are empty, $\langle n_i \rangle = 0$. This situation is referred to as the *Fermi sea*.

An attempt to calculate the particle number now gives

$$N = \frac{Vg}{(2\pi)^3} \int d^3k \theta(\mu - \epsilon_k) = \frac{Vg}{(2\pi)^3} \int_{\epsilon_k < \mu} d^3k = \frac{Vg}{(2\pi)^3} \frac{4\pi}{3} \left(\frac{2m\mu}{\hbar^2} \right)^{3/2}, \quad (6.40)$$

which makes use of the volume in \mathbf{k} space of the sphere with the radius $k = \sqrt{2m\mu/\hbar^2}$. Solving for the chemical potential, we find the value

$$\mu(T=0) \equiv \epsilon_F = \frac{\hbar^2 k_F^2}{2m}, \quad k_F = \left(\frac{6\pi^2 n}{g} \right)^{1/3}. \quad (6.41)$$

The chemical potential at zero temperature is called *Fermi energy*, hence the standard notation ϵ_F . Similarly, k_F is the Fermi wave number, $p_F = \hbar k_F$ is the Fermi momentum, $v_F = p_F/m$ is the Fermi velocity, $\theta_F = \epsilon_F/k_B$ is the Fermi temperature, and so on.

Example: Average energy per particle. As a further example of the limit $T \rightarrow 0$, consider the average energy per particle. We find

$$\frac{\langle E \rangle}{\langle N \rangle} = \frac{\int_{k < k_F} d^3k \epsilon_k}{\int_{k < k_F} d^3k} = \frac{\int_{k < k_F} dk k^2 \epsilon_k}{\int_{k < k_F} dk k^2} = \frac{\frac{\hbar^2}{2m} \int_{k=0}^{k_F} dk k^4}{\int_{k=0}^{k_F} dk k^2} = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} = \frac{3}{5} \epsilon_F. \quad (6.42)$$

Lots of common factors cancel in this way, and a depressing soup of constants is avoided.

$T \neq 0$, but “small”

Close to zero temperature the limits $\mu \rightarrow \epsilon_F > 0$ and $\beta \rightarrow \infty$ apply, and so we also have the limit $z = e^{\beta\mu} \rightarrow \infty$. It should be clear that for $T \rightarrow 0$ and/or $n \rightarrow \infty$ the expansions in fugacity we have pursued in a few of the preceding sections become useless.

In the limit of low temperature another expansion is therefore employed that is based on the observation that the Fermi function $f(\epsilon) = (1 + e^{\beta(\epsilon - \mu)})^{-1}$ changes very quickly from 1 to 0 when the energy grows past the chemical potential. In fact the derivative of f is almost a delta function, $f'(\epsilon) \simeq -\delta(\epsilon - \mu)$. This line of thought leads to what is called *Sommerfeld theory of metals*. We will not go into the details but simply quote one landmark result, namely, the heat capacity of an idea Fermi gas at low temperature. It is

$$C_V = \frac{\pi^2}{3} Nk \left(\frac{kT}{\epsilon_F} \right) + \mathcal{O} \left[\left(\frac{kT}{\epsilon_F} \right)^2 \right]. \quad (6.43)$$

Besides predicting that at low temperature the heat capacity is proportional to temperature, this expression also reveals the meaning of the concept of low temperature. Namely, in order to decide whether the temperature is low, it has to be compared with something. The correct comparison is between thermal energy kT and Fermi energy ϵ_F .

6.4 Ideal Bose gas

As far as expansions in fugacity go, the difference between Bose and Fermi gases is a couple of sign changes in the basic starting points, which translate to some sign changes in the final results. For instance, we have

$$n = \frac{g}{\lambda^3} g_{3/2}(z); \quad U = \frac{3}{2} \frac{gV kT}{\lambda^3} g_{5/2}(z); \quad \Omega = -pV = -\frac{2}{3}U; \quad \dots \quad (6.44)$$

Here the family of functions $g_\alpha(z)$ is defined as

$$g_\alpha(z) = \sum_{\ell=1}^{\infty} \frac{z^\ell}{\ell^\alpha}. \quad (6.45)$$

Inversion of a power series exactly analogously to the case of fermions gives in the high temperature/low density limit the results

$$z = z_0 - \frac{z_0^2}{2\sqrt{2}} + \dots, \quad U = \frac{3}{2} N k T \left[1 - \frac{1}{4\sqrt{2}} \left(\frac{n\lambda^3}{g} \right) + \dots \right]. \quad (6.46)$$

In the low-density limit, the same Maxwell-Boltzmann statistics emerges that was already found for fermions. The condition for the effects of quantum statistics to emerge, i.e., the condition for quantum degeneracy, again is that the distance between the particles is comparable to, or less than, the thermal de Broglie wavelength. For boson the convention is to downright introduce the parameter

$$\xi = \frac{n\lambda^3}{g} \quad (6.47)$$

called *phase space density*, and characterize degeneracy by the condition that $\xi \geq 1$.

6.4.1 Bose-Einstein condensation

In the limit of low temperature or high density, the ideal Bose gas shows a behavior much different from the behavior of the Fermi gas.

To start with, let us write the equation of density as

$$n = \frac{g}{\lambda^3} g_{3/2}(z). \quad (6.48)$$

The special function here is

$$g_{3/2}(z) = z + \frac{z^2}{2^{3/2}} + \frac{z^3}{3^{3/2}} + \dots \quad (6.49)$$

This is obviously an increasing function of its argument z for all values $0 \leq z < 1$, and besides, the series converges also at $z = 1$ to the value $g_{3/2}(1) = 2.612$. On the other hand, the fugacity z , positive by definition, must remain less than one. Otherwise the occupation number of the ground state

$$\langle n_0 \rangle = \frac{z}{1-z} \quad (6.50)$$

would be either infinite or negative. We then have from (6.48) the apparent result that

$$n < n_c = \frac{g}{\lambda^3} g_{3/2}(1). \quad (6.51)$$

Somehow, there should be an upper limit to the density of an ideal Bose gas, with no interaction at all between the particles. That obviously cannot be right.

The explanation is that the continuum limit (6.20) used to derive (6.48) fails. When the occupation number of the ground state gets singular with $z \rightarrow 1$, one can no longer replace the sum over the wave vectors \mathbf{k} with an integral. Instead, the ground state $\mathbf{k} = 0$ acquires a macroscopic occupation: the ratio $\langle n_0 \rangle / N$ does not tend to zero even in the thermodynamical limit. The particles in the ground state make what is called the *Bose-Einstein condensate* (BEC). The density above the critical density simply is for particles that reside in the condensate;

$$n_0 = n - n_c. \quad (6.52)$$

To address this phenomenology quantitatively, the starting point is the following theorem: Suppose that the dispersion relation for the particles in a Bose-Einstein ideal gas is quadratic in the limit $\mathbf{k} \rightarrow 0$, and that $f(\mathbf{k})$ is a continuous function of the wave vector in some neighborhood of $\mathbf{k} = 0$. In three dimensions, in the thermodynamical limit, one may then write

$$\sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle f(\mathbf{k}) = \frac{V}{(2\pi)^3} \int d^3k \frac{1}{z^{-1} e^{\beta \epsilon_{\mathbf{k}}} - 1} f(\mathbf{k}) + \frac{z}{1-z} f(0). \quad (6.53)$$

The proof is an exercise in classical analysis, and will not be given here. The physical meaning is that the continuum approximation takes care of all other states except the ground state, which must be treated separately.

Armed with the theorem, let us now discuss in more detail what happens when one attempts to exceed the critical density, so that $n > n_c$. By (6.53), the density is

$$n = \frac{g}{\lambda^3} g_{3/2}(z) + \frac{g}{V} \frac{z}{1-z}. \quad (6.54)$$

Let us tentatively assume that the fugacity will be close to one, and write $z = 1 - \Delta$ with $0 < \Delta \ll 1$. Equation (6.54) is easily manipulated into the form

$$\Delta = \frac{\frac{g}{V}(1-\Delta)}{n - \frac{g}{\lambda^3} g_{3/2}(1-\Delta)} \simeq \frac{g}{V(n - n_c)}, \quad (6.55)$$

where the second form is the more accurate, the smaller Δ is. But now, in the thermodynamical limit $V \rightarrow \infty$ holds, and we clearly have $\Delta \rightarrow 0$. So, in the thermodynamical limit the fugacity approaches unity as

$$z \simeq 1 - \frac{K}{V} \quad (6.56)$$

for some positive constant K .

The immediate consequence is that the continuum-approximation part in (6.54) approaches the critical density. Correspondingly, the density of the particles in the ground state, as per the second term on the right-hand side of (6.54), approaches the BEC density n_0 as in (6.52).

As another example of the application of (6.53) we consider entropy. For an ideal Bose gas it is given by

$$S = k \sum_i [(1 + \langle n_i \rangle) \ln(1 + \langle n_i \rangle) - \langle n_i \rangle \ln \langle n_i \rangle], \quad (6.57)$$

which may also be expressed as a sum in the form (6.53). However, the function $f(\mathbf{k})$ as in (6.53),

$$f(\mathbf{k}) = \frac{k(1 + \langle n_{\mathbf{k}} \rangle) \ln(1 + \langle n_{\mathbf{k}} \rangle) - \langle n_{\mathbf{k}} \rangle \ln \langle n_{\mathbf{k}} \rangle}{\langle n_{\mathbf{k}} \rangle}, \quad (6.58)$$

is here expressed in terms of the occupation numbers $\langle n_{\mathbf{k}} \rangle$ themselves, and could potentially behave badly around $\mathbf{k} = 0$. Fortunately, it does not, as is easily seen by studying this expression for large values of $\langle n_{\mathbf{k}} \rangle$. In fact, we have $f(\mathbf{k}) \rightarrow 0$ as $\langle n_{\mathbf{k}} \rangle \rightarrow \infty$, and in the thermodynamical limit $f(\mathbf{k})$ converges to a continuous (albeit not analytic) function of \mathbf{k} . The theorem (6.53) applies.

Let us look at the entropy of the BEC, which comes from the occupation number $\langle n_0 \rangle = z/(1 - z)$. In the thermodynamical limit it scales as

$$\begin{aligned} S_0 &= gk \left[\left(1 - \frac{z}{1 - z}\right) \ln \left(1 - \frac{z}{1 - z}\right) - \frac{z}{1 - z} \ln \frac{z}{1 - z} \right] \\ &= \ln V + \mathcal{O}(V^0), \end{aligned} \quad (6.59)$$

where (6.56) has been applied. The entropy of the condensate tends to zero logarithmically in the thermodynamical limit, which means that it may be neglected.

In fact, the condensate has virtually no thermodynamics except particle number. Not only the entropy of the condensate, but also its internal energy, pressure, etc., vanish in the thermodynamic limit. Truth to be told, these results are only true for an ideal gas. Interactions between the particles may change the picture substantially, but this problem area is beyond the scope of our discussion.

6.5 Massive ideal gases

Our discussions would be of very limited interest if there were no systems in Nature that at least to a fair approximation are ideal Fermi or Bose gases. Here we give a few examples of massive near-ideal quantum gases.

6.5.1 Conduction electrons in a metal

Nearly free conduction electrons in a metal are an abundant example of an everyday-life Fermi gas. It is easy to estimate the density of conduction electrons from considerations such as density of the metal, atom number, and valence, and then come up with an estimate of Fermi energy. The result is $\epsilon_F \sim 1$ eV. On the other hand, at room temperature $T \sim 300$ K, the thermal energy to the order of magnitude is $kT \sim 0.01$ eV. The conduction electrons are in the low-temperature limit.

One might quibble about the approximations of our theory. Unlike we have assumed, the electrons in a metal are not free, but move on the background of the lattice of positively charged ions. It turns out, however, that the effect of the crystal lattice is to modify the dispersion relation of the electrons, and that many a time a sufficient account of the modifications is to simply replace the mass of the electron m by some effective mass m^* that may differ from m by, say, a factor of two or three.

Moreover, the electrons are not an ideal gas, but have the Coulomb repulsion between them; as strong an interaction as there is outside of the realm of nuclear and particle physics. Several features, however, mitigate the Coulomb interactions. First, the positive ions of the atoms that have released the electrons provide a neutralizing background, so that the gas of the electrons does not fly apart because of the Coulomb repulsions. Second, at low temperature the Fermi statistics strongly suppresses the Coulomb interactions. While the details are again outside the scope of the present exposition, you may think of Coulomb interactions as scattering between pairs of electrons. It turns out that when you take into account the constraints of electron scattering, namely, conservation of energy and momentum, at low temperatures most scattering events attempt to put electrons to states that are already occupied. By the Pauli exclusion principle that will not work, and so the Coulomb interactions have little effect.

All told, to zeroth order you may think of conduction electrons in a metal as a free, noninteracting, low temperature Fermi gas. Along with this concept comes the prediction of linear heat capacity characteristic of the Fermi gas at low temperature. This can actually be observed, although only at temperatures below 10 K to the order of magnitude. This is not because the heat capacity would not be linear also at room temperature, but because in crystalline matter there are other heat capacities that overwhelm the heat capacity of the Fermi gas.

6.5.2 Bose-Einstein condensation

The superfluidity in liquid ^4He observed in 1937 is presently thought to be a manifestation of Bose-Einstein condensation. These days, however, the most active direction of research is BEC in alkali metal vapors, first observed in 1995.

In a BEC experiment one typically starts by laser-cooling an alkali metal vapor to a temperature on the order of $100\ \mu\text{K}$, the density being $\sim 10^{12}\text{ cm}^{-3}$. The phase space densities obtained in this way fall short of BEC by a factor of about 10^6 , so that further cooling and compression of the gas is due. To this end the gas is most often transferred to a magnetic trap. The idea is that the angular momentum (spin) of an atom locks on to the direction of the magnetic field and follows it even when the atom moves and the direction of the local field changes accordingly. An atom then experiences a potential energy of the form $U = \mu|\mathbf{B}(\mathbf{r})|$, and may get trapped around a minimum of the absolute value of the magnetic field. The third ingredient in the experiments is evaporative cooling. In very basic terms, the depth of the trap is finite, so sufficiently energetic atoms escape and the rest thermalize at a lower temperature. You lose some atoms, but the atoms get colder and therefore also collect closer to the minimum of the trapping potential. In this way a sufficient phase space density for condensation, $\xi \simeq 2.612$, may be reached. Typically, 10^7 atoms condense at the temperature of 100 nK , and the largest dimension of the condensate is about 0.1 mm .

The condensate is in a trap, not in translationally invariant free space, which modifies some features of the physics. For instance, if the trap is very strong in one or two spatial directions, the system may effectively be two- or one-dimensional. Atoms also interact, and the interaction modifies the properties of the condensate in an essential way. The interactions may be “weak” in the many-body sense, in which case simple methods suffice for an excellent theoretical description of the condensate, or the interactions may be “strong”, whereupon cutting-edge methods for “strongly correlated systems” from condensed-matter physics are called upon. Moreover, it is possible to achieve quantum degenerate Fermi gases as well, which add many more issues to study. For instance, the isotopes of potassium ^{39}K and ^{40}K make, respectively, bosonic and fermionic atoms.

All told, since 1995, Bose-Einstein condensation in alkali metal vapors has completely transformed atomic, molecular and optical physics, and given a new lease on life to many branches of condensed matter physics as well.

6.6 Ideal gases of bosonic excitations

A standard theme in many branches of physics is to find the ground state of a system first, and then the small excitations around the ground state. For instance, one can envisage solving the stationary configuration of a solid, obtain the forces that the atoms exert on each other when displaced slightly from the equilibrium, and finally obtain the normal modes of small oscillations of

the atoms around the equilibrium. As we all know from classical mechanics, the normal modes are formally harmonic oscillators. In fact, in the harmonic approximation, the normal modes are *independent* harmonic oscillators. One can then quantize the harmonic oscillators, and obtain a quantum mechanical description of the excitations of the system.

Now, although this is not obvious without a more detailed many-body analysis, there is a one-to-one mapping between bosons and harmonic oscillators. The bosons are basically the excitations of the oscillators. So, if an oscillator is moved from an eigenstate with the quantum number $n = 1$ to the quantum number $n = 2$, in the corresponding boson description the number of bosons in the oscillator mode has increased from 1 to 2. The quantized oscillators have their own quantum statistical mechanics. But, given the connection to fictitious bosonic particles, the bosons have a corresponding statistical mechanics. Inasmuch as the harmonic oscillators are uncoupled, as they are in the harmonic approximation to lattice vibrations, the corresponding bosons make a noninteracting ideal gas. The theme of this section is such ideal gases of bosonic excitations.

One common feature of the excitations is that, if the energy of the state without excitations (the vacuum) is defined at zero, the chemical potential of the bosons is $\mu = 0$. Ultimately this is because we are dealing with harmonic oscillators, which in and of itself already determines the statistical mechanics and makes the corresponding boson theory have the zero chemical potential. A rule of thumb to memorize the zero chemical potential states that, since there is no number conservation for such bosons, there is no need for a chemical potential.

6.6.1 Photons

Here we are looking for the eigenmodes of the electromagnetic fields analogous to the stationary solutions of the Schrödinger equation. We study electric and magnetic fields in free space, in the absence of charges and currents. We again introduce an auxiliary cube of side L and volume $V = L^3$, and require that the electric and magnetic fields obey periodic boundary conditions across the cube. There obviously is no physical cavity that realizes such boundary conditions, but the idea is to eventually make the cube infinitely large whereupon the boundary conditions cannot, and will not, matter.

Under such conditions the eigenmodes of Maxwell's equations are plane waves characterized by the mode index $q \equiv (\mathbf{q}, \lambda)$. Here the wave vector can only have discrete values according to

$$\mathbf{q} = \frac{2\pi}{L}(n_1\hat{\mathbf{e}}_1 + n_2\hat{\mathbf{e}}_2 + n_3\hat{\mathbf{e}}_3), \quad n_i = 0, \pm 1, \pm 2, \dots, \quad (6.60)$$

exactly like the plane waves in quantum mechanics. The index λ stands for the polarization of the plane-wave mode. It has two possible values that could stand, for instance, for left- and right-circular polarization. The index λ is an

analog of the magnetic quantum number m_s for particles with a spin. These eigenmodes oscillate at the frequencies

$$\omega_q = c|\mathbf{q}|. \quad (6.61)$$

In quantum theory we have bosonic particles called *photons* with the quantum numbers q and the corresponding one-particle energies

$$\epsilon_q = \hbar\omega_q = \hbar c|\mathbf{q}|. \quad (6.62)$$

Bosons should have integer spin and an odd number of spin components while we have only two values for the polarization λ . This discrepancy is resolved by the rule that for massless particles, dispersion relation of the form (6.62), there cannot be a component of the spin in the direction of motion of the boson. Electromagnetic radiation in thermal equilibrium, *black body radiation*, is an ideal gas of photons with the chemical potential $\mu = 0$.

Let us begin by considering the energy of the photon gas. We have

$$\begin{aligned} U &= \sum_q \langle n_q \rangle \epsilon_q = 2 \frac{V}{(2\pi)^3} \int d^3q \frac{\hbar c|\mathbf{q}|}{e^{-\beta\hbar c|\mathbf{q}|} - 1} = \frac{V}{\pi^2} \int dq q^2 \frac{\hbar c q}{e^{-\beta\hbar c q} - 1} \\ &= \frac{\hbar V}{\pi^2 c^3} \int_0^\infty \frac{d\omega \omega^3}{e^{\frac{\hbar\omega}{kT}} - 1}. \end{aligned} \quad (6.63)$$

Continuum approximation works here without a problem, and in the last step the integration variable has been changed from q to $\omega = cq$. The energy density is therefore written

$$E = \frac{U}{V} = \int_0^\infty d\omega E(\omega); \quad E(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\frac{\hbar\omega}{kT}} - 1}. \quad (6.64)$$

The quantity $E(\omega)$ is obviously the energy density in the radiation per unit frequency (angular frequency) interval.

The result (6.64) is called Planck's radiation law, and it correctly reproduces the observed spectrum of black body radiation. The energy density cuts off both at low and at high frequencies, and has a maximum at about $\omega = kT/\hbar$.

The total energy in black body radiation is obtained by integration,

$$E = \frac{k^4 T^4}{\pi^2 \hbar^3 c^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{4}{c} \sigma T^4, \quad (6.65)$$

where the known value of the integral $\pi^4/15$ gives the *Stefan-Boltzmann constant* σ as

$$\sigma = \frac{\pi^2 k^4}{60 \hbar^3 c^2} = 5.7 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}. \quad (6.66)$$

It remains to resolve the odd-looking factor $4/c$ in the density of black body radiation. Now, to establish a black body in a laboratory, one might take a cavity whose walls are blackened, say, with soot, and keep it at the temperature

T . In equilibrium the electromagnetic field inside also attains the temperature T , but the radiation inside a cavity might not be so useful for experiments. Therefore one imagines opening a small hole in the wall of the cavity, so that some of the radiation escapes. The hole is assumed “small” because then its presence does not significantly alter the black body radiation inside.

Suppose that the environment is at zero temperature or at least at a much lower temperature than the cavity, then the only flux of energy is from the inside of the cavity out. Flux of energy equals energy density times the velocity at which the energy carriers move, here the speed of light c . However, in black body radiation the photons move in all directions and only the component of the velocity perpendicular to the hole toward the outside contributes to the energy flux out of the cavity. Setting the z axis normal to the hole, we have the energy flux or *intensity* of the black body radiation

$$I = cE \frac{1}{4\pi} \int_{-\pi}^{\pi} d\phi \int_0^{\pi/2} d\theta \sin \theta \cos \theta = \sigma T^4. \quad (6.67)$$

The 4π in the denominator is to effect the result cE if by some fiat the directions did not matter, $\cos \theta$ is for the projection of the photon velocity to the direction straight out of the hole, and the polar integral is capped $\pi/2$ because at larger angles θ the purportedly escaping photons would actually move toward the inside of the cavity.

Equation (6.67) is the basic result for the radiation of energy from hot bodies. A body that is not black radiates less. Namely, if the surface absorbs the fraction α of the radiation falling on it, it can only emit at the fraction α of the black body intensity. Otherwise the surface could not remain in equilibrium with thermal radiation at the same temperature, but there would be a net flux of energy in to, or out of, the surface.

6.6.2 Phonons

To get at the thermal properties of solids, a little recap of the structure of crystalline matter is in order. Crystal structure results from a combination of two elements: lattice, and at each lattice point an identical set of atom called basis. Our crystal is comprised of N lattice sites, and has a basis with p atoms.

As we already have alluded to, in the harmonic approximation, in which the potential energy of the atoms in the crystal is expanded only up to second order in the displacements from equilibrium, there will be independent normal modes of lattice vibrations. In what are known as periodic Born-von Karman boundary conditions, such a mode can be characterized once more by a wave vector \mathbf{q} and a polarization index λ . The wave vectors are quantized on a lattice whose structure depends on the structure of the original lattice, but remarkably, the density of states in \mathbf{q} space and the volume V of the crystal are still related by $V/(2\pi)^3$, so that the continuum approximation works almost as before.

The caveat “almost” is because there is only a finite number of lattice vibration modes; in three dimensions there are $3pN$ degrees of freedom for the motion

of the atoms, and lattice vibrations for sure cannot have more. In fact, there are N possible values for the wave vectors \mathbf{q} . The choice of the wave vectors is not unique, but the usual convention is to require that they reside in what is known as the first Brillouin zone of the reciprocal lattice. The polarization modes λ enumerate how the lattice atoms move with respect to one another in a plane-wave vibration mode with the wave number \mathbf{q} . In a three-dimensional crystal structure there are $3p$ polarization modes λ .

The polarization modes further fall in two categories. There will be 3 *acoustic modes* λ such that in the long-wavelength limit $\mathbf{q} \rightarrow 0$ the dispersion relation is linear: $\omega_{\mathbf{q},\lambda} \simeq c_\lambda |\mathbf{q}|$, where c_λ is the speed of sound for the given acoustic mode. The frequencies of the rest of the modes, *optical modes*, behave as $\omega_{\mathbf{q},\lambda} \rightarrow \omega_\lambda > 0$ as $\mathbf{q} \rightarrow 0$.

In the quantum theory the normal modes of lattice vibrations are replaced by an ideal gas of bosons called *phonons*. The energy of a phonon with the quantum numbers (\mathbf{q}, λ) is $\epsilon_{\mathbf{q},\lambda} = \hbar\omega_{\mathbf{q},\lambda}$, and, as usual, the chemical potential of a phonon gas is $\mu = 0$. For instance, the thermal occupation numbers and the energy of the phonon gas are then

$$\langle n_{\mathbf{q},\lambda} \rangle = \frac{1}{e^{\frac{\hbar\omega_{\mathbf{q},\lambda}}{kT}} - 1}, \quad U = \sum_{\mathbf{q},\lambda} \frac{\hbar\omega_{\mathbf{q},\lambda}}{e^{\frac{\hbar\omega_{\mathbf{q},\lambda}}{kT}} - 1}. \quad (6.68)$$

High temperature

In the limit of $T \rightarrow \infty$ we have the approximation

$$e^{\frac{\hbar\omega_{\mathbf{q},\lambda}}{kT}} - 1 \simeq \frac{\hbar\omega_{\mathbf{q},\lambda}}{kT}. \quad (6.69)$$

Moreover, since there is only a finite number of modes and the mode frequencies are therefore limited from above, the high-temperature limit applies uniformly to all modes. We have the high-temperature limit

$$U \simeq \sum_{\mathbf{q},\lambda} kT = 3pNkT, \quad (6.70)$$

and hence the heat capacity

$$C_V = 3pNk. \quad (6.71)$$

This is called the *Dulong-Petit* law.

Low temperature

The frequencies of an optical mode λ are limited both from above and from below, as in

$$0 < \omega_\lambda \leq \omega_{\mathbf{q},\lambda} \leq \Omega_\lambda < \infty. \quad (6.72)$$

We therefore have the upper limit for the energy in the optical mode λ

$$U_\lambda = \sum_{\mathbf{q}} \frac{\hbar\omega_{\mathbf{q},\lambda}}{e^{\frac{\hbar\omega_{\mathbf{q},\lambda}}{kT}} - 1} \leq \sum_{\mathbf{q}} \frac{\hbar\Omega_\lambda}{e^{\frac{\hbar\omega}{kT}} - 1} = \frac{N\hbar\Omega_\lambda}{e^{\frac{\hbar\omega}{kT}} - 1} \sim \exp\left[-\frac{\hbar\omega}{kT}\right], \quad (6.73)$$

which becomes exponentially small as $T \rightarrow 0$. In the low-temperature limit the optical modes contribute no energy, nor heat capacity, to the phonon gas.

As it comes to an acoustic mode λ , only the wave vectors for which $\hbar\omega_{\mathbf{q},\lambda} \leq kT$ effectively contribute. In the low-temperature limit we may therefore write $\omega_{\mathbf{q},\lambda} \simeq c_\lambda|\mathbf{q}|$. Second, when we sum over the wave vectors, technically we should sum over the first Brillouin zone; but since only modes close to $\mathbf{q} \sim 0$ contribute anyway, we may just as well sum over all of the \mathbf{q} space. And finally, the continuum approximation works as usual. We therefore have

$$U = \frac{V}{(2\pi)^3} \int d^3q \frac{\hbar c_\lambda |\mathbf{q}|}{e^{\frac{\hbar c_\lambda |\mathbf{q}|}{kT}} - 1} = \frac{V \pi^2 k^4 T^4}{15 \hbar^3 c_\lambda^3} \propto T^4. \quad (6.74)$$

Notice that the calculation is materially the same as the calculation of the energy of a photon gas. The heat capacity of a phonon gas at low temperatures comes from acoustic modes, and behaves with temperature as

$$C_V \propto T^3. \quad (6.75)$$

Arbitrary temperature

As an interesting and useful variation of the analysis of the energy of acoustic modes, one could construct the *Debye model* according to which $\omega_{\mathbf{q},\lambda} = c_\lambda|\mathbf{q}|$ holds for all wave vectors, and the first Brillouin zone is replaced with a sphere in \mathbf{q} space with the radius q_D chosen in such a way that the sphere contains the right number N of phonon modes. We will not show the details, but merely note that the temperature corresponding to q_D , $\theta_D = \hbar c_\lambda q_D / k$, is called the Debye temperature.

Moving from the low temperature limit to the high temperature limit, the temperature dependence of the heat capacity of the phonon gas in a solid switches from T^3 to T^0 . This behavior of the heat capacity is actually observed. There is a characteristic temperature at which the switchover takes place, the precise definition of which could be stated in various ways. For instance, it could be the Debye temperature in the Debye theory that best fits the observed temperature dependence of the heat capacity. Interestingly, the characteristic temperature tends to be on the order of ordinary room temperature.

As a final note, the linear (in temperature) heat capacity of an electron gas is difficult to see in metals because metals are (usually) crystalline solids, and the heat capacity of the phonons greatly exceeds the heat capacity of the electrons. However, at low enough temperatures the phonon heat capacity $\propto T^3$ eventually becomes smaller than the heat capacity of the electrons $\propto T$.

6.7 Exercises

6.1. A system consists of two particles, each of which has two possible quantum states with energies $\pm\epsilon/2$. Write down the canonical partition function if (a)

the particles are distinguishable; (b) the particles are bosons; (c) the particles are fermions.

To solve part (a), remember (and forever memorize) that for two (or more) independent systems the partition function factorizes; e.g., $Z_{1+2} = Z_1 Z_2$.

6.2. Consider noninteracting fermions having just three possible one-particle states $|1\rangle$, $|2\rangle$, and $|3\rangle$ with the respective energies $\epsilon_1 = 0$, $\epsilon_2 = \epsilon_3 = \epsilon$. Find all canonical partition functions for different numbers of particles, and from them the grand partition function as a function of temperature T and chemical potential μ . Is the result as it should be?

6.3. Calculate the integrals

$$\int_0^\infty dx x^n e^{-\lambda x}, \quad \int_0^\infty dx x^n e^{-\lambda x^2}.$$

For the purposes of this exercise it is enough to find an algorithm for the values of the integrals for $n = 0, 1, \dots$, assuming $\lambda > 0$; however, the results are good for a complex parameter λ as well, provided $\Re(\lambda) > 0$.

6.4. Think about a quantum mechanical ideal gas of massive particles; in fact, this argument works equally well for Bose-Einstein, Fermi-Dirac, and Maxwell-Boltzmann gases. It is well known that, in a monatomic ideal gas where center-of-mass motion is the only degree of freedom, the energy of a one-particle state scales with the containing volume as $V^{-2/3}$ ($\frac{\hbar^2}{mL^2}$...), and in an ideal gas the same naturally applies to many-body energy eigenstates.

(a) If the container contracts or expands, the quantum states change. Suppose, however, that the probabilities of the quantum states remain unchanged, i.e., in the expression $\rho(V) = \sum_n |n(V)\rangle p_n \langle n(V)|$ the p_n remain constants. This is a model for an adiabatic process. Explain why.

(b) Show that the state of the gas satisfies $pV = \frac{2}{3} U$.

6.5. Show (for instance by starting from the grand partition function) that the entropies of Bose-Einstein and Fermi-Dirac ideal gases are

$$S = -k \sum_i [\langle n_i \rangle \ln \langle n_i \rangle \mp (1 \pm \langle n_i \rangle) \ln(1 \pm \langle n_i \rangle)] .$$

6.6. In a magnetic field the spin degeneracy of the electron gas is lifted: the magnetic energy for the z component of the spin m_s is $E = -g\mu_B B m_s$. Hence, in equilibrium the number of electrons in the spin-up and spin-down states is no longer the same, and the electron gas acquires a macroscopic magnetization M (magnetic moment/unit volume). Show that the resulting “Pauli paramagnetic susceptibility” $\chi = \frac{\partial M}{\partial B}$ at zero temperature and zero magnetic field is

$$\chi = \frac{g^2 \mu_B^2}{4V} \mathcal{D}(\epsilon_F) = \frac{3g^2 \mu_B^2 n}{8\epsilon_F} .$$

Here $g \simeq 2$, and \mathcal{D} is the density of one-particle energy eigenstates

$$\mathcal{D}(\epsilon) = \frac{\partial \mathcal{N}}{\partial \epsilon}.$$

$\mathcal{N}(\epsilon)$, in turn, is the number of one-particle energy eigenstates with kinetic energy $\leq \epsilon$, as if there were no magnetic field at all.

Things to consider: This is obviously a case in which the different magnetic states are *not* degenerate. Moreover, since you are finding the susceptibility at zero magnetic field, an analysis up to the lowest nontrivial order in B suffices.

6.7. For the extremely relativistic ideal Fermi gas one may write the dispersion relation $\epsilon_{\mathbf{k}} = \hbar c |\mathbf{k}|$. Find the Fermi energy, pressure and average energy per particle at zero temperature.

6.8. A degenerate gas of $T = 0$ fermions with the spin $s = \frac{5}{2}$ occupies an infinitely large (three-dimensional) volume. Dependence of the particle energy on the wave number is given by the relativistic formula

$$\epsilon(\mathbf{k}) = \sqrt{(mc^2)^2 + (\hbar kc)^2},$$

where m is the particle mass and c is the speed of light.

(a) Calculate the number density n and the energy density u of the particles if the Fermi wave number is k_F .

(b) Find the dependence of the gas pressure p on the Fermi momentum for the asymptotic case $\hbar k_F/(mc) \rightarrow \infty$.

6.9. Find the entropy $S = Ns(n, T)$ and the chemical potential $\mu = \mu(n, T)$ of an ideal Fermi gas in the limit of a classical ideal gas ($n \rightarrow 0$ and/or $T \rightarrow \infty$). There are situations in which the absolute value of entropy matters; this is one way of getting it right.

6.10. In *two* spatial dimensions the chemical potential of an ideal Fermi gas may be obtained exactly, analytically in closed form, as a function of density and temperature. What is $\mu(n, T)$?

6.11. By comparing in two and three dimensions the expressions of the density of the ideal, massive, free Bose gas as a function of fugacity and temperature, argue that a free two-dimensional gas is not likely to undergo Bose-Einstein condensation at any nonzero temperature.

6.12. For an isotropic 3D harmonic oscillator with the angular frequency ω the energy levels are characterized by a triplet of nonnegative integers $\mathbf{n} \equiv (n_1, n_2, n_3)$, the energy being $\epsilon_{\mathbf{n}} = \hbar\omega(n_1 + n_2 + n_3 + \frac{3}{2})$. In this problem we always assume that the harmonic oscillator quantum $\hbar\omega$ is the smallest energy scale in the problem, so that the integers n_i are large and anything that depends on these integers varies little with $n_i \rightarrow n_i + 1$.

(a) Argue that in this case the continuum approximation reads

$$\sum_{\mathbf{n}} f(\mathbf{n}) \rightarrow \int d^3n f(\mathbf{n}),$$

where the integral runs over the first octant (all Cartesian coordinates positive) of the 3D space.

(b) Now a little exercise in geometry: Show that the area of the part of the plane $x + y + z = k > 0$ that belongs to the first octant is $\frac{1}{2} k^2$.

(c) As usual, the density of energy eigenstates $D(\epsilon)$ is defined as the function of energy ϵ such that for an arbitrary function of energy $g(\epsilon)$ we have

$$\sum_i g(\epsilon_i) = \int_0^\infty d\epsilon D(\epsilon) g(\epsilon).$$

Show that for the 3D harmonic oscillator the energy density, in fact, is

$$D(\epsilon) = \frac{\epsilon^2}{2(\hbar\omega)^3}.$$

6.13. As in the previous problem, the definition of the density of energy eigenstates $D(\epsilon)$ says that the sum over energies of one-particle states may be written

$$\sum_i G(\epsilon_i) = \int_0^\infty d\epsilon D(\epsilon) G(\epsilon),$$

and the particular form of $D(\epsilon)$ for an isotropic 3D harmonic oscillator (frequency ω) is

$$D(\epsilon) = \frac{\epsilon^2}{2(\hbar\omega)^3}.$$

This is a form of the continuum approximation, good when the characteristic energy scale of the problem, say, temperature, is much larger than the spacing of the oscillator energies $\hbar\omega$. Find the BEC transition temperature for a scalar ($s = 0$) Bose gas in such an isotropic 3D harmonic oscillator trap.

6.14. The familiar density of energy eigenstates $D(\epsilon)$ is defined so that it converts sums over energy eigenstates into integrals,

$$\sum_i g(\epsilon_i) \simeq \int d\epsilon D(\epsilon) g(\epsilon).$$

The approximation is the better, the more states are involved in the sum. Suppose now that $D(\epsilon) \propto \epsilon^\alpha$ as $\epsilon \rightarrow 0$.

(a) Argue that noninteracting bosons are liable to condense in a system only if the density of states is characterized by an exponent $\alpha > 0$.

(b) Show that for a free massive particle in D dimensions, $\alpha = (D - 2)/2$ holds true.

(c) The density of states is related to the number of energy eigenstates with energy less than or equal to ϵ , $N(\epsilon)$, by $D(\epsilon) = \frac{dN(\epsilon)}{d\epsilon}$. On the basis of this observation, argue that for massive particles in a harmonic-oscillator potential in D dimensions the exponent is $\alpha = D - 1$.

(d) Do you have Bose-Einstein condensation in $D = 1, 2$ or 3 dimensions if the particles are free? What if they are trapped in a harmonic potential well?

6.15. Show by direct (and unnecessarily clumsy) calculation that energy density E and pressure p of black body radiation satisfy $p = \frac{1}{3} E$.

6.16. Show that entropy and energy in black body radiation satisfy

$$S = \frac{4U}{3T}.$$

6.17. The Sun is the most conspicuous black body in sight. Its surface temperature and radius are 5700 K and 7.0×10^8 m. The Earth is orbiting around the Sun at the mean distance of 1.49×10^{11} m, basking in sunlight. The ratio of reflected to absorbed sunlight energy (“albedo”) of the Earth is 0.3.

(a) Assume that the absorbed energy is all reradiated as black body radiation. What would be the average surface temperature of the Earth?

(b) If you did it right, the answer of part (a) clearly disagrees with common sense. Which assumption in part (a), do you think, went awry? Why?

6.18. The simplest Debye model for the heat capacity of an acoustic phonon mode in a solid is based on two assumptions: (i) The dispersion relation is $\omega_{\mathbf{q}} = c|\mathbf{q}|$ for all phonons, c being the speed of sound. (ii) Instead of the proper first Brillouin zone, the phonon modes are confined to a sphere in \mathbf{q} space with the radius q_D (Debye wave number) such that the sphere contains the correct number N of wave vectors.

(a) Show that the heat capacity of the phonon mode in the Debye approximation is

$$\frac{C_V}{Nk} = 3 \left(\frac{T}{\theta_D} \right)^3 \int_0^{\theta_D/T} dx \frac{x^4 e^x}{(e^x - 1)^2},$$

with the Debye temperature $\theta_D = \hbar c q_D / k$ acting as the sole scale of temperature.

(b) Demonstrate that this expression has plausible high- T and low- T limits.

6.19. From the distance measured by stellar parallax and from the apparent brightness it can be inferred that the luminosity (total radiated power) of the brightest star in the night sky, Sirius, is 23 times that of our sun, and the spectrum tells us that the temperature is 9300 K (5700 K for the sun). What is the radius of Sirius, in units of the radius of our Sun?

6.20. Collective elementary excitations called *spin waves* or *magnons* determine the low-T specific heat of a spin system that has undergone the ferromagnetic phase transition. For a small wave vector \mathbf{q} the dispersion relation of magnons is $\omega_q \propto \mathbf{q}^2$. What is the temperature dependence of the heat capacity of magnons at low temperatures?

Chapter 7

Classical limit of statistical mechanics

Up to this point our development has been about quantum statistical mechanics. Inasmuch as one wants to use ideas such as the no-a-priori-probabilities as a cornerstone of statistical mechanics, quantum theory is really a necessity. A discrete state as in quantum mechanics can have a probability, but the issues get murky when one starts talking about, say, probability of a continuous position or momentum of a classical particle where there is no unique way of counting states and identifying probabilities for individual states. Nonetheless, much of the world around us appears to be classical, and it would seem natural to have a version of statistical mechanics also for the domain of classical physics. Our strategy is to view classical statistical mechanics as a limit of quantum statistical mechanics.

When the mission is stated in this way, the next question is, what limit. One could technically take the limit of quantum statistical mechanics for $\hbar \rightarrow 0$ using auxiliary tools such as Wigner functions. We will not proceed in this way, however, but resort to a heuristic approach.

One also has to ask when classical statistical mechanics is valid in practice. We have partial answers from our studies of ideal gases: In the limit of high temperature or low density, such that the occupation numbers of the one-particle quantum states are small. At other times, one might note that the temperature has to be high enough so that the thermal energy kT is at least comparable to the quantum mechanical energy scale of the problem; say, the quantum energy $\hbar\omega$ of a harmonic oscillator, or the Fermi energy for a Fermi gas. Instead of attempting to state such conditions beforehand, it is often easier to see when the classical result should be valid after the fact.

At any rate, besides being compatible with classical physics, classical statistical mechanics also bring with it the advantage that solving it is usually *much* easier than solving the corresponding problem in quantum statistical mechanics. The latter is our main motivation for classical statistical mechanics.

7.1 Probabilities and classical states

In quantum statistical mechanics one of the key ideas was that the state of the system is random. Clearly we need a device to characterize a random state also in classical mechanics.

In our development we use as an example an N -particle system in three spatial dimensions, but the dimensionality is in no way critical to the development. At least for a Hamiltonian classical-mechanics system a deterministic state may always be specified by giving the values of all generalized coordinates and of the corresponding conjugate momenta. In the process of laying out the connection between quantum mechanics and classical mechanics, we work exclusively using Cartesian coordinates \mathbf{r}_n and the corresponding momenta \mathbf{p}_n for the particles. The coordinates and momenta are denoted in different ways,

$$q = \{\mathbf{r}_1, \dots, \mathbf{r}_N\} = \{q_1, q_2, \dots, q_{3N}\}, \quad p = \{\mathbf{p}_1, \dots, \mathbf{p}_N\} = \{p_1, p_2, \dots, p_{3N}\}. \quad (7.1)$$

For future use we also define the corresponding differentials, e.g.,

$$dq = d^3r_1 d^3r_2 \dots d^3r_N, \quad dp = d^3p_1 d^3p_2 \dots d^3p_N. \quad (7.2)$$

To handle a random state we introduce the *phase space density* $f(p, q)$ for the particle system. It satisfies

$$f(p, q) \geq 0, \quad \int dp dq f(p, q) = 1, \quad (7.3)$$

as appropriate for a *probability density*, and in fact has the operational definition that the probability for the system to be in the phase space volume $dp dq$ around the phase-space coordinates (p, q) equals

$$dP = f(p, q) dp dq. \quad (7.4)$$

Now, according to the usual rules of probability, for a random variable with *discrete* values ξ_n that occur at the probabilities P_n , the *mean value* is

$$\langle \xi \rangle = \sum_n P_n \xi_n. \quad (7.5)$$

The version of this rule for the case when we have continuous-valued random momenta and positions obviously is

$$\langle A(p, q) \rangle = \int dp dq f(p, q) A(p, q). \quad (7.6)$$

For the given N -particle system this is operationally the same as the expectation value in quantum mechanics. Henceforth we drop the distinction in the terminology between mean and expectation values.

The conditions of positivity and normalization of the phase space density, as well as the rule for calculating averages, have obvious analogs with the quantum mechanical density operator. In fact, the set of phase space densities is even convex. We have therefore taken our first step toward the parallels between quantum mechanics and classical mechanics:

- QM: $\rho \leftrightarrow \text{CL: } f(p, q)$.

If the state of the classical-mechanics system is known with certainty to be (p_0, q_0) , the corresponding phase space density is obviously

$$f_{p_0 q_0}(p, q) = \delta(p - p_0)\delta(q - q_0). \quad (7.7)$$

Now, in quantum mechanics there were pure states, technically, corners of the convex set of states that cannot be obtained by mixing any other states. The classical counterpart is the state with no uncertainty:

- QM: ρ pure \leftrightarrow CL: $f_{p_0 q_0}(p, q) = \delta(p - p_0)\delta(q - q_0)$

If the classical system is in a precisely known state (p_0, q_0) at some reference time $t = 0$ and if its time future time evolution is deterministically determined by the present state, we can at least in principle find the mapping that gives the time evolution over the time t both in terms of the deterministic state and the corresponding phase space density,

$$\begin{aligned} (p_0, q_0) &\rightarrow (\tilde{p}, \tilde{q}) = (\tilde{p}(p_0, q_0, t), \tilde{q}(p_0, q_0, t)), \\ f_{p_0 q_0}(p, q) &\rightarrow f_{\tilde{p}, \tilde{q}}(p, q) = \delta(p - \tilde{p})\delta(q - \tilde{q}). \end{aligned} \quad (7.8)$$

From now on we assume that the classical system is Hamiltonian; and even more, that there exists a *time independent* Hamilton's function $\mathcal{H}(p, q)$ that generates the time evolution of the coordinates and momenta according to Hamilton's equations of motion

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}. \quad (7.9)$$

We then have the equation of motion for the pure states

$$\begin{aligned} \frac{\partial}{\partial t} f_{\tilde{p}, \tilde{q}}(p, q) &= \frac{\partial}{\partial t} \delta(p - \tilde{p})\delta(q - \tilde{q}) = \left(-\dot{\tilde{p}}_i \frac{\partial}{\partial p_i} - \dot{\tilde{q}}_i \frac{\partial}{\partial q_i} \right) \delta(p - \tilde{p})\delta(q - \tilde{q}) \\ &= \left(-\frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial}{\partial q_i} + \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial}{\partial p_i} \right) f_{\tilde{p}, \tilde{q}}(p, q). \end{aligned} \quad (7.10)$$

We adopt the summation convention, in that a repeated index is automatically summed over, and for future reference write the result in a homogeneous form

$$\left(\frac{\partial}{\partial t} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial}{\partial p_i} \right) f_{\tilde{p}, \tilde{q}}(p, q) = 0. \quad (7.11)$$

Now, take an arbitrary phase space density at the initial reference time $t = 0$, $f_0(p, q)$. We may, of course, identically write it in the form

$$f_0(p, q) = \int dp_0 dq_0 f_0(p_0, q_0) f_{p_0 q_0}(p, q). \quad (7.12)$$

To proceed from here, we need to take a little detour into the probability theory. Denote momentarily $(p, q) = z$, and assume that at the initial time $t = 0$

the system is in a discrete set of pure states z_n^0 with the probabilities P_n . Over the time t these states evolve into \tilde{z}_n . Also, take an arbitrary function $A(z)$. Clearly, the expectation value of the function A must evolve from

$$\langle A \rangle_0 = \sum_n P_n A(z_n^0) = \int dz \left[\sum_n P_n f_{z_n^0}(z) \right] A(z) \quad (7.13)$$

to

$$\langle A \rangle = \sum_n P_n A(\tilde{z}_n) = \int dz \left[\sum_n P_n f_{\tilde{z}_n}(z) \right] A(z). \quad (7.14)$$

Since $A(z)$ is arbitrary, this can be possible if *and only if* the initial distribution function of the states z evolves into the final distribution according to

$$\sum_n P_n f_{z_n^0}(z) \rightarrow \sum_n P_n f_{\tilde{z}_n}(z). \quad (7.15)$$

This must hold no matter what probabilities P_n are. So, it must even hold if the distribution P_n is continuous and is characterized by an arbitrary probability density f_0 :

$$\int dz_0 f_0(z_0) f_{z_0}(z) \rightarrow \int dz_0 f_0(z_0) f_{\tilde{z}}(z). \quad (7.16)$$

Returning to the issue on hand, we have shown that the initial state (7.12) must evolve over time t to the state

$$f(p, q, t) = \int dp_0 dq_0 f_0(p_0, q_0) f_{\tilde{p}\tilde{q}}(p, q). \quad (7.17)$$

But acting on this result with the differential operator on the left-hand side of (7.11), we see that the only dependence on the variables p, q and t is in the function inside the integral $f_{\tilde{p}\tilde{q}}(p, q)$, and the differential operator gives zero when acting on it. We immediately have the results that the evolution of an arbitrary phase space density obeys

$$\left(\frac{\partial}{\partial t} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial}{\partial p_i} \right) f(p, q, t) = 0. \quad (7.18)$$

This is called the *Liouville equation*. We have another parallel between quantum mechanics and classical mechanics:

- QM: Liouville-von Neumann equation \leftrightarrow CL: Liouville equation

One more item is needed before we are fully armed for classical statistical mechanics. A function $G(p, q)$ that does not depend explicitly on time is a constant of the motion if its value does not change along the trajectory of the system in the phase space, in other words if

$$\frac{d}{dt} G(p(t), q(t)) = \left(\dot{p}_i \frac{\partial G}{\partial p_i} + \dot{q}_i \frac{\partial G}{\partial q_i} \right) = \left(\frac{\partial G}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial G}{\partial p_i} \right) = [G, \mathcal{H}] = 0, \quad (7.19)$$

where the last form stands for the Poisson bracket. In particular, an arbitrary function of the Hamiltonian $f(\mathcal{H}(p, q))$ is a constant of the motion. Moreover, we see from the Liouville equation that the phase space density is stationary,

$$\frac{\partial}{\partial t} f(p, q, t) = 0, \quad (7.20)$$

and so of the form $f(p, q)$ if and only if $[f, \mathcal{H}] = 0$. A stationary solution to the Liouville equation $f(p, q)$ must therefore be a constant of the motion. If now the stationary phase space density is positive, its logarithm exists and is also a constant of the motion. We have a minor theorem: Positive stationary phase space densities are precisely of the form

$$f(p, q) = e^{G(p, q)}, \quad (7.21)$$

where G is a constant of the motion.

7.2 Going to the classical limit

Let us consider a particle, massive or not, moving in one spatial dimension using periodic boundary conditions over a distance L , which, as usual, will eventually be taken to be large. The momentum operator

$$\hat{p} = -i\hbar \frac{\partial}{\partial x} \quad (7.22)$$

will, of course have the quantized eigenvalues of the form

$$p_k = \frac{2\pi\hbar}{L} k, \quad k = 0, \pm 1, \dots \quad (7.23)$$

One may associate to each momentum eigenstate a volume in p space equal to $V_p = 2\pi\hbar/L$. On the other hand, in the real x space the volume associated with the state, which is a homogeneous plane wave, is $V_x = L$. We associate for each eigenstate of momentum a volume in the phase space equal to $\mathcal{V} = V_p V_x = 2\pi\hbar = h$.

Given N particles in three dimensions, one may think of the motion of each particle in each spatial direction as a separate degree of freedom. The dimension of the phase space is now $2d = 6N$ instead of the 2 of our initial example, and the volume per eigenstate of momentum for all the N particles in the phase space is $\mathcal{V} = h^{3N}$. If we are to somehow replace the sum over all quantum states, eigenstates of momentum, with an integral over the phase space, it would have to be done according to the skeleton rule

$$\sum_{\text{quantum states}} \rightarrow \frac{1}{h^d} \int d^d p d^d q \quad (7.24)$$

to ensure that we count one state per each volume h^d in the phase space for d degrees of freedom, i.e., with d position and momentum coordinates.

Obviously, for N particles in three dimensions, $d = 3N$, but we are not yet done. Namely, classically, we would be counting separately all states for which the labels of the particles are exchanged, but quantum mechanics does not permit such a distinction for indistinguishable particles. All of the $N!$ classical states in which the particle labels are permuted should represent the same quantum state. Correcting such overcounting of the state, we have the fundamental rule for taking the classical limit for N indistinguishable particles:

$$\sum_{\text{quantum states}} \rightarrow \frac{1}{N!h^{3N}} \int d^{3N}p d^{3N}q \equiv \int d\tau. \quad (7.25)$$

In the last form we have simply changed the name of the integration measure.

This argument basically says two things: The phase space volume per quantum state for each degree of freedom is h , and all configurations in which the particle labels are permuted represent the same quantum state. Clearly, if there is any proof in here at all, it is less than rigorous. However, a more careful analysis and experience has told us that the rule (7.25) and the results we are going to extract from it are correct.

7.2.1 Classical statistical mechanics

In quantum mechanics, in the canonical ensemble, the expectation value of any observable A is calculated according to

$$\langle A \rangle = \text{Tr}(\rho A) = \frac{1}{Z} \text{Tr}(e^{-\beta H} A), \quad (7.26)$$

which involves sums over quantum states. To transfer to classical statistical mechanics, we have to decide what is the classical counterpart of the quantum Hamiltonian and of the observable A . The quantum Hamiltonian obvious is represented classically by the classical Hamiltonian, and frequently the classical counterpart of the operator A is an equally obvious function of the phase space variable p and q . We then have the classical counterpart of the quantum average in (7.26)

$$\langle A \rangle = \int d\tau \frac{e^{-\beta \mathcal{H}(p,q)}}{Z} A(p,q), \quad Z = \int d\tau e^{-\beta \mathcal{H}(p,q)}. \quad (7.27)$$

Comparison with (7.6) shows that we can identify the N particle phase space density

$$f(p,q) = \frac{e^{-\beta \mathcal{H}(p,q)}}{Z}, \quad Z = \int d\tau e^{-\beta \mathcal{H}(p,q)}. \quad (7.28)$$

The phase space density is an exponential of basically the Hamiltonian, a constant of the motion, and so is stationary, as it should be in thermal equilibrium. While these expressions use the $d\tau$ integration measure, in fact it does not matter in the discussions of the phase space density whether one uses $d\tau$ or $dp dq$, as long as the usage is consistent.

In principle the problem of classical statistical mechanics is completely solved when the phase space density is laid down. In practice, however, one again seeks a direct connection to thermodynamics.

Connection to thermodynamics: Ideal gas

The classical counterpart of the canonical partition function obviously reads

$$Z = \sum_i e^{\beta E_i} \rightarrow Z = \frac{1}{N! h^{3N}} \int dp dq e^{-\beta \mathcal{H}(p,q)}. \quad (7.29)$$

Here the particular integration measure needs to be used, otherwise the counting of the quantum states is not done correctly for the classical limit. The canonical partition function obtained in this way is supposedly the classical approximation (e.g., high-temperature limit) of the quantum mechanical partition function. Inasmuch as the classical limit in itself is justified, the connection to thermodynamics is, of course, exactly the same as in QSM:

$$F = -kT \ln Z. \quad (7.30)$$

We discuss a free ideal gas as an example. “Free” here means that there is no external binding potential energy, so that the Hamiltonian reads

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m}. \quad (7.31)$$

The partition function is therefore

$$Z = \frac{1}{N! h^{3N}} \int d^{3N}p d^{3N}q e^{-\beta \sum_i \frac{\mathbf{p}_i^2}{2m}}. \quad (7.32)$$

Nothing inside the integral depends on the position coordinates, so the position integral gives

$$\int d^{3N}q = \left[\int_0^L dq \right]^{3N} = L^{3N} = V^N. \quad (7.33)$$

Although it is not explicitly indicated in the notation, the integral over the positions is limited to a $3N$ dimensional hypercube with the side length L . After all, the original argument about phase space volume per quantum state was done in a certain quantization volume. To develop the momentum integrals, we first take separately the $3N$ Cartesian coordinates p_j of the N momenta \mathbf{p}_i , and proceed as follows,

$$\begin{aligned} \int d^{3N}p e^{-\beta \sum_i \frac{\mathbf{p}_i^2}{2m}} &= \int d^{3N}p e^{-\beta \sum_j \frac{p_j^2}{2m}} = \int d^{3N}p \prod_j e^{-\frac{p_j^2}{2mkT}} \\ &= \int dp_1 e^{-\frac{p_1^2}{2mkT}} \int dp_2 e^{-\frac{p_2^2}{2mkT}} \dots \int dp_{3N} e^{-\frac{p_{3N}^2}{2mkT}} \\ &= \left[\int_{-\infty}^{\infty} dp e^{-\frac{p^2}{2mkT}} \right]^{3N} \\ &= (2\pi mkT)^{3N/2}. \end{aligned} \quad (7.34)$$

The partition function therefore is

$$Z = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N, \quad \lambda = \sqrt{\frac{2\pi\hbar^2}{mkT}}, \quad (7.35)$$

which again shows the familiar thermal de Broglie wavelength.

Helmholtz free energy reads in the Stirling approximation

$$F = -kT \ln Z \simeq -NkT \left[\ln \frac{V}{\lambda^3} - \ln N + 1 \right]. \quad (7.36)$$

We may find pressure as

$$p = - \left(\frac{\partial F}{\partial V} \right)_{N,T} = \frac{NkT}{V}, \quad pV = NkT. \quad (7.37)$$

Boyle's law pops out as if by magic. As another example, the entropy is

$$S = - \left(\frac{\partial F}{\partial S} \right)_{V,T} = Nk \left[\ln \frac{1}{n\lambda^3} + \frac{5}{2} \right], \quad (7.38)$$

complete with the correct additive constant for a monatomic ideal gas. Namely, the constant can in principle be determined experimentally, and in some cases actually has been determined. One basically starts at a very low temperature, whereupon $S = 0$, and tracks both the heat put in and the temperature as the substance is converted all the way to a nearly ideal gas. The same expression can also be found, for instance, in the high-temperature limit of the entropies of both the Bose-Einstein and the Fermi-Dirac ideal gas.

7.2.2 Few-particle distribution functions

A *one-particle observable* is a straight sum of quantities referring to individual particles:

$$A(p, q) = \sum_{i=1}^N a(\mathbf{r}_i, \mathbf{p}_i) = \sum_i a(z_i), \quad (7.39)$$

with the notation $z_i = (\mathbf{r}_i, \mathbf{p}_i)$. Examples include total kinetic energy and total momentum of a particle system. The expectation value of such an observable for the given phase space density is

$$\begin{aligned} \langle A \rangle &= \int dz_1 \dots dz_N f(z_1, \dots, z_N) a(z_1) + \dots + \int dz_1 \dots dz_N f(z_1, \dots, z_N) a(z_N) \\ &= \int dz_1 \dots dz_N f(z_1, z_2, \dots, z_N) a(z_1) + \int dz_1 \dots dz_N f(z_2, z_1, \dots, z_N) a(z_2) \\ &\quad + \dots + \int dz_1 \dots dz_N f(z_N, z_2, \dots, z_1) a(z_N) \\ &= \int dz_1 \dots dz_N f(z_1, z_2, \dots, z_N) a(z_1) + \int dz_1 \dots dz_N f(z_1, z_2, \dots, z_N) a(z_1) \end{aligned}$$

$$\begin{aligned}
& + \dots + \int dz_1 \dots dz_N f(z_1, z_2, \dots, z_N) a(z_1) \\
& = \int dz_1 f_1(z_1) a(z_1),
\end{aligned} \tag{7.40}$$

where the *one-particle distribution function* is basically obtained by integrating away the coordinates and momenta of all but one particle from the phase space density,

$$f_1(z_1) = N \int dz_2 \dots dz_N f(z_1, z_2, \dots, z_N). \tag{7.41}$$

In (7.40) the first equality is because we are dealing with identical particles. Quantum mechanically, one cannot distinguish states that differ only in the identity, label, of the particles. Correspondingly, classically, the phase space density must be invariant under the permutations of the particle labels. The second equality is obtained by renaming the integration variables.

In short, to calculate the expectation value of any one-particle observable one does not have to know the entire phase space density, but only the one-particle distribution function. The one-particle distribution function has other uses as well. As we have defined it above, it is normalized to particle number,

$$\int d^3r d^3p f_1(\mathbf{r}, \mathbf{p}) = N \int dz_1 \dots dz_N f(z_1, \dots, z_N) = N. \tag{7.42}$$

Rearranging, we have the equality

$$\int d^3r \left(\int d^3p f_1(\mathbf{r}, \mathbf{p}) \right) = N. \tag{7.43}$$

But this expresses the particle number as an integral of a positive function. Obviously, the positive function is the particle density,

$$n(\mathbf{r}) = \int d^3p f_1(\mathbf{r}, \mathbf{p}). \tag{7.44}$$

Likewise,

$$f(\mathbf{p}) = \frac{1}{N} \int d^3r f_1(\mathbf{r}, \mathbf{p}) \tag{7.45}$$

is the momentum distribution, normalized to one.

Normalization is the sore point with distribution functions of all ilk. When in doubt, normalize explicitly! Here we use a unit-normalized one-particle distribution to demonstrate a useful property of the phase space density of an ideal gas, and not necessarily a free gas; the Hamiltonian reads

$$\mathcal{H}(p, q) = \sum_i^N h(z_i). \tag{7.46}$$

Phase space density normalized to unity with respect to the $dp dq$ measure reads

$$f(z_1, \dots, z_n) = \frac{e^{-\beta h(z_1)} \dots e^{-\beta h(z_N)}}{[\int dz e^{\beta h(z)}]^N}, \tag{7.47}$$

which gives the one-particle distribution function, normalized to unity, as

$$f_1(z_1) = \int dz_2 \dots dz_N f(z_1, \dots, z_N) = \frac{e^{-\beta h(z_1)}}{\int dz e^{-\beta h(z)}}. \quad (7.48)$$

This is the same as the phase space density for a system with exactly one particle. Moreover, the phase-space density for N particles satisfies

$$f_N(z_1, z_2, \dots, z_N) = f_1(z_1)f_1(z_2) \dots f_1(z_N). \quad (7.49)$$

Factorization of probability density tells us that the particles are statistically independent. This is as it should be, given that we are dealing with an ideal gas of noninteracting particles.

To be very very specific, let us write the one-particle distribution for a free ideal gas. Normalized to unity, it reads

$$f_1(\mathbf{r}, \mathbf{p}) = \frac{e^{-\frac{\mathbf{p}^2}{2mkT}}}{\int d^3r d^3p e^{-\frac{\mathbf{p}^2}{2mkT}}} = \frac{e^{-\frac{\mathbf{p}^2}{2mkT}}}{V(2\pi mkT)^{3/2}}. \quad (7.50)$$

To corresponding one-particle distribution normalized to the number of particles is

$$f(\mathbf{r}, \mathbf{p}) = \frac{n}{(2\pi mkT)^{3/2}} e^{-\frac{\mathbf{p}^2}{2mkT}}, \quad (7.51)$$

where n refers to the uniform density of the gas. This is a standard way of writing the Maxwell-Boltzmann distribution.

One frequently also sees *two-particle* observable, such as the interaction potential energy, which are of the form

$$V(p, q) = \frac{1}{2} \sum_{i \neq j} v(z_i, z_j). \quad (7.52)$$

Without going into the details, expectation values of two-particle observables may be expressed in terms of the *two-particle distribution function*. A unit-normalized version reads

$$f_2(z_1, z_2) = \int dz_3 dz_4 \dots dz_N f(z_1, z_2, z_3, \dots, z_N). \quad (7.53)$$

The expectation value is then

$$\langle V \rangle = \frac{N(N-1)}{2} \int dz_1 dz_2 f_2(z_1, z_2) v(z_1, z_2). \quad (7.54)$$

Higher multiparticle distributions are possible, but are rarely encountered in practice.

7.3 Kinetic theory

Thermodynamics and statistical mechanics describe the equilibrium of a system. The present chapter is about evolution of nonequilibrium states in many-body systems. We briefly discuss *kinetic theory*. There are quantum mechanical versions of kinetic theory, but here we only take up elementary classical topics. Evolution of a nonequilibrium many-body system is obviously a difficult subject, even classically, and kinetic theory more often than not deals with near-equilibrium situations in near-ideal gases.

7.3.1 Boltzmann equation

To get started, let us consider the evolution of the phase space density in a special form of ideal gas,

$$\left(\frac{\partial}{\partial t} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial}{\partial p_i} \right) f(p, q, t) = 0 \quad (7.55)$$

with

$$\mathcal{H} = \sum_{i=1}^{3N} h(p_i, q_i). \quad (7.56)$$

A normal ideal gas where the sum is particle by particle, not by degrees of freedom, would do as well, we just make the example as simple as possible.

Let us integrate away one pair of canonical coordinates, call it (p_1, q_1) from the Liouville equation. The ensuing development runs thus,

$$\begin{aligned} 0 &= \int \left(\frac{\partial}{\partial t} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial}{\partial p_i} \right) f(p, q, t) dp_1 dq_1 \\ &= \left[\frac{\partial}{\partial t} + \sum_{i \neq 1} \left(\frac{\partial h(p_i, q_i)}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial h(p_i, q_i)}{\partial q_i} \frac{\partial}{\partial p_i} \right) \int f(p, q, t) dp_1 dq_1 \right] \\ &\quad + \int dp_1 dq_1 \left(\frac{\partial h(q_1, p_1)}{\partial p_1} \frac{\partial}{\partial q_1} - \frac{\partial h(p_1, q_1)}{\partial q_1} \frac{\partial}{\partial p_1} \right) f(p, q, t). \end{aligned} \quad (7.57)$$

But the last line vanishes by partial integration. The net result is that we have integrated away the canonical coordinate pair (p_1, q_1) from the Liouville equation. Obviously this process may be repeated until only the coordinates and momenta of one single particle are left, whereupon we get an equation of motion for the one-particle distribution $f(\mathbf{r}, \mathbf{p}) = f_1(z)$. As appropriate for Cartesian coordinates, we denote

$$\frac{\partial h(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \dot{\mathbf{r}} = \mathbf{v}, \quad \frac{\partial h(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} = \frac{\partial \mathcal{H}}{\partial \mathbf{r}} = -\dot{\mathbf{p}} = -\mathbf{F}. \quad (7.58)$$

The result is

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial}{\partial \mathbf{p}} \right) f(\mathbf{r}, \mathbf{p}, t) = 0. \quad (7.59)$$

In an ideal gas the one-particle distribution evolves as if only one particle were present. The convention in kinetic theory is to normalize the one-particle distribution to the total number of particles. The derivative on the left-hand side may be identified as the total rate of change of the distribution function in a frame that moves with the local flow of the particles. Correspondingly the terms *convective derivative* and *flow derivative* are used for this combination of derivatives.

When the gas is not ideal, in principle we could still write

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial}{\partial \mathbf{p}} \right) f(\mathbf{r}, \mathbf{p}, t) = \left(\frac{\partial f}{\partial t} \right)_C. \quad (7.60)$$

In practice it would usually be impossible to obtain an exact expression for the *collision derivative* on the right-hand side. The standard approximation strategy, which may be applied in a variety of ways, is based on the assumption that any “spectator” particle collides with the other particles, but the presence of the spectator particle does not perturb the flux of the incoming colliders. This is called the assumption of molecular chaos, or in German, Stosszahlansatz. Obviously, the assumption works best in a dilute gas, when the molecules do not spend a large fraction of their time within the range of their molecule-molecule interactions. The end result is that the collision derivative gets expressed in terms of the one-particle distribution function f , which gives a closed equation for f .

Although a quite simple derivation would be possible, we do not go into the details but only state the resulting *Boltzmann equation*, the main workhorse in kinetic theory. To begin with, in a two-body elastic collision the momenta of the incoming particles \mathbf{p}_1 and \mathbf{p}_2 and two angles Ω in the center-of-mass frame of the colliding particles completely determine the momenta of the outgoing particles \mathbf{p}'_1 and \mathbf{p}'_2 . Besides, there is also a differential cross section for the collision, denote it by $\frac{\partial \sigma}{\partial \Omega}$, that governs the probability distribution of the outgoing particles, given a uniform flux of incoming particles. Further, denoting

$$f(\mathbf{r}, \mathbf{p}_1, t) \equiv f_1, \quad f(\mathbf{r}, \mathbf{p}'_1, t) \equiv f'_1, \quad (7.61)$$

and so on, the collision derivative in the Boltzmann equation reads

$$\left(\frac{\partial f}{\partial t} \right)_C = \int d^3 p_2 |\mathbf{v}_1 - \mathbf{v}_2| \frac{\partial \sigma}{\partial \Omega} (f'_1 f'_2 - f_1 f_2). \quad (7.62)$$

The collision derivative is nonlinear in the distribution function f , which elevates the complications coming with the general case way above the level of our present discussion. Two results arising basically from the symmetry properties of the collisions, including time reversal symmetry, bear a mention, however. First, defining what is called Boltzmann’s H function,

$$H(t) = \int d^3 r d^3 p f(\mathbf{r}, \mathbf{p}, t) \ln f(\mathbf{r}, \mathbf{p}, t), \quad (7.63)$$

it may be shown that under the Boltzmann equation

$$\frac{dH}{dt} \leq 0. \quad (7.64)$$

This is called Boltzmann's H theorem. The similarities to von Neumann entropy and to the second law of thermodynamics are inviting.

Second, basically as a corollary of the proof of the H theorem, it may be shown that any distribution functions $f(\mathbf{r}, \mathbf{p}, t)$ that gives a zero collision derivative satisfies

$$f_1 f_2 = f'_1 f'_2. \quad (7.65)$$

for all momentum arguments that represent allowed collisions. This means that the distribution functions are of the form

$$f = e^{g(\mathbf{p})}, \quad (7.66)$$

where $g(\mathbf{p})$ is a "locally conserved quantity" that satisfies the equation

$$g(\mathbf{p}_1) + g(\mathbf{p}_2) = g(\mathbf{p}'_1) + g(\mathbf{p}'_2). \quad (7.67)$$

for all allowed binary collision. The most general such $g(\mathbf{p})$ is a linear combination of a constant, the three components of the momentum, and the kinetic energy, i.e.,

$$g(\mathbf{p}) = A + \mathbf{B} \cdot \mathbf{p} + C \frac{\mathbf{p}^2}{2m}. \quad (7.68)$$

The collision derivative is local in position, so that the "constants" A , \mathbf{B} and C may depend on the position \mathbf{r} .

As a tangible example, suppose that the force acting on the particles \mathbf{F} is conservative and may be derived from a potential $V(\mathbf{r})$ as $\mathbf{F} = -\nabla V$. Then a function of the form

$$f(\mathbf{r}, \mathbf{p}) = K e^{-\beta[\frac{\mathbf{p}^2}{2m} + V(\mathbf{r})]} \quad (7.69)$$

for any constant β is of the form (7.66), so that it zeros the collision derivative. Moreover it also zeros the convective derivative, so that it satisfies the Boltzmann equation as a whole. It simply remains to make the connection between the constant β and temperature from some, say, thermodynamics argument. On the other hand, for a free gas with a constant potential energy equal to, say, zero, a distribution of the form

$$f(\mathbf{r}, \mathbf{p}) = K e^{\frac{(\mathbf{p} - m\mathbf{u})^2}{2mkT}} \quad (7.70)$$

is a solution to Boltzmann's equation for any constants K and \mathbf{u} . This obviously represents a gas flowing at the speed \mathbf{u} . In fact, in free space the difference between a flowing and stationary gas is simply a choice of the Galilean reference frame.

Interesting as this kind of observations about the stationary solution might be, we do not bother with detailed proofs. The reason is that all equilibrium

results that could be obtained in this manner may just as well be obtained from statistical mechanics. There is a philosophical point here, though, that is worth remembering. Namely, in a dilute gas (otherwise Stosszahlansatz would not be valid) particle-particle interactions support the precisely same equilibrium as is predicted from statistical mechanics of an ideal gas. In this sense, the statistical-mechanics one-particle distribution function is valid even to the leading order in particle-particle interactions. Besides, it turns out that not only does the Boltzmann equation maintain the one-particle distribution function as per the ideal gas, the time evolution even tends to drive the one-particle distribution toward the thermal form if it starts out of equilibrium.

Linearized Boltzmann equation and relaxation-time approximation

The nonlinearity of the Boltzmann equation may be avoided in principle if the one-particle distribution function is “close” to the stationary distribution. We write

$$f = \tilde{f} + g, \quad (7.71)$$

where $\tilde{f}(\mathbf{r}, \mathbf{p})$ is stationary, and $g(\mathbf{r}, \mathbf{p}, t)$ is “small.” Expansion to the first order in g gives

$$f'_1 f'_2 - f_1 f_2 \simeq \tilde{f}'_2 g'_1 + \tilde{f}'_1 g'_2 - \tilde{f}_2 g_1 - \tilde{f}_1 g_2; \quad (7.72)$$

the leading term $\tilde{f}'_1 \tilde{f}'_2 - \tilde{f}_1 \tilde{f}_2$ will not contribute because \tilde{f} is a stationary distribution. The Boltzmann equation then reads

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial}{\partial \mathbf{p}} \right) g = \int d^3 p_2 |\mathbf{v}_1 - \mathbf{v}_2| \frac{\partial \sigma}{\partial \Omega} (\tilde{f}'_2 g'_1 + \tilde{f}'_1 g'_2 - \tilde{f}_2 g_1 - \tilde{f}_1 g_2), \quad (7.73)$$

which is linear. The theory of linear systems may be brought to bear.

However, even the linearized version of the Boltzmann equation is still cumbersome to deal with. Here we go further, and introduce the relaxation time approximation according to which the collision derivative is

$$\left(\frac{\partial f}{\partial t} \right) = -\frac{f - f_0}{\tau}. \quad (7.74)$$

Here f_0 is some physically motivated ansatz appropriate for the problem, and τ is the relaxation time. Namely, suppose both f and f_0 are homogeneous in the position variables, that there is no external force, and that f_0 is time independent, then the Boltzmann equation in the relaxation time approximation and its solution are

$$\frac{\partial f}{\partial t} = -\frac{f - f_0}{\tau}; \quad f(t) = f_0 + [f(t=0) - f_0]e^{-t/\tau}. \quad (7.75)$$

This indeed relaxes to f_0 over a time scale τ .

Suppose now that the solution to the Boltzmann equation remains “close” to the ansatz f_0 , so that $f = f_0 + g$ and g is “small.” Within the relaxation

time approximation we have

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial}{\partial \mathbf{p}} \right) (f_0 + g) = -\frac{g}{\tau}. \quad (7.76)$$

In order that our assumption of a “small” g be consistent, the differential operator on the left must produce a “small” result when acting on f_0 . So, the differential operator in itself denotes a “small” quantity, and, when acting on g , it should give a second-order “small” quantity. Upon this reasoning, we drop the g on the left-hand side, and have an immediate solution for the correction g to the ansatz f_0 ,

$$g = -\tau \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial}{\partial \mathbf{p}} \right) f_0. \quad (7.77)$$

Whether this procedure may work obviously depends on an intelligent choice of the ansatz f_0 .

As an example, let us consider a dilute gas of charged (q) molecules under a small constant (in space and time) electric field \mathbf{E} . We know that according to Ohm’s law we should have the current density \mathbf{j} determined by the conductivity of the medium σ as $\mathbf{j} = \sigma \mathbf{E}$. We want to find the conductivity in the relaxation time approximation.

The zeroth-order ansatz here obviously is the homogeneous Maxwell-Boltzmann gas as described by

$$f_0 = \frac{n}{(2\pi mkT)^{3/2}} e^{-\frac{\mathbf{p}^2}{2mkT}}. \quad (7.78)$$

The supposedly small correction owing to the electric field arises from the electrostatic force $\mathbf{F} = q\mathbf{E}$, and reads

$$g = -\tau \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial}{\partial \mathbf{p}} \right) f_0 = \frac{\tau q}{mkT} \mathbf{E} \cdot \mathbf{p} f_0 = \frac{\tau q E}{mkT} p_1 f_0, \quad (7.79)$$

where we have denoted the fixed direction of the electric field by 1. Obviously the electric current density will be in the direction 1 as well, and, as usual, equals charge times density times drift velocity:

$$\begin{aligned} j_1 &= q \left(\int d^3p f \right) \langle v_1 \rangle = q \left(\int d^3p f \right) \frac{\int d^3r d^3p f \frac{p_1}{m}}{\int d^3r d^3p f} = q \left(\int d^3p f \right) \frac{\int d^3p f \frac{p_1}{m}}{\int d^3p f} \\ &= q \int d^3p (f_0 + g) \frac{p_1}{m} = q \int d^3p g \frac{p_1}{m} = \frac{q^2 \tau E}{mkT} \int d^3p f_0 \frac{p_1^2}{m} = \frac{n_0 q^2 \tau}{m} E. \end{aligned} \quad (7.80)$$

The first step expresses density in terms of the one-particle distribution function, in the second step we calculate the average velocity and normalize the distribution just to make sure, in the third step we have noticed that the integrand in fact does not depend on position at all, in the fourth step we have canceled the normalization integral, in the fifth step we note that $f_0 p_1$ is an odd

function of the integration variable and does not contribute, and finally we have completed a Gaussian integral. The conductivity is

$$\sigma = \frac{n_0 q^2 \tau}{m}. \quad (7.81)$$

The correction g here is proportional to the electric field strength, and the convective derivative acting on g would give something proportional to the square of the electric field strength. We have therefore verified, after the fact, that the procedure as in (7.77) is correct to the leading order in the strength of the electric field.

7.4 Exercises

7.1.

(a) Show that

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2 + \beta x} = \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{4\alpha}}$$

for $\alpha > 0$

(b) Show that the same holds for complex α and β as long as $\Re(\alpha) > 0$.

7.2. Show that for noninteracting particles the thermal N -body density function is simply the straight product of one-particle density functions, i.e.,

$$f(z_1, z_2, \dots, z_N) = f_1(z_1) f_1(z_2) \dots f_1(z_N).$$

7.3. Given a variable x in a system and a function $f(x)$ such that (i) $f(x) \geq 0$, (ii) $\int f(x) dx = 1$, and that (iii) the expectation value of an arbitrary function of the variable x , $g(x)$, may be expressed as $\langle g \rangle = \int dx f(x) g(x)$, $f(x)$ obviously is the probability density of the variable x . Find the probability distribution of kinetic energy in the Maxwell-Boltzmann (classical) free (no confining potential) ideal gas.

7.4. The *two-particle* density function is defined as

$$f_2(z_1, z_2) = \int dz_3 \dots dz_N f(z_1, z_2, z_3, \dots, z_N).$$

Quite analogously to the one-particle density function, this can be used to calculate the expectation values of two-particle operators that refer to pairs of particles.

(a) Find the joint distribution function of two momenta in a free classical ideal gas at the temperature T .

(b) Given the two momenta \mathbf{p}_1 and \mathbf{p}_2 , one can define the total and relative momenta as $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$ and $\mathbf{p} = \frac{1}{2}(\mathbf{p}_1 - \mathbf{p}_2)$. In collisions, it is the relative momenta that matter. What is the distribution of relative momenta?

7.5. Let us try our hand on the classical limit in the microcanonical ensemble.

(a) Take any function $f(x)$ of the absolute value $x = \sqrt{x_1^2 + \dots + x_n^2}$ of the vector $\mathbf{x} = (x_1, \dots, x_n)$ in n -dimensional space. For a suitable constant C_n we obviously have the spherical integral in n dimensions like this:

$$\int d^n x f(x) = C_n \int_0^\infty dx x^{n-1} f(x).$$

Find the constant C_n by studying the particular function $f(x) = e^{-x^2}$.

(b) Show that the volume of a n -dimensional sphere of radius R is $\Omega(R) = \pi^{n/2} R^n / \Gamma(n/2 + 1)$. Here $\Gamma(x)$ is the gamma function, in effect a continuous-variable version of the factorial; for positive-integer arguments, $\Gamma(n + 1) = n!$.

(c) Consider energy eigenstates for N identical particles in an isotropic harmonic oscillator potential with the frequency ω using classical-limit arguments. Show that the number of many-body energy eigenstates with the energy less than E equals

$$\Sigma(E) = \frac{1}{N!(3N)!} \left(\frac{E}{\hbar\omega} \right)^{3N}.$$

This is the expression given in P. 5.2. The formula in P. 5.2 is of similar origin.

7.6. Assuming that the Earth is flat (good) and the atmosphere is a constant-temperature classical ideal gas of molecules with external (center-of-mass) degrees of freedom only (bad), find the heat capacity of the atmosphere per molecule.

7.7. Consider N atoms trapped in an isotropic harmonic oscillator potential characterized by the angular frequency ω , so that the potential energy is

$$V(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{2} m \omega^2 \sum_{i=1}^N \mathbf{r}_i^2.$$

Find the heat capacity in the classical limit. The result is useful when $kT \gg \hbar\omega$.

7.8. Let us study, again N noninteracting atoms trapped in an isotropic harmonic oscillator potential characterized by the angular frequency ω . What is the density of the gas at the temperature T according to the Maxwell-Boltzmann statistics?

7.9.

(a) Find the chemical potentials of two- and three-dimensional ideal gases, as functions of temperature and area density or density as appropriate, in the classical limit of statistical mechanics.

(b) Suppose atoms adsorbed at a surface may be regarded as a two-dimensional ideal gas, and also assume that the energy of an atom is lowered by some amount ϵ when it gets adsorbed to the surface. Given the density of the gas outside of the surface ρ , what's the area density of the adsorbed atoms n ?

7.10.

(a) Consider a simple chemical reaction $A + B \leftrightarrow AB$. As the double arrow indicates, all chemical reactions go both ways. Show that in a mixture of reactants the condition for chemical equilibrium is expressed in terms of chemical potentials as $\mu_A + \mu_B = \mu_{AB}$.

(b) Show that in an ideal gas where the only relevant degree of freedom is the center-of-mass motion of the atoms or molecules, the chemical potential in the classical (high temperature/low density) limit is $\mu = kT \ln(n\lambda^3)$, where $\lambda = (2\pi\hbar^2/mkT)^{1/2}$ is the usual thermal de Broglie wavelength.

(b) Consider the formation of diatomic molecules A_2 out of atoms A . Assume that the binding energy of the molecule is I , i.e., the difference in internal energy between a molecule and a pair of atoms is $-I$. Show that in the limit when the atoms and molecules may be regarded as classical ideal gases, except for the formation of the molecules, the equilibrium densities of the atoms and molecules satisfy $n_{A_2}/n_A^2 = \sqrt{8}\lambda_A^3 e^{\frac{I}{kT}}$. The chemical-equilibrium densities of atoms and molecules depend manifestly on the de Broglie wavelength, i.e., on quantum mechanics, even at room temperature!

.

7.11. In a possible formulation of statistical mechanics the equilibrium density operator $\hat{\rho}$ is the one that maximizes the entropy $S = -k \text{Tr}(\hat{\rho} \ln \hat{\rho})$ under the constraints of the ensemble. The usual constraint for the canonical ensemble is that the total energy $E = \langle \hat{H} \rangle$ is fixed. However, if it so happens that the system is (or may be regarded as) translationally invariant, total momentum $\hat{\mathbf{P}} = \sum_i \hat{\mathbf{p}}_i$ is also a constant of the motion. The density operator of the canonical ensemble, in a generalization of the usual form, is then

$$\rho = \frac{1}{Z} e^{-\beta(\hat{H} - \mathbf{v} \cdot \hat{\mathbf{P}})}.$$

The vector constant \mathbf{v} is obviously the overall flow velocity of the gas.

(a) Take a near-ideal gas with atoms of mass m at temperature T . Show that in the classical limit the chemical potential is

$$\mu = -\frac{m\mathbf{v}^2}{2} + kT \ln(n\lambda^3),$$

where λ is the usual thermal de Broglie wavelength.

(b) Consider now a pipe with gas flowing at the speed v connected by a tight channel to a chamber containing the same gas, all of this at the temperature T . Given that the density of the gas in the flow is n , what is the density in the chamber?

N.B. This problem may evoke all sorts of tricky fundamental questions; bypass them.

7.12. (a) Consider in the classical limit of statistical mechanics two ideal gases such that the atoms are indistinguishable within each gas, but different between the gases. Initially the gases of N_1 and N_2 atoms are held in volumes V_1 and V_2 such that the densities and temperatures are equal for both gases, then the gases are allowed to mix within the total volume $V_1 + V_2$. While the temperature and density obviously remain unchanged, the total entropy increases by the amount

$$\Delta S = -k \left(N_1 \ln \frac{N_1}{N_1 + N_2} + N_2 \ln \frac{N_2}{N_1 + N_2} \right),$$

the entropy of mixing. Demonstrate this.

Mixing entropy is a general phenomenon that occurs in any mixture. Moreover, when the mixture is sufficiently dilute, mixing entropy is the dominant nonadditive effect of the mixing on the free energy of the mixture.

(b) Consider the case of a pure solvent and the same solvent with some solute added, on opposite sides of a semipermeable membrane that passes solvent molecules but not the molecules of the solute. Show in the limit of a dilute solution that, in order to maintain thermodynamic equilibrium, there must be a pressure increase, *osmotic pressure*, on the side of the solution that equals the pressure that the solute would have if it were an ideal gas with the given density and temperature.

Osmotic pressure can be substantial. In sea water it is about 30 atm. Thus, if you attempt to construct a device to purify sea water by “reverse osmosis”, by squeezing the water through a membrane that does not pass the salts, you must apply a pressure of more than 30 atm.

7.13. Let us consider a diatomic molecule under the assumption that the distance between the atoms is (somewhat miraculously) isolated from all other degrees of freedom. There is a potential energy $V(\xi)$ to the deviation ξ of the distance between the atoms from the zero-temperature equilibrium that is quadratic to the leading order, but here we take into account also the cubic correction valid near the equilibrium and write

$$V(\xi) \simeq K\xi^2 - a\xi^3, \quad (7.82)$$

where $K > 0$ is the force constant in the harmonic approximation and $a > 0$ is “small.” Calculate from classical statistical mechanics how much the molecule stretches as a function of temperature from its zero-temperature equilibrium configuration. This is also basically how thermal expansion works in solids.

7.14. Study a trap for atoms. We assume that the gas is effectively homogeneous along the trap axis, so that the direction along the trap axis may be separated out and we have a two-dimensional situation. The trapping potential is the rotationally symmetric $V(\mathbf{x}) = \frac{1}{2}m\omega^2(x_1^2 + x_2^2)$. Let us also introduce the 3 component of the angular momentum $L_3 = x_1p_2 - x_2p_1$, and write

$$f(\mathbf{x}, \mathbf{p}) = K \exp \left[-\frac{1}{kT} \left(\frac{\mathbf{p}^2}{2m} + \frac{m\omega^2 \mathbf{x}^2}{2} - \Omega L_3 \right) \right]$$

with some constant Ω .

(a) Show that the convective derivative acting on $f(\mathbf{x}, \mathbf{p})$ gives zero. Since everything inside the exponential is conserved locally in a collision, $f(\mathbf{x}, \mathbf{p})$ is therefore also a stationary solution to the Boltzmann equation.

(b) Find the local flow velocity

$$\mathbf{u}(\mathbf{x}) = \frac{\int d^2p f(\mathbf{x}, \mathbf{p})(\mathbf{p}/m)}{\int d^2p f(\mathbf{x}, \mathbf{p})},$$

and explain what the constant Ω means.

(b) How does the root-mean-square radius $r = \sqrt{\langle \mathbf{x}^2 \rangle}$ of the gas behave with the parameter Ω ? Please explain.

7.15. A centrifuge is basically a cylinder with some radius R that is being rotated at the angular velocity ω about its axis. When churned long enough, normal (not superfluid) matter inside settles to thermal equilibrium in a frame rotating with the cylinder. On the other hand, transformation to a rotating (hence, accelerated) frame adds a term $-\boldsymbol{\omega} \cdot \mathbf{L}$ to the Hamiltonian, where $\boldsymbol{\omega}$ is the angular velocity vector and \mathbf{L} is the angular momentum with respect to a point on the axis of rotation. Based on these observations, find the equilibrium density at temperature T for a nearly ideal gas of molecules with mass m inside the centrifuge, given that the density without the rotation would be n .

7.16. One way of modeling the collision derivative is to state that

$$\left(\frac{df}{dt} \right)_c = \gamma \frac{\partial}{\partial \mathbf{p}} \cdot (\mathbf{p}f) + D \frac{\partial^2 f}{\partial \mathbf{p}^2};$$

the odd second derivative stands for Laplacian. Nothing essential is lost by working in one dimension, so henceforth do so.

(a) Consider the Boltzmann equation in free space, in the absence of external forces. Find a position independent stationary solution to the ensuing simple *Fokker-Planck* equation.

(b) The result shows that you cannot choose any old damping and diffusion coefficients γ and D , but these coefficients must satisfy the *Einstein relation*. What is it?

7.17. Assume that a density gradient exists in a constant-temperature free gas. The system sets up a “particle current” such that in the limit of small density gradient the particle current density equals $\mathbf{j} = -D \nabla n$, where D is the *coefficient of diffusion*. Analyze the situation using the relaxation time approximation to the Boltzmann equation.

(a) Pause and carefully consider what you should write down for the distribution f_0 that is “almost” the solution to the Boltzmann equation.

(b) Find the coefficient of diffusion.

7.18. *Hall effect in an ionic conductor.* Consider an ionic conductor, say, a liquid with a low concentration of charged particles dispersed in it. Suppose that some external electric field sets up an electric current density j in the x direction. In addition, a magnetic field B in the z direction is applied to the liquid. Because the magnetic field bends the orbits of moving charges [n.b. $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$], there initially is a current in the $\pm y$ direction, until the charges accumulated on the surfaces of the liquid set up an electric field with the strength E that cancels the current.

- (a) Given the density n of the carriers with charge q , what is the flow velocity of the charges in the x direction?
- (b) Write down the thermal one-particle distribution function $f_0(\mathbf{r}, \mathbf{p})$ that incorporates the known current $j\hat{\mathbf{e}}_z$, but does not yet account for the effects of the magnetic field.
- (c) Using the Boltzmann equation in the relaxation time approximation, calculate the “Hall coefficient” $R = E/jB$.

7.19. Assume that a density gradient exists in a constant-temperature gas. Now, in addition to particle current, the system also sets up an energy current or “heat current”.

- (a) Pause and carefully consider what you should write down for the distribution f_0 that is “almost” the solution to the Boltzmann equation.
- (b) The heat current equals $\mathbf{q} = n\langle \mathbf{v} \frac{\mathbf{p}^2}{2m} \rangle$. Justify this!
- (c) If the density gradient is small enough, the heat current equals $\mathbf{q} = -K \nabla n$. Calculate the (unnamed) transport coefficient K in the relaxation time approximation.

7.20. For the sake of simplicity, consider the motion of a particle with mass m in one dimension under the potential energy $V(x)$ using wave function quantum mechanics. Let us define the Wigner function in terms of the wave function $\psi(x, t)$ as

$$W(x, p, t) = \frac{1}{2\pi\hbar} \int du e^{-i\frac{up}{\hbar}} \psi^*(x - \frac{1}{2}u, t) \psi(x + \frac{1}{2}u, t).$$

- (a) Show that the following holds true:

$$\int dp W(x, p, t) = |\psi(x, t)|^2.$$

Similarly, the integral over the x coordinate gives the absolute square of the momentum space wave function, $|\psi(p, t)|^2$.

(b) Starting from the time dependent Schrödinger equation, show that in the limit $\hbar \rightarrow 0$ the Wigner function obeys the equation of motion

$$\left(\frac{\partial}{\partial t} + \frac{p}{m} \frac{\partial}{\partial x} + F \frac{\partial}{\partial p} \right) W(x, p, t) = 0,$$

where $F = -\frac{\partial}{\partial x} V(x)$ is the force.

So, not only do the marginal distributions behave as they should for the phase space density, but in the classical limit the Wigner function even has the same equation of motion as the classical phase space density! Wigner function is probably the most elegant way known to man to study the classical limit of quantum mechanics. Be advised that at our present level the mathematics in part (b) is a challenge.

Chapter 8

Interacting gas

All systems we have studied so far in a truly microscopic fashion are ideal gases, in other words, the Hamiltonian is a straight sum of the Hamiltonians of the constituent particles. Here we will discuss a basically perturbative treatment of a classical almost-ideal gas.

Of course, the whole point of many systems is that they are *not* ideal gases. Unfortunately, while the degree of interest of the statistical-mechanics analysis then rises steeply, so does also the degree of difficulty. The present *virial expansion* is subject to the usual limitations of perturbation theory in physics; if one needs to go past the lowest nontrivial order in the perturbation, chances are that one needs to sum to all orders. A traditional solution to this problem is to sum some terms in the perturbation theory to infinite order using Feynman diagrams. At this age of computers other methods have become available. For instance, one may solve problems in classical statistical mechanics basically by integrating Newton's equations of motion for a large number of particles. Analogously, in quantum statistical mechanics there are the numerical Quantum Monte Carlo simulations. We will not go into any discussions of this kind of methods in the present elementary course, but merely point out that almost all of the research into strongly interacting or "strongly correlated" statistical mechanics systems is nowadays carried out numerically.

8.1 Virial expansion

For the most part the interactions between the atoms or molecules are two-body interactions, so that in our example we write the Hamiltonian for N particles as

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} U(\mathbf{r}_i - \mathbf{r}_j) \quad (8.1)$$

$$= \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} U(|\mathbf{r}_i - \mathbf{r}_j|) \quad (8.2)$$

The canonical partition function is as usual,

$$Z_N = \frac{1}{N! \lambda^{3N}} Q_N; \quad (8.3)$$

$$Q_N = \int d^{3N} q e^{-\frac{1}{2kT} \sum_{i \neq j} U(|\mathbf{r}_i - \mathbf{r}_j|)} \equiv \int d^{3N} q e^{-\beta \sum_{i<j} U_{ij}}, \quad (8.4)$$

where λ once more is the thermal de Broglie wavelength. The assumption of a central two-body potential $U(|\mathbf{r}_i - \mathbf{r}_j|)$ could be relaxed, but we keep it for the sake of a concise argument. Of course, the problem remains that the integrals Q_N in general cannot be carried out.

Here we proceed in the following way. First notice the trivial rewrite

$$Q_N = \int d^{3N} q \prod_{i<j} e^{-\beta U_{ij}}. \quad (8.5)$$

Second, the usual potential energies satisfy $U \rightarrow \infty$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow 0$ and (by convention) $U \rightarrow 0$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow \infty$, so that the limits $e^{-\beta U_{ij}} \rightarrow 0$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow 0$ and $e^{-\beta U_{ij}} \rightarrow 1$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow \infty$ follow. This inspires us to define

$$e^{-\beta U_{ij}} = 1 + f_{ij}, \quad (8.6)$$

where the newly defined f_{ij} satisfy $f_{ij} \rightarrow -1$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow 0$ and $f_{ij} \rightarrow 0$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow \infty$. The first few integrals Q_N are then written

$$\begin{aligned} Q_0 &= 1, \\ Q_1 &= \int d^3 r_1, \\ Q_2 &= \int d^3 r_1 d^3 r_2 (1 + f_{12}), \\ Q_3 &= \int d^3 r_1 d^3 r_2 d^3 r_3 (1 + f_{12})(1 + f_{13})(1 + f_{23}), \\ &= \int d^3 r_1 d^3 r_2 d^3 r_3 (1 + f_{12} + f_{13} + f_{23} + f_{12}f_{23} + f_{23}f_{13} + f_{12}f_{13}f_{23}), \\ &\dots \end{aligned} \quad (8.7)$$

The first equation $Q_0 = 1$ is a convention that makes $Z_0 = 1$, and, among other things, will ensure that the pressure tends to zero in the limit of zero density.

Now, a perusal of the results will give a graphical representation of the quantities Q_N . Let us denote each integral by a bubble \circ , and each integrand f_{ij} with a line — that connects the bubbles for the integrals over \mathbf{r}_i and \mathbf{r}_j . We then have

$$Q_1 = \circ$$

$$\begin{aligned}
Q_2 &= \text{diagram 1} + \text{diagram 2} \\
Q_3 &= \text{diagram 3} + \text{diagram 4} + \text{diagram 5} + \text{diagram 6} \\
&\quad + \text{diagram 7} + \text{diagram 8} + \text{diagram 9} + \text{diagram 10} \\
&= \text{diagram 11} + 3 \left(\text{diagram 12} \right) + 3 \left(\text{diagram 13} \right) + \text{diagram 14} \\
&\dots
\end{aligned} \tag{8.8}$$

Obviously, to obtain Q_N , one has to calculate all graphs obtained by combining N bubbles in all possible ways with $0, \dots, N$ lines. Many graphs have the same value, as they differ only in the names of the integration variables.

A rather elegant result emerges if we go to the grand canonical ensemble. To begin with, we have

$$\ln \mathcal{Z} = \ln \sum_{N=0}^{\infty} \left(\frac{z}{\lambda^3} \right)^N \frac{Q_N}{N!} = \ln \left(1 + \xi Q_1 + \frac{1}{2} \xi^2 Q_2 + \frac{1}{6} \xi^3 Q_3 + \dots \right), \tag{8.9}$$

with the obvious temporary notation

$$\xi = \frac{z}{\lambda^3}. \tag{8.10}$$

Using the expansion

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \tag{8.11}$$

we have the expansion in the parameter ξ as

$$\ln z = \xi Q_1 + \frac{1}{2!} \xi^2 (Q_2 - Q_1^2) + \frac{1}{3!} [Q_3 - 3Q_1 Q_2 + 2Q_1^3] + \dots \tag{8.12}$$

Let us look at this term by term in the powers of ξ . We have, using the obvious rules for multiplying the graphs Q_N , the following:

$$\begin{aligned}
1! \xi^1 &: Q_1 = \text{bubble} = V \equiv b_1 \\
2! \xi^2 &: Q_2 - Q_1^2 = \text{diagram 1} + \text{diagram 2} - \text{diagram 3} \\
3! \xi^3 &: \text{diagram 4} + 3 \text{diagram 5} + 3 \text{diagram 6} + \text{diagram 7} \\
&\quad - 3 \text{diagram 8} - 3 \text{diagram 9} + 2 \text{diagram 10} \\
&= 3 \text{diagram 11} + \text{diagram 12} \\
&\dots
\end{aligned} \tag{8.13}$$

The coefficients are precisely the graphs Q_N , except that all graphs that break apart into pieces that are not connected by a line are omitted. Let us denote such connected graphs by \tilde{Q}_N . We therefore have the expansion of the logarithm of the grand partition function as

$$\ln \mathcal{Z} = \sum_{N=0}^{\infty} \frac{\tilde{Q}_N}{N!} \left(\frac{z}{\lambda^3} \right)^N. \quad (8.14)$$

Although we have demonstrated (8.14) only up to the order $N = 3$, it holds true to all orders. This result constitutes what is known as *linked-cluster theorem*, a common tool in all sorts of Feynman graph expansions. Since the integrand f_{ij} effectively cuts off at large distances between the integration variables \mathbf{r}_i and \mathbf{r}_j , one expects that in the connected graphs the integrand is effectively nonzero only when all integration variables are close to one another. Such a cluster can be anywhere, so that one of the integrals will give a result proportional to the volume V . However, thereafter the remaining integrals are over a bounded volume. One therefore expects that $Q_N \propto V$ for all $N > 0$.

To obtain the equation of state for the interacting gas is now a routine exercise in statistical mechanics. We first have

$$N = z \frac{\partial}{\partial z} \ln \mathcal{Z} = \sum_{N=1}^{\infty} \frac{\tilde{Q}_N}{(N-1)!} \left(\frac{z}{\lambda^3} \right)^N = \tilde{Q}_1 \xi + \tilde{Q}_2 \xi^2 + \frac{1}{2} \tilde{Q}_3 + \dots \quad (8.15)$$

But $\tilde{Q}_1 = V$, so that we find for the density of the gas the power series expansion in ξ

$$n = \frac{N}{V} = \xi + \frac{\tilde{Q}_2}{V} \xi^2 + \dots, \quad (8.16)$$

which is readily inverted as

$$\xi = n - \frac{\tilde{Q}_2}{V} n^2 + \dots \quad (8.17)$$

On the other hand

$$pV = -\Omega = kT \ln \mathcal{Z} = kT \left(\tilde{Q}_1 \xi + \frac{\tilde{Q}_2}{2} \xi^2 + \dots \right). \quad (8.18)$$

Substitution of (8.17) immediately gives

$$\frac{p}{kT} = n - \frac{\tilde{Q}_2}{2V} n^2 + \dots \quad (8.19)$$

We have therefore the first two *virial coefficients*

$$B_1(T) = 1, \quad B_2(T) = \frac{\tilde{Q}_2}{2V} \quad (8.20)$$

in the virial expansion of the equation of state in powers of density

$$\frac{p}{kT} = \sum_{N=1}^{\infty} B_N(T) n^N. \quad (8.21)$$

This is the one and only truly ab-initio result we are going to derive here for any interacting statistical-mechanics system. The second virial coefficient is easy to calculate numerically, if not downright analytically, when the interaction potential between the particles $U(|\mathbf{r}_i - \mathbf{r}_j|)$ is known. However, the effort required to find all connected graphs and to calculate the respective integrals (numerically) tends to grow rapidly with the order N , and neither does one expect to gain a whole lot from such an enterprise: As noted before, if one needs to go beyond the leading nontrivial perturbative correction, $B_2(T)$, chances are that one would have to calculate all $B_N(T)$. A partial cure in many-body perturbation theory is to calculate some specifically picked subclass of the graphs up to infinite order. As a purely arbitrary example, here one could easily sum up all “ring diagrams” in which the lines between the blobs make a closed loop that does not intersect with itself. We, however stay away from this business.

8.2 Exercises

8.1. Let us consider the *Lennard-Jones potential*

$$U(r) = E_0 \left[\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right].$$

This one looks like a typical molecular potential (plot to see!), with $-E_0$ being the minimum value of the potential energy at the equilibrium distance r_0 , and even has the form of the prototypical van der Waals interaction between the atoms, $\propto 1/r^6$, at large distances. Find the second virial coefficient numerically. The result should be expressed in the form of a dimensional constant times a dimensionless function of the dimensionless variable E_0/kT .

Chapter 9

Mean field theories

As was already noted in the previous Chapter, when the particles in a many-body system interact, an explicit and more or less analytical solution of the statistical mechanics and thermodynamics can only be found in a few very special cases. In the present Chapter the focus is on one particular consequence of the interactions: phase transitions. Power-series expansions in basically the interaction strength, as in the previous Chapter, are a valuable tool if one wants to analyze the leading nontrivial correction to the noninteracting case. Unfortunately, at a phase transition point the state variables of a system are typically nonanalytic functions of thermodynamics parameters such as temperature or pressure, and there is even in principle no hope that any perturbation expansion could produce a nonanalytic result.

Monte Carlo simulations on a computer, both in the classical theory and in quantum mechanics, have completely changed the playing field for a many-body theorist. However, Monte Carlo simulations are beyond our scope. *Renormalization group* methods may tame the nonanalyticity of continuous phase transitions, but we also lack the time to address them. We are left with *mean field theories*, which we will discuss on an elementary level.

Mean field theories are an ad-hoc method that often correctly captures the qualitative phenomenology of a phase transition, but one should not expect quantitatively accurate results. They also suffer from the shortcoming that generally there is no systematic way to find corrections to a mean field theory that would make it more accurate quantitatively. Nonetheless, mean field theories are usually the first tool when, say, a new phase transition is encountered in some system; and on many an occasion, the last and final tool as well.

9.1 Model of magnetic system

9.1.1 Ising model

A ferromagnetic phase transition entails that below a certain critical or *Curie* temperature T_C spins such magnetic moments of atoms more or less align and

produce a macroscopic magnetization in the sample that prevails even in the absence of an external magnetic field. Soft iron is a ferromagnet at room temperature, albeit your ordinary piece of iron is not automatically magnetized. The reason is not that the atomic magnetic moments were not aligned, but that they are aligned in the same direction only in a small region called domain and the magnetization cancels when averaged over the domains. We ignore the domains as just one among the numerous approximations in our discussion.

We study the Ising model for a ferromagnet. For concreteness we talk about a one-dimensional chain, although the results are stated in a form that applies in an arbitrary number of dimensions. We have a chain of N sites i , and a classical two-valued spin sitting at each site, $s_i = \pm 1$, with the Hamiltonian

$$\mathcal{H} = -\epsilon \sum_{\langle ij \rangle} s_i s_j - H \sum_i s_i. \quad (9.1)$$

Here $\epsilon > 0$ is the strength of the spin-spin interactions, H is the applied magnetic field in suitable units, and the brackets denote summation over nearest-neighbor pairs. In one dimension this basically means that we sum $\sum_i s_i s_{i+1}$, with some rules for the ends of the spin chain. For instance, periodic boundary conditions dictate that the 1st and $N + 1$ th site are the same. We denote by γ the number of nearest neighbors; $\gamma = 2$ in one dimension.

In the present formal setting the idea of the mean field theory is to introduce the statistical-mechanics average of the spins, here $L = \langle s_i \rangle$, and regard the deviation of the spin from its average $s_i - L$ as small. This gives the calculation

$$\begin{aligned} \sum_{\langle ij \rangle} s_i s_j &= \sum_{\langle ij \rangle} [(s_i - L) + L][(s_j - L) + L] \simeq \sum_{\langle ij \rangle} [L^2 + L(s_i - L) + L(s_j - L)] \\ &= 2L \sum_i s_i - L^2 = \gamma L \sum_i s_i - \frac{1}{2} N \gamma L^2. \end{aligned} \quad (9.2)$$

The approximation is to drop the terms that are second order in the assumedly small quantities $s_i - L$. We then have the mean field version of the Hamiltonian

$$\begin{aligned} \mathcal{H}_M &= -(\gamma \epsilon L + H) \sum_i s_i + \frac{1}{2} \gamma \epsilon N L^2 \\ &= \sum_i \{-(\gamma \epsilon L + H) s_i + \frac{1}{2} \gamma \epsilon L^2\} \equiv \sum_i h(s_i). \end{aligned} \quad (9.3)$$

The first version in (9.3) emphasizes the mean field idea: The surrounding spins generate a magnetic field that adds to the external magnetic field. This is an approximation, and may be quite crude because it ignores the correlations between the spins. If a particular spin s_i is pointing a particular way, chances are that the neighboring spins are more likely to point in the same direction than the average value for the spins would imply, and are not accurately represented by the mean field.

The second form in (9.3) shows that the mean-field approximation makes a non-interacting ideal gas out of the spins, an easily solvable model. In particular,

the expectation value of the spins may be obtained as

$$\langle s_i \rangle = \frac{\sum_{s=\pm 1} s e^{-\beta h(s)}}{\sum_{s=\pm 1} e^{-\beta h(s)}} = \tanh[\beta(H + \gamma \epsilon L)]. \quad (9.4)$$

But this expectation value is supposed to equal the expectation value of the spins L , so that we have a self-consistency condition for the mean field theory

$$L = \tanh[\beta(H + \gamma \epsilon L)]. \quad (9.5)$$

In mean field theories it is a frequent theme to assume a mean field, and then calculate the value of the mean field from statistical mechanics to get a self-consistency conditions such as (9.5). As an important aside, while the additive terms in the Hamiltonian that do not depend on the dynamical variable s_i do not contribute in our present argument, they should nonetheless be kept. The value of the mean field may and does depend, say, on temperature, so the additive “constant” in the Hamiltonian $\frac{1}{2}\gamma\epsilon NL^2$ depends on the parameters of the problem. It may affect the results, depending on how the analysis is done.

Spontaneous magnetization at $H = 0$

At zero applied magnetic field, $H = 0$, the self-consistency condition (9.5) for the mean field theory reads

$$L = \tanh \left[\frac{\gamma \epsilon}{kT} L \right]. \quad (9.6)$$

Let us write this as $y_1(L) = y_2(L)$, with $y_1(L) = L$, $y_2(L) = \tanh xL$, and $x = \gamma\epsilon/(kT)$. We plot in Fig. 9.1 both the line $y_1(L)$ and three curves $y_2(L)$, the latter for $x = 0.8$, $x = 1$ and $x = 1.3$ from bottom to top. The equation $y_1(L) = y_2(L)$ obviously has a nontrivial solution for spontaneous magnetization L if and only if the initial slope of the tanh function as a function of the mean field L is larger than unity, i.e.,

$$x = \frac{\gamma \epsilon}{kT} > 1 \Rightarrow T < T_C = \frac{\gamma \epsilon}{k}. \quad (9.7)$$

Suppose we are operating below, but close to, the critical temperature, so that $x > 1$ but $x \simeq 1$, and of course, $L \ll 1$. Then we may expand (9.7) as

$$L = \tanh xL \simeq xL - \frac{1}{3}(xL)^3, \quad (9.8)$$

and obtain in the limit $x \rightarrow 1$

$$L = \sqrt{3 \left(1 - \frac{1}{x} \right)} = \sqrt{3 \left(1 - \frac{T}{T_C} \right)}. \quad (9.9)$$

Therefore, when the critical temperature is approached from below, the spontaneous magnetization behaves as

$$L \propto (T_C - T)^{1/2}. \quad (9.10)$$

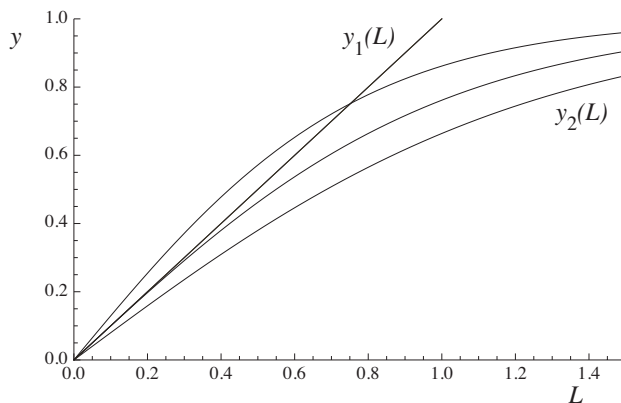


Figure 9.1: Graphical construction for the self-consistency condition of the mean field theory in the Ising model.

A further analysis of the thermodynamics shows that at temperatures below the critical temperature the free energy of the solution with a spontaneous magnetization is lower than the free energy of the nonmagnetized state, so that the mean field theory predicts a transition from a nonmagnetized to a ferromagnetic state at the Curie temperature. The approximation to the transition temperature from mean field theory may be poor, but the qualitative picture is correct. Besides, just like in the mean field theory, in real physics the magnetization tends to zero nonanalytically according to a power-state law as the critical temperature is approached from below. The actually observed *critical exponent* for magnetization is not exactly $1/2$ as in mean field theory, but again the qualitative behavior of the mean field result is correct.

Heat capacity at $H = 0$

The energy is easily calculated from the mean-field Hamiltonian, and at zero applied magnetic field it equals

$$E = \langle \mathcal{H}_M \rangle = N \langle h(s_i) \rangle = N \langle -\gamma \epsilon L s_i + \frac{1}{2} \gamma \epsilon L^2 \rangle = -\frac{1}{2} \gamma \epsilon N L^2. \quad (9.11)$$

Above the transition temperature we have $L = 0$, hence $E = 0$ and $C_H = 0$. There is a subtlety here that occurs in magnetic systems and also has to do with the sign of the work term in $dW = M dB$. It transpires that this work term is correct if the energy in the magnetization of the sample is taken to be a reservoir *outside* of the internal energy. But here we have written the Hamiltonian *with* the magnetization term $-H \sum_i s_i \cong -BM$. This means that, what we thought was internal energy is thermodynamically actually the enthalpy. Because of the abundance of H s in the argument, we temporarily denote enthalpy by E , as in $E = U - BM$. We wrote E for enthalpy in the expectation value of the Hamiltonian, and obtained heat capacity at constant magnetic field not constant magnetization. Likewise, if a straightforward statistical mechanics

versus thermodynamics argument seems to give you Helmholtz free energy, for a magnetic systems you actually got the Gibbs free energy.

On the other hand, for temperatures below, albeit close to, the critical temperature, we have from (9.9) and (9.11)

$$E = -\frac{3}{2}\gamma\epsilon N \left(1 - \frac{T}{T_C}\right); \quad C_H = \frac{3\gamma\epsilon N}{2T_C} = \frac{3}{2}Nk. \quad (9.12)$$

In a result typical to mean field theory, the heat capacity jumps when the temperature is lowered through the critical temperature. Realistic systems typically show a logarithmic divergence of the heat capacity. A logarithm diverges at zero argument slower than any power law, so one could say that in such a case the heat capacity diverges with the critical exponent 0.

Magnetization at critical temperature

In the next example we assume that the temperature is at the exact critical temperature, and ask how the magnetization behaves as a function of magnetic field in the limit of small fields. At $T = T_C$ and $H = 0$ we have $L = 0$, so that the assumption of small L appears warranted. We therefore expand the self-consistency condition (9.5) as

$$L = \tanh(L + \beta_C H) \simeq L + \beta_C H - \frac{1}{3}(L + \beta_C H)^3. \quad (9.13)$$

Apparently L tends to zero more slowly than H , so that we have

$$L^3 \simeq 3\beta_C H, \quad L \propto H^{1/3}. \quad (9.14)$$

This shows another nonanalytic behavior with another critical exponent that is in a qualitative but not in a full quantitative agreement with experiments.

Susceptibility at $H = 0$

We leave it as an exercise to show that another critical exponent governs the behavior of the susceptibility in the neighborhood of the critical temperature:

$$\chi = \left(\frac{\partial L}{\partial H} \right)_{H=0} \propto |T_C - T|^{-1}. \quad (9.15)$$

This applies on both sides of the critical temperature, albeit with different numerical coefficients. Experiments find a divergence and a critical exponent as well, although it is unmistakably different from -1 .

9.1.2 Spontaneously broken symmetry

The discussion of the Ising model already implicitly brought up a very general topic with continuous phase transitions that we discuss next, *spontaneous symmetry breaking*. Namely, below the Curie temperature and at zero field, if L is a

solution to the self-consistency condition (9.5), then so is $-L$. Which one will be realized in Nature, given that there is no apparent reason for either particular pick?

The situation gets even more mysterious when one notices that at zero magnetic field the Hamiltonian of the Ising model is such that the configurations $\{s_i\}$ and $\{-s_i\}$ have the exact same energy, so that they also have the exact same weights in statistical-mechanics sums over the states. The inescapable conclusion is that statistical mechanics gives $\langle s_i \rangle = 0$.

A nonzero magnetization is a case of *spontaneously broken symmetry*. The symmetry of the state is less than the symmetry of the Hamiltonian. In the Ising model the spin inversion symmetry $s_i \rightarrow -s_i$ is broken. Coming with spontaneously broken symmetry is *order parameter*, some system variable X such that $X = 0$ in the symmetric phase and $X \neq 0$ in the broken-symmetry phase. In the Ising model $L = \langle s_i \rangle$ would be an order parameter.

Also invariably associated with spontaneously broken symmetry, there will be a *Goldstone mode* or *Goldstone boson*, an excitation mode such that $\omega_{\mathbf{q}} \rightarrow 0$ as $|\mathbf{q}| \rightarrow 0$. For ferromagnets this boson is called magnon.

Often there may be a *symmetry breaking field* in a system, an external field that causes a nonzero value of the order parameter. In the Ising model the magnetic field H serves as the symmetry breaking field.

While giving lofty names to a few observations, our discussion of the terminology associated with spontaneously broken symmetry does nothing to resolve the problem that broken symmetry flat out violates statistical mechanics. We will discuss three approaches to this conundrum.

First, in practice one usually puts the spontaneously broken symmetry, a nonzero value of an order parameter, in to the theory by hand, then finds a self-consistency condition that gives the value for the order parameter. This is what we did with the Ising model. This gimmick, of course, in no way explains the origin of symmetry breaking.

A piece of reasoning that does substantially better goes as follows for the Ising model. Let us write

$$H_0 = -\epsilon \sum_{\langle ij \rangle} s_i s_j, \quad H_I = -H \sum_i s_i. \quad (9.16)$$

By symmetry the result

$$\langle s_i \rangle_0 = \frac{\sum_{\{s_i\}} s_i e^{-\beta H_0}}{\sum_{\{s_i\}} e^{-\beta H_0}} = 0 \quad (9.17)$$

inevitably holds. But, if the mathematics could be done, it might be that

$$L = \lim_{H \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\sum_{\{s_i\}} \left[\left(\frac{1}{N} \sum_{i=1}^N s_i \right) e^{-\beta(H_0 + H_I)} \right]}{\sum_{\{s_i\}} e^{-\beta(H_0 + H_I)}} \neq 0. \quad (9.18)$$

It could be (there are known examples) that if one first takes the thermodynamic limit and then the limit of zero symmetry breaking field, the order parameter

tends to a nonzero limit. Physically, in a large sample, a tiniest magnetic field could elicit a substantial magnetization.

This idea is plausible while there is a physical symmetry breaking field, but that is not always the case. Consider a Bose-Einstein condensate. The condensate is not only marked with a macroscopic occupation number of a quantum state, but it should also acquire a phase. This is analogous to the phase of a quantum state. A global phase of a wave function is unobservable, but interference phenomena will reveal relative phases of wave functions. Accordingly, two dilute Bose-Einstein condensates, when overlaid, produce an interference pattern with regions of higher and lower atom density. The presence of the interference tells us that the condensates seemingly have phases, although the interference pattern varies at random from experiment to experiment as if the relative phase of the condensates were random in each experiment.

The phase violates a symmetry of the Hamiltonian that it is invariant under the choice of the phase of the wave function, or “gauge symmetry,” so that we have here another case of spontaneous symmetry breaking. Moreover, in this case there is no known physical symmetry breaking field, so that the idea that the phase is a result of some minute residual field does not wash. The maybe surprise explanation for the symmetry breaking in this case is that the symmetry actually is not broken. Instead, the reduction of the quantum state of the two overlapping condensates while the atoms are being detected produces the import of the phase. Basically, a measurement of the phase will produce a phase!

9.2 van der Waals equation of state

As a more ad-hoc approach to mean field theories we study the classical *van der Waals* equation of state for an interacting gas. We begin by reciting the interacting Hamiltonian (8.1) and the result obtained when the momentum variables have been integrated out, (8.3):

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} U(\mathbf{r}_i - \mathbf{r}_j), \quad (9.19)$$

$$Z_N = \frac{1}{N! \lambda^{3N}} \int d^{3N} \mathbf{q} e^{-\frac{1}{2kT} \sum_{i \neq j} U(\mathbf{r}_i - \mathbf{r}_j)}. \quad (9.20)$$

The van der Waals equation of state results from the following approximate argument. Typically, the two-body potential $U(\mathbf{r})$ has a hard core, call the radius a , such that the potential is effectively positive and infinite for $r < a$. Second, instead of attempting to account for the precise potential energy of the other particles on the i^{th} particle as a function of the positions of the other particles, we assume that the other particles are evenly distributed outside of the hard core of the potential. This gives for a fixed “spectator” particle i the

estimate

$$\begin{aligned} \sum_{j \neq i} U(\mathbf{r}_i - \mathbf{r}_j) &\simeq \sum_j \frac{1}{V} \int_{|\mathbf{r}_j - \mathbf{r}_i| > a} d^3 r_j U(\mathbf{r}_i - \mathbf{r}_j) \simeq \frac{N}{V} \int_{r > a} d^3 r U(\mathbf{r}) \\ &= -\frac{N\tilde{U}}{V}. \end{aligned} \quad (9.21)$$

The potential $U(\mathbf{r})$ typically has an “attractive tail”, i.e., it is negative for large r , so that the constant \tilde{U} defined in this way is positive. We have the approximation

$$Z_N = \frac{1}{N! \lambda^3} \int d^3 N q e^{\frac{N^2 \tilde{U}}{2KT V}} \simeq \frac{(V - NV_0)^N}{N! \lambda^{3N}} e^{\frac{N^2 \tilde{U}}{2KT V}}. \quad (9.22)$$

In the last step we have considered that for each particle the volume over which the integral is taken is the total volume minus the volume of the hard cores of the remaining particles, with $V_0 = \frac{4\pi}{3} a^3$.

Run-of-the mill thermodynamics then gives

$$F(T, V, N) = -kT \ln Z_N = -kTN \left[\ln \frac{V - NV_0}{\lambda^3} + \frac{N\tilde{U}}{2KT V} - \ln N + 1 \right], \quad (9.23)$$

$$p = -\left(\frac{\partial F}{\partial V} \right)_{T, N} = \frac{kTN}{V - NV_0} - \frac{N^2 \tilde{U}}{2V^2}. \quad (9.24)$$

Traditionally one introduces the *number of moles* $n = N/N_A$ and the *gas constant* $R = N_A k = 8.31 \text{ JK}^{-1} \text{ mol}^{-1}$, $N_A = 6.022 \times 10^{23}$ being the Avogadro number, and besides there is also a standard redefinition of the constants \tilde{U} and V_0 so that the equation of state of the gas is written

$$\left(p + \frac{an^2}{V^2} \right) (V - nb) = nRT. \quad (9.25)$$

This is the celebrated *van der Waals equation of state*.

Let us now rewrite

$$p = \frac{nRT}{V - nb} - \frac{an^2}{V^2}. \quad (9.26)$$

Figure 9.2 shows three isotherms, $p(V)$ for a fixed temperature, for three temperatures that decrease as indicated in the figure. Obviously, when the temperature is low enough, the isotherms allow two (or three) different volumes for the same pressure. We will return to this point below in more detail, but this circumstance tells us that a liquid-gas phase transition has taken place. There will be multiple solutions when $p(V)$ has an extremum; with decreasing T this situation first appears at the values of V and T for which both equations

$$\frac{\partial p}{\partial V} = 0, \quad \frac{\partial^2 p}{\partial V^2} = 0 \quad (9.27)$$

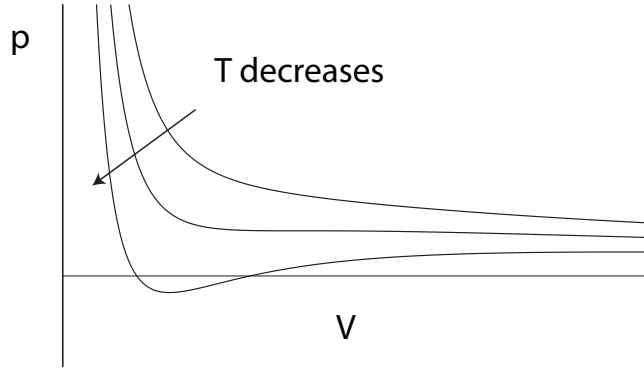


Figure 9.2: Three isotherm of the van der Waals equation of state.

have a real solution. These equations together with (9.26) constitute three equations out of which three quantities, p , V and T , can be determined. These particular values make the critical point of the liquid-gas phase transition, and are easily found to be

$$p_c = \frac{a}{27b^2}, \quad V_c = 3nb, \quad RT_c = \frac{8a}{27b}. \quad (9.28)$$

The isotherms in Fig. 9.2, however, have a problem: there are regions with $\partial p / \partial V > 0$, i.e., regions of negative compressibility, indicating thermodynamic instability. In fact, though, such isotherm do not predict an unphysical behavior. To see this, consult Fig. 9.3. On the top left we have a graph of $p(V)$ that has negative-compressibility piece in it between the local minimum and maximum of $p(V)$. The top right panel draws the same graph as $V(p)$. Between the local minima and maxima of $p(V)$ denoted by p_l and p_u , this is a multivalued “function” (citation marks because a function, by definition, is single-valued) of p ; we denote the three branches as $V_u(p)$, $V_m(p)$ and $V_l(p)$. The question is, given a p between p_l and p_u (and a fixed T), which of the possibly three different values of V will realized? The answer, of course, is, the value of V that gives the minimum of Gibbs free energy.

Now, we have

$$V = \left(\frac{\partial G}{\partial p} \right)_T, \quad (9.29)$$

so that along an isotherm with a constant T

$$G = \int^P dp V(p) \quad (9.30)$$

holds true. Here the integral is a path integral along the isotherm, and P denotes the end point of the path. In the lower-right panel of Fig. 9.3 we have drawn the value $G(P)$ as a “function” of the pressure $p(P)$ as the point P progresses along the isotherm. A few arrows indicate how the curve is formally traveled as

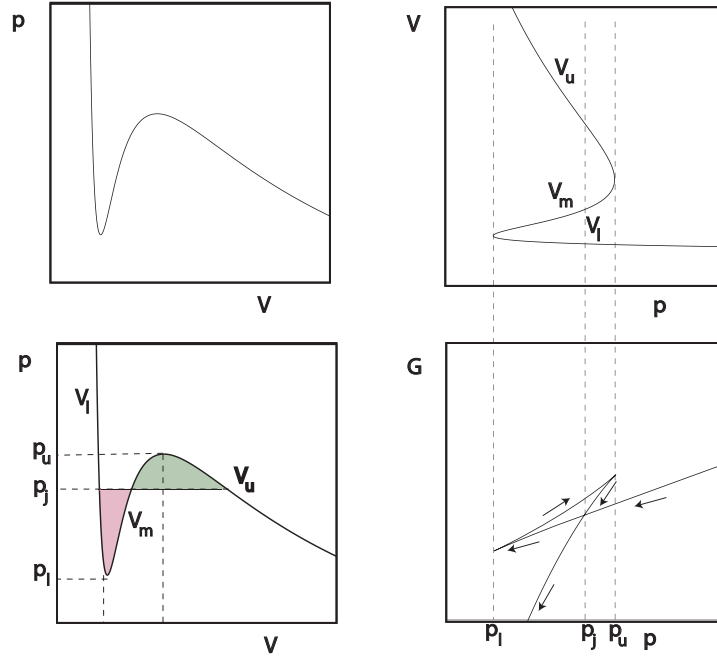


Figure 9.3: Maxwell construction.

the volume is increased. The qualitative behavior of this curve can be surmised easily by considering the variation of V with p as shown in the upper-left panel. The sharp corners in $G(p)$ correspond to the extrema of $p(V)$. The piece of $G(p)$ coming in to the first corner corresponds to the branch $V_l(p)$, the piece between the sharp corners to $V_m(p)$, and the piece exiting from the second corner to $V_u(p)$. However, a physical system would not follow the curve as the arrows indicate if for a given p there is a state available with a lower the Gibbs free energy. This is clearly the case for the closed-loop part of $G(p)$. Thus, at the pressure p_j the system jumps from the branch V_l to the branch V_u .

Finally, to determine p_j we note that the total change of the Gibbs free energy along the closed loop equals zero. Now refer to the lower-left panel in Fig. 9.3. The total change of Gibbs free energy when one travels along the curve from the pressure p_j on the branch V_l to the same pressure on the branch V_u , to be set equal to zero, gives the argument

$$\Delta G = - \int_{p_l}^{p_j} V_l dp + \int_{p_l}^{p_j} V_m dp + \int_{p_j}^{p_u} V_m dp - \int_{p_j}^{p_u} V_l dp = 0. \quad (9.31)$$

These are with judicious definitions of the upper and lower limits of the integration and the signs in front, so that we have the condition

$$\int_{p_l}^{p_j} (V_m - V_l) dp = \int_{p_j}^{p_u} (V_u - V_m) dp. \quad (9.32)$$

This one says that the areas of the reddish and greenish regions in the lower-left

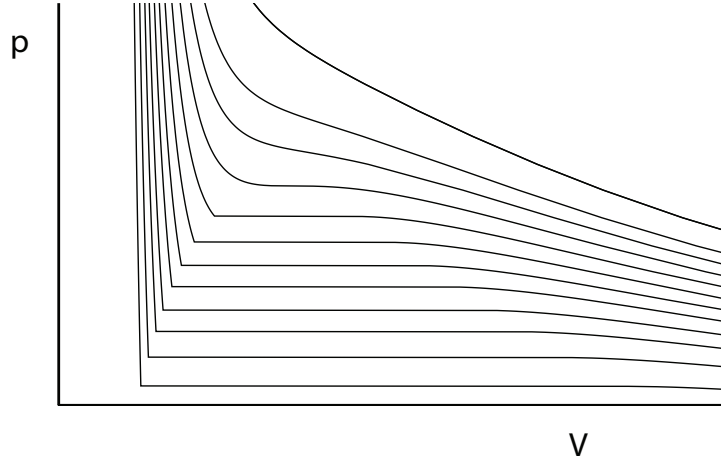


Figure 9.4: Isotherms from the Maxwell construction. The temperature increases in the curves from top to bottom. The critical isotherm with $T = T_c$ is the fourth one from the top.

panel are the same.

The result is what is known as the *Maxwell construction*: Instead of entering into an unstable regime with $\partial p / \partial V > 0$, the isotherm takes a shortcut from V_l to V_u at a p_j such that the colored area as in the lower-left panel are equal. Isotherms from the Maxwell construction are illustrated in Fig. 9.4. An isotherm with a flat part, in other words, a pressure independent of volume, clearly indicates a phase transition: The volume may vary for a given pressure because it is possible to divide the molecules freely between the liquid and gas phases with different volumes per molecule. The area with flat parts in the isotherms is the *coexistence region* of gas and liquid in the Vp plane.

9.3 Exercises

9.1. By taking the derivative of the self-consistency condition of our 1D Ising model with respect to the magnetic field, show that the susceptibility diverges like $|T - T_c|^{-1}$ on both sides of the critical temperature.

9.2. As we all well know, in the Bragg-Williams mean field approximation the Hamiltonian for the one-dimensional Ising model is

$$\mathcal{H}_{\text{MF}} = -(\gamma\epsilon L + H) \sum_i s_i + \frac{1}{2} \gamma\epsilon N L^2,$$

where $\gamma = 2$, $\epsilon > 0$ is a constant characterizing the spin-spin interactions, H is the applied magnetic field, N is the number of lattice sites i containing the spins $s_i = \pm 1$, and $L = \langle s_i \rangle$ is the mean field.

(a) Show that the free energy is

$$G = -kTN \ln\{2 \cosh[\beta(\gamma\epsilon L + H)]\} + \frac{1}{2} \gamma\epsilon NL^2.$$

It is a subtlety of magnetic systems that when you calculate the free energy according to $-kT \ln Z$, the result is indeed Gibbs not Helmholtz!

(b) Now regard the mean field as an independent parameter, and find the extremum of free energy with respect to the value of the mean field. Show that the result is the usual self-consistency condition of the mean field theory, $L = \tanh[\beta(\gamma\epsilon L + H)]$.

(c) In the absence of the magnetic field, below the critical temperature the system has a spontaneous magnetization $L \neq 0$. Show that the nontrivial solution, in fact, gives the minimum of free energy, and the trivial solution $L = 0$ is a maximum of free energy.

(d) In the absence of the magnetic field, below the critical temperature the system has a spontaneous magnetization $L \neq 0$; besides, if L is a possible magnetization, then so is $-L$. Show, however, that a tiniest external field $H \neq 0$ suffices to determine which way the equilibrium magnetization points.

(e) Argue that the ferromagnetic phase transition in this system is continuous at the critical point $T_c, H = 0$, but first order for any $H \neq 0$.

9.3. Consider a one-dimensional chain of Ising spins (values ± 1) whose interaction attempts to turn them antiparallel. It is convenient to imagine that there are two interlaced lattices of spins s_i and S_i , N each. The Hamiltonian is written

$$\mathcal{H} = \frac{1}{2} \epsilon \sum_i (s_i S_i + s_{i+1} S_i) - H \left(\sum_i s_i + \sum_i S_i \right),$$

where $\epsilon > 0$ is the coupling coefficient of the spin-spin interaction, and H is the magnetic field.

(a) Derive in the (Bragg-Williams) mean-field approximation the equations for the averages of the spins $s = \langle s_i \rangle$ $S = \langle S_i \rangle$, and the expression for the free energy per spin

$$\frac{G}{N} = -\epsilon s S - kT \ln\{2 \cosh[\beta(\epsilon S - H)]\} - kT \ln\{2 \cosh[\beta(\epsilon s - H)]\}.$$

Which free energy is a slightly tricky question with magnetic systems; let me just note, without proof, that what you derive from statistical mechanics thinking it is the Helmholtz FE is thermodynamically actually the Gibbs FE.

(b) Show (in the mean-field approximation, of course) that at $H = 0$ and $T = T_c = \epsilon/k$ an *antiferromagnetic* phase transition takes place in which the alternating spins attain non-zero expectation values with opposite signs.

(c) How does the staggered magnetization behave as a function of temperature close to the critical temperature, of course at $H = 0$.

9.4. In a simple example of *Landau's theory for continuous phase transitions*, write the free energy (call it Gibbs, in analogy to magnetic systems) as

$$G = \int dx \left[\alpha_2(T - T_c)\phi(x)^2 + \gamma \left(\frac{d}{dx}\phi(x) \right)^2 + \alpha_4\phi(x)^4 - h(x)\phi(x) \right],$$

where $\phi(x)$ is the order parameter such as magnetization, $h(x)$ is the symmetry-breaking field such as the magnetic field, α_2 , γ and α_4 are positive constants, T is the temperature, and T_c is the purported critical temperature for the phase transition. The idea is now to choose the (here, real) field $\phi(x)$ to minimize the free energy G .

(a) In the absence of external perturbations the field that minimizes Gibbs free energy obviously is independent of position, $\phi(x) \equiv \phi$, and so minimizing Gibbs free energy is the same as finding the real number ϕ that minimizes

$$g(T, \phi) = \alpha_2(T - T_c)\phi^2 + \alpha_4\phi^4.$$

Do it, and show that the result is

$$\phi = 0, T \geq T_c; \quad \phi = \pm \sqrt{\frac{\alpha_2(T_c - T)}{2\alpha_4}}, T < T_c.$$

T_c is the transition temperature, and the order parameter varies with temperature in the typical mean-field fashion.

(b) The purpose of the term proportional to γ is to increase energy for the situation in which the order parameter varies with position. This particular form is the lowest-order derivative [i.e., most sensitive to slow variations of $\phi(x)$] that does not create a preferred direction ($+\phi$ or $-\phi$), and neither goes away as a surface term. Minimizing G is a traditional variational problem with the Euler-Lagrange equation

$$\left[-\gamma \frac{d^2}{dx^2} + 2\alpha_2(T - T_c) + 4\alpha_4\phi(x)^2 \right] \phi(x) = h(x).$$

Suppose for simplicity that $T > T_c$ so that there is no spontaneous order parameter, and that the symmetry breaking field is of the form $h(x) = h_0\delta(x)$ where h_0 is small, then $\phi(x)$ remains “small” and the ϕ^2 term inside the square brackets may be ignored. Solve the remaining problem using your favorite method (e.g., Fourier transformations), and show that the “correlation length” over which there is an induced value of the order parameter tends to infinity when $T \rightarrow T_c+$.

9.5.

(a) Derive for the van der Waals gas the following lowest-order expansion in the constants a and b :

$$V = \frac{nRT}{p} + nb - \frac{na}{RT}.$$

- (b) For a diatomic gas one may set $C_V = \frac{5}{2}nR$. Derive an expression for $S = S(T, p)$ to the lowest nontrivial order in the constant a and b .
- (c) Model air with N_2 , for which $a = 0.1408 \text{ Pa m}^6 \text{ mol}^{-1}$ and $b = 0.03913 \text{ m}^3 \text{ mol}^{-1}$. How much does the air in a pump heat up when you try to put twice the ambient pressure in a bicycle tube, starting at 20° C ? You know the answer for an ideal gas already, P. 2.7. The question is, how much difference does it make that the gas is not quite ideal.
- 9.6. Consider the van der Waals equation below the critical temperature, but close to the critical point. Show that the difference in the densities of the liquid and gas phases scales with temperature as $\rho_l - \rho_g \propto (T_c - T)^{1/2}$. This is once more the characteristic mean-field scaling of the order parameter with temperature.

Chapter 10

Fluctuations and response

A system that just sits there alone is of little interest in physics. As it comes to many-body systems, most of the interesting information is obtained in two ways: Measuring a value of an observable, or perturbing the system and seeing how it responds. This is the general area of the present chapter. The final goal here is the fluctuation-dissipation theorem, which maybe surprisingly says that knowing how the system fluctuates is sufficient to predict how it will respond to perturbations, and vice versa. Our mathematical development is uncompromising, so that we will have little time for physics examples. In exchange, the foundation we build is rock solid, and will hold up if the reader ever finds him/herself analyzing linear response, or fluctuations, or a combination thereof.

10.1 Linear response

We consider a system with the *unperturbed Hamiltonian* H_0 driven by the *external fields* $h_j(t)$ that couple to the system observables A_j . This means that the overall Hamiltonian is

$$H = H_0 + H_I \equiv H_0 - \sum_j h_j(t) A_j. \quad (10.1)$$

What we will say below applies as written only if the unperturbed Hamiltonian H_0 does not depend explicitly on time, although a generalization to a time dependent H_0 is possible. The driving fields may be construed very broadly. For instance it could be that $h_j(t) \cong \mathbf{E}(\mathbf{r}, t)$, where the index j denotes positions, and the sum in (10.1) is actually an integral. The present scheme is extremely general, and applies almost always when any system is perturbed by classical electric, magnetic, gravitational, etc., fields. Variations of the theme exist even if the driving fields are quantized, though we do not address them.

As an example, think of a motion of a particle in one dimension under a time dependent but position independent force $F(t)$. We write the interaction

Hamiltonian that couples to the position of the particle x ,

$$H_I = -F(t)x. \quad (10.2)$$

In fact, classically, this would just add a term to the Hamilton's equation of motion of the momentum

$$\dot{p} = \frac{\partial H}{\partial x} = \dots + F(t) \quad (10.3)$$

in perfect accord with Newton's second law. Practical quantum mechanical Hamiltonians are basically obtained by taking the classical Hamiltonian and replacing coordinates and momenta with quantum mechanical position and momentum operators. In this spirit we declare that (10.2) is good also in a quantum system, with x being the position operator.

10.1.1 Perturbation theory

Although much of the most interesting modern work in physics does *not* conform to this assumption, here we think that the fields $h_j(t)$ are a small perturbation rather than substantially alter the whole system. We therefore develop perturbation theory in the strength of the fields. We will find is a version of the usual time dependent perturbation theory adapted for density operators.

Interaction picture

We begin by defining the interaction picture, which maps an arbitrary system operator \mathcal{O} to the operator $\tilde{\mathcal{O}}$ according to

$$\tilde{\mathcal{O}} = e^{i\frac{H_0 t}{\hbar}} \mathcal{O} e^{-i\frac{H_0 t}{\hbar}} \Leftrightarrow \mathcal{O} = e^{-i\frac{H_0 t}{\hbar}} \tilde{\mathcal{O}} e^{i\frac{H_0 t}{\hbar}}. \quad (10.4)$$

By cyclic invariance of the trace we have

$$\text{Tr}[\tilde{\mathcal{O}}] = \text{Tr}[e^{i\frac{H_0 t}{\hbar}} \mathcal{O} e^{-i\frac{H_0 t}{\hbar}}] = \text{Tr}[e^{-i\frac{H_0 t}{\hbar}} e^{i\frac{H_0 t}{\hbar}} \mathcal{O}] = \text{Tr}[\mathcal{O}]. \quad (10.5)$$

Moreover, we have

$$\widetilde{\mathcal{O}\mathcal{P}} = e^{i\frac{H_0 t}{\hbar}} \mathcal{O}\mathcal{P} e^{-i\frac{H_0 t}{\hbar}} = \left(e^{i\frac{H_0 t}{\hbar}} \mathcal{O} e^{-i\frac{H_0 t}{\hbar}} \right) \left(e^{i\frac{H_0 t}{\hbar}} \mathcal{P} e^{-i\frac{H_0 t}{\hbar}} \right) = \tilde{\mathcal{O}}\tilde{\mathcal{P}}, \quad (10.6)$$

and as an immediate consequence,

$$[\widetilde{\mathcal{P}}, \widetilde{\mathcal{O}}] = [\tilde{\mathcal{O}}, \tilde{\mathcal{P}}]. \quad (10.7)$$

Transformation to the interaction picture preserves the most essential structure of operator algebra.

Let us next develop the Liouville-von Neumann equation as follows,

$$\begin{aligned}
i\hbar \frac{\partial \rho}{\partial t} &= i\hbar \frac{\partial}{\partial t} \left[e^{-i\frac{H_0 t}{\hbar}} \tilde{\rho} e^{i\frac{H_0 t}{\hbar}} \right] \\
&= H_0 e^{-i\frac{H_0 t}{\hbar}} \tilde{\rho} e^{i\frac{H_0 t}{\hbar}} - e^{-i\frac{H_0 t}{\hbar}} \tilde{\rho} e^{i\frac{H_0 t}{\hbar}} H_0 + e^{-i\frac{H_0 t}{\hbar}} \left(i\hbar \frac{\partial \tilde{\rho}}{\partial t} \right) e^{i\frac{H_0 t}{\hbar}} \\
&= [H_0, \rho] + e^{-i\frac{H_0 t}{\hbar}} \left(i\hbar \frac{\partial \tilde{\rho}}{\partial t} \right) e^{i\frac{H_0 t}{\hbar}} \\
&= [H, \rho] = [H_0, \rho] - \sum_j h_j [A_j, \rho].
\end{aligned} \tag{10.8}$$

We immediately see that

$$e^{-i\frac{H_0 t}{\hbar}} \left(i\hbar \frac{\partial \tilde{\rho}}{\partial t} \right) e^{i\frac{H_0 t}{\hbar}} = - \sum_j h_j [A_j, \rho], \tag{10.9}$$

or, multiplying this from the left by $e^{i\frac{H_0 t}{\hbar}}$ and from the right by $e^{-i\frac{H_0 t}{\hbar}}$,

$$i\hbar \frac{\partial \tilde{\rho}}{\partial t} = - \sum_j h_j [\tilde{A}_j, \tilde{\rho}]. \tag{10.10}$$

In complete analogy to the use of the interaction picture in the derivation of time dependent perturbation theory in the ordinary state vector quantum mechanics, in the interaction picture the density operator only evolves as a result of the external perturbations.

Perturbation expansion

In a silent premise that is always there in the usual perturbation theory, we assume that in the distant past the driving fields vanished. Therefore, in the distant past the interaction picture the density operator was a constant, $\tilde{\rho}(t = -\infty) \equiv \tilde{\rho}_0$. At any time t the density operator is formally given by

$$\tilde{\rho}(t) = \tilde{\rho}_0 + \int_{-\infty}^t dt' \frac{\partial \tilde{\rho}}{\partial t'} = \tilde{\rho}_0 + \frac{i}{\hbar} \int_{-\infty}^t dt' h_j(t') [\tilde{A}(t'), \tilde{\rho}(t')], \tag{10.11}$$

where the first step is an identity and the second step follows from the interaction picture equation of motion for the density operator (10.10). We have, in fact, accomplished quite a lot by converting a differential equation into an integral equation. Namely, one may iterate the integral equation as follows,

$$\begin{aligned}
\tilde{\rho}_0(t) &= \tilde{\rho}_0, \\
\tilde{\rho}_{n+1}(t) &= \tilde{\rho}_0 + \frac{i}{\hbar} \int_{-\infty}^t dt' h_j(t') [\tilde{A}(t'), \tilde{\rho}_n(t')], \quad n \geq 0.
\end{aligned} \tag{10.12}$$

Under reasonable conditions the sequence of functions obtained in this way converges to a solution of (10.10), and the result comes out as a power series in

the strength of the perturbing fields $h_j(t)$; $\tilde{\rho}_n(t)$ has the powers of the perturbing fields up to the order n .

As already mentioned, the usual case is to assume that the perturbing fields are weak probes not a major influence on the dynamics of the system, and the perturbation series is only taken up to the leading nontrivial, usually first, order in the strength of the fields $h_j(t)$. In this way we have the result

$$\tilde{\rho}_1(t) = \tilde{\rho}_0 + \frac{i}{\hbar} \int_{-\infty}^t dt' h_j(t') [\tilde{A}(t'), \tilde{\rho}_0]. \quad (10.13)$$

Perturbation induced expectation value

Even if we have the state of the system in the bag, as per perturbation theory, one rarely attempts to determine the state experimentally. Rather, one would measure or make use of some observable that changes as a result of the perturbation. Let us call such an observable B . Some statements we will make about the observable B are true only if it does not depend explicitly on time in the original Schrödinger picture (as opposed to the interaction picture); to be on the safe side, let us always assume so. From now on we assume that the “unperturbed” expectation value of the operator equals zero, $\text{Tr}[\tilde{\rho}_0 \tilde{B}] = 0$; if not, one simply replaces the observable by a version in which the nonzero expectation value has been subtracted, $B \rightarrow B - \text{Tr}[\tilde{\rho}_0 \tilde{B}]$.

The expectation value of the operator at the time t up to first order in perturbation theory is

$$\begin{aligned} \langle B(t) \rangle &\equiv \langle B \rangle(t) = \text{Tr}[\rho(t)B] = \text{Tr}[\tilde{\rho}(t), \tilde{B}(t)] \simeq \text{Tr}[\tilde{\rho}_1(t), \tilde{B}(t)] \\ &= \frac{i}{\hbar} \int_{-\infty}^t dt' h_j(t') \text{Tr}\{[\tilde{A}(t'), \tilde{\rho}_0] \tilde{B}(t)\}. \end{aligned} \quad (10.14)$$

The first form is a standard notation, dubious as it might be in the present context. Further progress is made by noting that cyclic permutations inside the trace give

$$\text{Tr}\{[A, \rho]B\} = \text{Tr}\{A\rho B - \rho AB\} = \text{Tr}\{BA\rho - AB\rho\} = \text{Tr}\{[B, A]\rho\}. \quad (10.15)$$

We then have

$$\langle B(t) \rangle = \frac{i}{\hbar} \int_{-\infty}^t dt' h_j(t') \text{Tr}\{[\tilde{B}(t), \tilde{A}_j(t')], \tilde{\rho}_0\}. \quad (10.16)$$

This form is perfectly valid as is, but under usual conditions will simplify significantly:

Lemma. If A and B do not depend on time in the Schrödinger picture and $[\tilde{\rho}_0, H_0] = 0$, then the expression $\text{Tr}\{[\tilde{B}(t), \tilde{A}_j(t')], \tilde{\rho}_0\}$ depends only on the difference between the times t and t' .

Proof.

$$\text{Tr}\{[\tilde{B}(t), \tilde{A}_j(t')], \tilde{\rho}_0\} = \text{Tr}\{\tilde{B}(t) \tilde{A}_j(t') \tilde{\rho}_0 - \dots\}$$

$$\begin{aligned}
&= \text{Tr} \left\{ e^{-i\frac{H_0 t}{\hbar}} B e^{i\frac{H_0 t}{\hbar}} e^{-i\frac{H_0 t'}{\hbar}} A e^{i\frac{H_0 t'}{\hbar}} \tilde{\rho}_0 - \dots \right\} \\
&= \text{Tr} \left\{ e^{-i\frac{H_0(t-t')}{\hbar}} B e^{i\frac{H_0(t-t')}{\hbar}} A \tilde{\rho}_0 - \dots \right\}
\end{aligned} \tag{10.17}$$

based on the assumptions of the lemma and on the cyclic invariance of the trace.

From now on we always take the assumption of the lemma to be valid. In particular, a typical thermal density operator is of the form $\rho = \tilde{\rho} = e^{-\beta H_0}/Z$, and as such satisfies the condition $[\tilde{\rho}, H_0] = 0$. In this light we develop the expectation value (10.16) further as follows,

$$\begin{aligned}
\langle B(t) \rangle &= \frac{i}{\hbar} \int_{-\infty}^t dt' h_j(t') \text{Tr} \{ [\tilde{B}(t-t'), \tilde{A}_j(0)], \tilde{\rho}_0 \} \\
&\equiv 2i \sum_j \int_{-\infty}^t dt' \chi''_{BA_j}(t-t') h_j(t') \\
&= \sum_j \int_{-\infty}^{\infty} dt' \chi_{BA_j}(t-t') h_j(t').
\end{aligned} \tag{10.18}$$

Here we have defined the *response functions* as

$$\chi''_{BA}(t) = \frac{1}{2\hbar} \text{Tr} \{ \tilde{\rho}_0 [\tilde{B}(t), \tilde{A}(0)] \} = \frac{1}{2\hbar} \langle [\tilde{B}(t), \tilde{A}(0)] \rangle_0, \tag{10.19}$$

$$\chi_{BA}(t) = 2i \theta(t) \chi''_{BA}(t). \tag{10.20}$$

The first form emphasizes the observation that the response function $\chi''(t)$ is a “two-time” expectation value in the “unperturbed” state $\tilde{\rho}_0$. The unprimed response function $\chi(t)$ is obtained from $\chi''(t)$ basically by multiplying with the Heaviside theta function, which forces the value of $\chi(t)$ to be zero at times $t < 0$.

At this point it is expedient to introduce Fourier transformations with respect to time. The usual convention in physics is

$$f(\omega) = \int e^{i\omega t} f(t) dt \quad \Leftrightarrow \quad f(t) = \frac{1}{2\pi} \int d\omega e^{-i\omega t} f(\omega). \tag{10.21}$$

Now, the last form in (10.18) is a convolution, and the Fourier transformation converts a convolution into a product. In another piece of dubious but obvious notation we then have the main result of the present development,

$$\langle B(\omega) \rangle = \sum_j \chi_{BA_j}(\omega) h_j(\omega). \tag{10.22}$$

Let us consider an example with a monochromatic driving field. As long as the theory is linear, as it is here, one may without ambiguity use complex values for physical quantities, even if real; one may adopt, for instance, the convention that the real part of the complex quantity is the value of the real physical quantity. We take only one coupling field that we write

$$h(t) = h_0 e^{-i\omega_0 t}. \tag{10.23}$$

Thus we have $h(\omega) = 2\pi\delta(\omega - \omega_0)$, $\langle B(\omega) \rangle = 2\pi\chi_{BA}(\omega)\delta(\omega - \omega_0)$, and

$$\langle B(t) \rangle = \frac{1}{2\pi} \int d\omega e^{-i\omega t} B(\omega) = \chi_{BA}(\omega_0) h_0 e^{-i\omega_0 t}. \quad (10.24)$$

It is a built-in property of linear response theory, and all of physics in the linear regime, that a monochromatic excitation with the amplitude h_0 will elicit a monochromatic stationary response, here with the amplitude $\chi_{BA}h_0$. On the other hand, while $\langle B(\omega) \rangle = \chi_{BA}(\omega)h(\omega)$ holds true, in general there is no susceptibility $X_{BA}(t)$ such that $\langle B(t) \rangle = X_{BA}(t)h(t)$. The parallels of these observations to the usual way of defining frequency dependent dielectric constants, refractive indices, susceptibilities, etc., in E&M should be obvious.

Mathematical properties of response functions

The response functions we have defined have a rich variety of general mathematical properties that derive from the physical symmetries and other physical requirements of the problem. Moreover, somewhat unfortunately, one has to be familiar with these mathematical properties to apply the response functions in a meaningful way. We therefore simply roll up our sleeves.

- $\chi''(t)$ is imaginary.

It should be noted, again, that for two hermitian operators A and B the commutator satisfies $(i[A, B])^\dagger = i[A, B]$. Therefore the expectation value $\langle i[A, B] \rangle$ is real, and the expectation value $\langle [A, B] \rangle$ is purely imaginary. All of the response functions $\chi''_{AB}(t)$ are purely imaginary.

- *Time reversal symmetry.*

Consider a state vector in the position representation, i.e., a wave function $\psi(r)$, where r may stand for the coordinates of a great many of particles. The *time reversal* operator T takes the complex conjugate of the wave function: $T\psi = \psi^*$. For multicomponent wave functions involving the spin, one also reverses the spin. Taking a complex conjugate twice has no effect on the wave function, so that $TT = I$. While the algebra involving time reversal is similar to the algebra of linear operators, one has to be careful since the time reversal operator is not linear: $T(a\psi) = a^*T\psi$.

Most observables behave neatly under time reversal: $TAT = \epsilon_A A$ with $\epsilon_A = \pm 1$, the sign depending on the operator. To wit, here are two examples:

$$T\mathbf{r}T\psi = T\mathbf{r}\psi^* = (\mathbf{r}\psi^*)^* = \mathbf{r}\psi; \quad \epsilon_{\mathbf{r}} = 1; \quad (10.25)$$

$$T\mathbf{p}T\psi = T\left(\frac{\hbar}{i}\frac{\partial}{\partial \mathbf{r}}\psi^*\right) = \left(\frac{\hbar}{i}\frac{\partial}{\partial \mathbf{r}}\psi^*\right)^* = -\frac{\hbar}{i}\frac{\partial}{\partial \mathbf{r}}\psi = -\mathbf{p}\psi; \quad \epsilon_{\mathbf{p}} = -1. \quad (10.26)$$

From now on we assume that the time reversal symmetries of both the unperturbed Hamiltonian and of the unperturbed states are even: $\epsilon_{H_0} = 1$,

$\epsilon_{\tilde{\rho}_0} = 1$. We then have the following argument:

$$\begin{aligned} T\tilde{A}(t)T &= Te^{i\frac{H_0 t}{\hbar}} A e^{-i\frac{H_0 t}{\hbar}} T = (Te^{i\frac{H_0 t}{\hbar}} T)(TAT)(Te^{-i\frac{H_0 t}{\hbar}} T) \\ &= \epsilon_A e^{-i\frac{H_0 t}{\hbar}} A e^{i\frac{H_0 t}{\hbar}} = \epsilon_A \tilde{A}(-t). \end{aligned} \quad (10.27)$$

We may therefore write

$$\tilde{A}(-t) = \epsilon_A T \tilde{A}(t) T. \quad (10.28)$$

This, in turn, opens the way for the following development,

$$\begin{aligned} \chi''_{AB}(-t) &= \frac{1}{2\hbar} \langle \tilde{A}(-t) \tilde{B}(0) - \dots \rangle_0 = \epsilon_A \epsilon_B \langle T \tilde{A}(t) T T \tilde{B}(0) T - \dots \rangle_0 \\ &= \frac{1}{2\hbar} \epsilon_A \epsilon_B \text{Tr} \{ \tilde{\rho}_0 T \tilde{A}(t) \tilde{B}(0) T - \dots \} \\ &= \frac{1}{2\hbar} \epsilon_A \epsilon_B \text{Tr} \{ T (T \tilde{\rho}_0 T) T T \tilde{A}(t) \tilde{B}(0) T - \dots \} \\ &= \frac{1}{2\hbar} \epsilon_A \epsilon_B \text{Tr} \{ T \tilde{\rho}_0 \tilde{A}(t) \tilde{B}(0) T - \dots \}. \end{aligned} \quad (10.29)$$

But, we have in the position representation

$$\begin{aligned} (\psi, T\phi) &= \int dr \psi^*(T\phi) = \int dr \psi^* \phi^* = \left(\int dr \psi \phi \right)^* \\ &= \left(\int dr (T\psi)^* \phi \right)^* = (T\psi, \phi)^*; \\ \text{Tr}(T\mathcal{O}T) &= \sum_n (\psi_n, T\mathcal{O}T\psi_n) = \sum_n (T\psi_n, \mathcal{O}T\psi_n)^* \\ &= \left[\sum_n (T\psi_n, \mathcal{O}T\psi_n) \right]^* = (\text{Tr } \mathcal{O})^*. \end{aligned} \quad (10.30)$$

The last step is because, evidently, if $\{\psi_n\}$ is an orthonormal basis, then so is the set of functions $\{T\psi_n\}$ made of the complex conjugates of the original functions. Continuing where (10.29) left off, we then have

$$\chi''_{AB}(-t) = \frac{1}{2\hbar} \epsilon_A \epsilon_B \left(\text{Tr} \{ \tilde{\rho}_0 [\tilde{A}(t), \tilde{B}(0)] \} \right)^*, \quad (10.31)$$

or, since the quantity inside the complex conjugate is purely imaginary,

$$\chi''_{AB}(-t) = -\epsilon_A \epsilon_B \chi''_{AB}(t). \quad (10.32)$$

In (10.32) we have the main result from our analysis of time reversal. More generally, understanding time reversal is important when you find yourself studying friction and other irreversible phenomena. For instance, if for $t > 0$ your model says $\chi''(t) \propto e^{-\gamma t}$, then for $t < 0$ it must be that $\chi''(t) \propto e^{\gamma t}$.

- *Causality.*

By the very definition (10.20), the response function $\chi(t)$ cuts off at negative times. From (10.18) one can then see that the induced response at time t depends on the driving fields $h_j(t')$ only at times $t' \leq t$, as it should to satisfy causality. As simple as these observations are, causality nevertheless imposes much mathematical structure on the theory of linear response.

Let us begin with a few tricks with Fourier transformations. We have

$$\begin{aligned}\chi_{AB}(\omega) &= \int_{-\infty}^{\infty} dt e^{i\omega t} \chi_{AB}(t) = 2i \int_0^{\infty} dt \chi_{AB}''(t) \\ &= 2i \int_0^{\infty} dt e^{i\omega t} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' e^{-i\omega' t} \chi_{AB}''(\omega') \right) \\ &= \frac{i}{\pi} \int_{-\infty}^{\infty} d\omega' \chi_{AB}''(\omega') \int_0^{\infty} dt e^{i(\omega - \omega')t}. \quad (10.33)\end{aligned}$$

The second integral is not convergent, but it can be handled with a neat trick: Take a number $\eta > 0$, calculate

$$\begin{aligned}\int_0^{\infty} dt e^{i(\omega - \omega')t} &= \int_0^{\infty} dt e^{[i(\omega - \omega') - \eta]t} = \frac{1}{i(\omega' - \omega) + \eta} \\ &= \frac{\eta}{(\omega' - \omega)^2 + \eta^2} - i \frac{(\omega' - \omega)}{(\omega' - \omega)^2 + \eta^2}, \quad (10.34)\end{aligned}$$

and let $\eta \rightarrow 0$ at the end of the calculations. A number η that behaves in this way is conventionally denoted by $\eta = 0+$. Physicists say that we have added a convergence factor to the divergent integral. This is also correct mathematics, albeit way beyond our league; it works for objects called tempered distributions, which include Fourier transformations of functions of time that grow with the argument t more slowly than any exponential as $t \rightarrow \pm\infty$.

Let us now consider (10.34) in the limit $\eta \rightarrow 0$. The first term on the second line is a Lorentzian with the integral over ω' (or ω) equal to π regardless of the value of $\eta > 0$, and it peaks more and more sharply as $\eta \rightarrow 0$. We therefore write

$$\frac{\eta}{(\omega' - \omega)^2 + \eta^2} = \pi \delta(\omega' - \omega), \quad \eta = 0+. \quad (10.35)$$

For the second term we introduce the placeholder

$$\frac{(\omega' - \omega)}{(\omega' - \omega)^2 + \eta^2} = \mathcal{P} \frac{1}{\omega' - \omega}, \quad \eta = 0+. \quad (10.36)$$

The \mathcal{P} denotes “principal value integral”; when the right-hand side appears inside an integral, it is replaced by the left-hand side, the integral is done, and finally the limit $\eta \rightarrow 0$ is taken. There are a couple of other

definitions of principal value integral in mathematics. In practice they all mean the same.

Putting the preceding considerations together we have

$$\chi_{AB}(\omega) = \frac{1}{\pi} \mathcal{P} \int d\omega' \frac{\chi''_{AB}(\omega')}{\omega' - \omega} + i \chi''_{AB}(\omega). \quad (10.37)$$

To make further progress, for the sake of the example let us assume that the time reversal symmetries of the observables A and B are the same. We then have $\chi''_{AB}(t) = -\chi''_{AB}(-t)$, and $\chi''_{AB}(t)$ is imaginary as always. These observations precipitate the following chain of manipulations,

$$\begin{aligned} [\chi''_{AB}(\omega)]^* &= \left[\int dt e^{i\omega t} \chi''_{AB}(t) \right]^* = - \int dt e^{-i\omega t} \chi''_{AB}(t) \\ &= - \int dt \chi''_{AB}(-t) e^{i\omega t} = \int dt \chi''_{AB}(t) e^{i\omega t} \\ &= \chi''_{AB}(\omega). \end{aligned} \quad (10.38)$$

The second line comes from the change of integration variable $t \rightarrow -t$. The response function $\chi''_{AB}(\omega)$ is real!

This allows us to develop (10.37) further. Everything on the right-hand side except the explicit imaginary unit is real, so that we can write the response function in the form

$$\chi_{AB}(\omega) = \chi'_{AB}(\omega) + i \chi''_{AB}(\omega). \quad (10.39)$$

where both $\chi'_{AB}(\omega)$ and our old acquaintance $\chi''_{AB}(\omega)$ are real. Moreover, we have

$$\chi''_{AB}(\omega) = \Im[\chi_{AB}(\omega)]. \quad (10.40)$$

In addition, the real part and the imaginary part of the response function are related by

$$\chi'_{AB}(\omega) = \frac{1}{\pi} \mathcal{P} \int d\omega' \frac{\chi''_{AB}(\omega')}{\omega' - \omega}, \quad (10.41)$$

so that the imaginary part of the response function determines the real part. In fact, it works the other way round, too, so that

$$\chi''_{AB}(\omega) = -\frac{1}{\pi} \mathcal{P} \int d\omega' \frac{\chi'_{AB}(\omega')}{\omega' - \omega}, \quad (10.42)$$

Equations (10.41) and (10.42) are known as the *Kramers-Kronig* relations.

For our remaining causality related item we write (10.33) for an as yet nonzero value of η , so that we have

$$\chi_{AB}(\omega) = \frac{i}{\pi} \int_{-\infty}^{\infty} d\omega' \frac{\chi''_{AB}(\omega')}{i(\omega' - \omega) + \eta}. \quad (10.43)$$

Now, the function of ω inside the integral is analytic for every ω in the upper half of the complex plane, so that the integral is likewise. In fact, this should be the case. Namely to obtain $\chi_{AB}(t)$ from the inverse Fourier transformation,

$$\chi_{AB}(t) = \frac{1}{2\pi} \int d\omega e^{-i\omega t} \chi_{AB}(\omega), \quad (10.44)$$

for negative times, one would have to close the path of the integration through the upper half of the complex plane. The result should be identically zero; and it is if $\chi_{AB}(\omega)$ is analytic in the upper half of the complex plane.

- *Time derivative operators.*

In an unperturbed system with $h_j(t) \equiv 0$ the interaction picture is in fact the same as the Heisenberg picture. Then the operators depend on time, the time derivative being

$$i\hbar \frac{d\tilde{A}}{dt} = [\tilde{A}, H_0] \equiv [\tilde{A}, \tilde{H}_0]. \quad (10.45)$$

Encouraged by this observation, for the present purposes we *define* the time derivative of an arbitrary system observable A as

$$\dot{A} \equiv \frac{1}{i\hbar} [A, H_0]. \quad (10.46)$$

Since the commutator is invariant in a transformation to either the interaction picture or to the Heisenberg picture, this definition does not depend on which of these pictures is used, and in the Heisenberg picture it actually coincides with what is known to be the time derivative of the operator A . Hence, we use it to represent the time derivative of the quantity A also in the Schrödinger picture.

For instance, suppose we have the standard one-particle Hamiltonian

$$H_0 = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}). \quad (10.47)$$

The time derivative corresponding to the position operator is

$$\dot{\mathbf{r}} = \frac{1}{2\hbar} \left[\mathbf{r}, \frac{\mathbf{p}^2}{2m} \right] = \frac{\mathbf{p}}{m}, \quad (10.48)$$

which makes preeminent sense.

In light of the definition, we start our analysis as follows,

$$\chi''_{AB}(t) = \frac{1}{2\hbar} \text{Tr} \left\{ \tilde{\rho}_0 [\tilde{A}, \tilde{B}(0)] \right\} = \frac{1}{2\hbar} \text{Tr} \left\{ \tilde{\rho}_0 \left[\frac{1}{i\hbar} [\tilde{A}(t), \tilde{H}_0], \tilde{B}(0) \right] \right\}. \quad (10.49)$$

The definition of the interaction picture gives

$$\frac{1}{i\hbar}[\tilde{A}(t), \tilde{H}_0] = \frac{1}{i\hbar}[\tilde{A}(t), H_0] = \frac{\partial}{\partial t}\tilde{A}(t), \quad (10.50)$$

so that the immediate finding is

$$\begin{aligned} \chi''_{\tilde{A}\tilde{B}}(t) &= \frac{1}{2\hbar} \text{Tr} \left\{ \tilde{\rho}_0 \left[\frac{\partial}{\partial t} \tilde{A}(t), \tilde{B}(0) \right] \right\} \\ &= \frac{\partial}{\partial t} \left(\frac{1}{2\hbar} \text{Tr} \left\{ \tilde{\rho}_0 [\tilde{A}(t), \tilde{B}(0)] \right\} \right) = \frac{\partial}{\partial t} \chi''_{AB}(t). \end{aligned} \quad (10.51)$$

In the same way one derives

$$\chi''_{\tilde{A}\tilde{B}}(t) = -\frac{\partial}{\partial t} \chi''_{AB}(t). \quad (10.52)$$

The Fourier transformation maps time derivative to the multiplicative factor $-i\omega$, so we have, for instance, the results

$$\chi''_{\tilde{A}\tilde{B}}(\omega) = -i\omega \chi''_{AB}(\omega), \quad \chi''_{\tilde{A}\dot{\tilde{B}}}(\omega) = i\omega \chi''_{AB}(\omega), \quad \chi''_{\dot{\tilde{A}}\tilde{B}}(\omega) = \omega^2 \chi''_{AB}(\omega). \quad (10.53)$$

10.2 Correlation functions

Let us look at the position of a mirror in an apparatus that is supposed to measure small displacements using laser interferometry. The mirror is subject to thermal fluctuations, and even quantum fluctuations. With the convention that the intended position of the mirror is $x = 0$, the fluctuations are characterized by $(\Delta x)^2 = \langle x^2 \rangle = \langle \tilde{x}(t) \tilde{x}(t) \rangle_0$.

However, if the mirror is set up to observe displacements in a certain frequency range, the instantaneous position is not really an issue. What would then be the proper way to characterize the fluctuations? The usual approach runs as follows: In your run-of-the mill system that is invariant under time translations the function $S_{xx}(t, t') = \langle \tilde{x}(t), \tilde{x}(t') \rangle$ is a function of the time difference $t - t'$ only, $S_{xx}(t - t')$, and one may take the Fourier transform of this function $S_{xx}(\omega)$. Conversely, the inverse Fourier transform gives

$$S_{xx}(t - t') = \frac{1}{2\pi} \int d\omega e^{-i\omega(t-t')} S_{xx}(\omega). \quad (10.54)$$

Put in this way, the squared fluctuations are

$$(\Delta x)^2 = S_{xx}(t = 0) = \int d\omega \frac{S_{xx}(\omega)}{2\pi}. \quad (10.55)$$

The quantity inside the integral behaves like the square of the amplitude of random fluctuations per unit frequency. The fluctuations affecting a detector with the pass band $\Delta\omega$ would then be

$$(\Delta x)_{\Delta\omega}^2 = \int_{\Delta\omega} d\omega \frac{S_{xx}(\omega)}{2\pi}. \quad (10.56)$$

The above reasoning works in practice, and could also be derived from physical/mathematical models of a spectrum analyzer. For the present purposes, though, the point of the argument is that it focuses our attention on the *correlation function* $S_{xx}(t) = \langle \tilde{x}(t)\tilde{x}(0) \rangle_0$, which for obvious reasons is called a *two-time* expectation value. Correlation functions are the theme of the present section.

10.2.1 Scattering measures correlation functions

Our immediate focus is on the observation that scattering experiments measure correlation functions. Since lots of experiments that probe the properties of all sorts of systems are scattering experiments, we have here the general framework for the discussion of many an experiment. In condensed matter physics the resulting correlation functions are often called *structure factors*, and the juxtaposition of theory and experiment is in the comparison of the respective structure factors.

Thus, consider some system under study, characterized by the Hamiltonian H_0 . We use a set of positions \mathbf{r}_j , $j = 1, \dots, N$ as the degrees of freedom of the system, although the position representation is purely for the convenience of our example. The system is initially in some energy eigenstate $|i\rangle$. In a scattering experiment a plane wave of some incoming particle with the wave vector \mathbf{k} and energy $\epsilon_{\mathbf{k}}$ is incident on the system. After the scattering one sees the particle coming out with the wave vector \mathbf{k}' and the energy $\epsilon_{\mathbf{k}'}$, so that one has the changes of wave vector and energy

$$\Delta\mathbf{k} = \mathbf{k} - \mathbf{k}', \quad \Delta\epsilon = \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}}'. \quad (10.57)$$

The counterintuitive signs are deliberate; when defined in this way, $\hbar\Delta\mathbf{k}$ is the change of the momentum of the system upon scattering, and likewise for energy. In the process the state of the system may have changed as well, to $|f\rangle$. Scattering takes place as a result of an interaction between the projectile and the target, written schematically in terms of the position of the projectile \mathbf{r} and of the target degrees of freedom \mathbf{r}_j as

$$H_I = V(\mathbf{r}, \{\mathbf{r}_j\}). \quad (10.58)$$

For instance, one could use neutron scattering to study the positions of the atoms in a solid. Neutrons only couple to the nuclei, so the interaction Hamiltonian could be of the form

$$H_I = \lambda \sum_j \delta(\mathbf{r} - \mathbf{r}_j), \quad (10.59)$$

where \mathbf{r}_j are the positions of the nuclei and λ characterizes the strength of the neutron-nucleus coupling. When viewed as a function of the neutron coordinate \mathbf{r} , this interaction Hamiltonian is basically the density of the nuclei, or of the atoms.

By the Golden Rule the rate for scattering with the specified initial and final states is

$$R_{i\mathbf{k} \rightarrow f\mathbf{k}'} \propto \delta[\Delta\epsilon - (\epsilon_f - \epsilon_i)] |\langle \mathbf{k}' | \langle f | H_I | i \rangle | \mathbf{k} \rangle|^2. \quad (10.60)$$

Using position representation for the coordinate of the projectile, we have the matrix element

$$\langle \mathbf{k}' | \langle f | H_I | i \rangle | \mathbf{k} \rangle \propto \int d^3r e^{i\Delta\mathbf{k} \cdot \mathbf{r}} \langle f | H_I(\mathbf{r}) | i \rangle, \quad (10.61)$$

and the delta function in energy may be written

$$\delta[\Delta\epsilon - (\epsilon_f - \epsilon_i)] \propto \int dt e^{\frac{i}{\hbar}[\Delta\epsilon - (\epsilon_f - \epsilon_i)]}. \quad (10.62)$$

These observations lead to a form of the scattering rate

$$\begin{aligned} R_{i\mathbf{k} \rightarrow f\mathbf{k}'} &\propto \int dt e^{\frac{i}{\hbar}[\Delta\epsilon - (\epsilon_f - \epsilon_i)]} \\ &\times \int d^3r d^3r' e^{-i\Delta\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \langle i | H_I(\mathbf{r}) | f \rangle \langle f | H_I(\mathbf{r}') | i \rangle. \end{aligned} \quad (10.63)$$

But we may write

$$\begin{aligned} e^{-\frac{i}{\hbar}(\epsilon_f - \epsilon_i)} \langle i | H_I(\mathbf{r}) | f \rangle \langle f | H_I(\mathbf{r}') | i \rangle &= \langle i | e^{i\frac{H_0 t}{\hbar}} H_I(\mathbf{r}) e^{-i\frac{H_0 t}{\hbar}} | f \rangle \langle f | H_I(\mathbf{r}') | i \rangle \\ &= \langle i | \tilde{H}_I(\mathbf{r}, t) | f \rangle \langle f | \tilde{H}_I(\mathbf{r}', 0) | i \rangle, \end{aligned} \quad (10.64)$$

which gives

$$R_{i\mathbf{k} \rightarrow f\mathbf{k}'} \propto \int dt d^3r d^3r' e^{\frac{i}{\hbar}\Delta\epsilon t - i\Delta\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \langle i | \tilde{H}_I(\mathbf{r}, t) | f \rangle \langle f | \tilde{H}_I(\mathbf{r}', 0) | i \rangle. \quad (10.65)$$

Ordinarily in a scattering experiment the final state of the system is not observed so that it should be summed over, conveniently using the identity

$$\sum_f |f\rangle \langle f| = I_S. \quad (10.66)$$

Moreover, the exact initial state $|i\rangle$ is not known either. Let us assume, though, that before the scattering event the density operator was diagonal in the energy eigenstates $\{|i\rangle\}$, with the probabilities p_i . We just have to average the results for the states $|i\rangle$ using the probabilities p_i . The result is the rate for the scattering of the projectile from the initial plane wave \mathbf{k} to the final plane wave \mathbf{k}' as follows,

$$R_{\mathbf{k} \rightarrow \mathbf{k}'} \propto \int dt d^3r d^3r' e^{\frac{i}{\hbar}\Delta\epsilon t - i\Delta\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \langle \tilde{H}_I(\mathbf{r}, t) \tilde{H}_I(\mathbf{r}', 0) \rangle_0, \quad (10.67)$$

where the expectation value is for the pre-scattering “unperturbed” state of the system.

Frequently the system is, or for the purposes of the scattering experiment may be considered to be, translationally invariant. In this case the correlation function is obviously a function of the differences of the positions only, so that one may write

$$\begin{aligned} R_{\mathbf{k} \rightarrow \mathbf{k}'} &\propto \int dt d^3r d^3r' e^{\frac{i}{\hbar} \Delta \epsilon t - i \Delta \mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \langle \tilde{H}_I(\mathbf{r} - \mathbf{r}', t) \tilde{H}_I(0, 0) \rangle_0 \\ &\propto V \int dt d^3r e^{\frac{i}{\hbar} \Delta \epsilon t - i \Delta \mathbf{k} \cdot \mathbf{r}} \langle \tilde{H}_I(\mathbf{r}, t) \tilde{H}_I(0, 0) \rangle_0. \end{aligned} \quad (10.68)$$

This is a Fourier transformation of the correlation function $S_{H_I(\mathbf{r}), H_I(0)}(t)$. The Fourier variables are essentially the momentum and the energy that the target receives from the projectile in the scattering. In our example of neutron scattering the operator $H_I(\mathbf{r})$ is basically the density of the atoms, so that we would be measuring the two-time correlation function of the density of the atoms at two different positions.

10.3 Fluctuation-dissipation theorem

Remarkably, linear response in a system also determines the fluctuations, and vice versa. The attendant fluctuation-dissipation theorem is the topic of the present chapter.

From now on we take the unperturbed state to be the thermal density operator

$$\tilde{\rho}_0 = \rho_0 = Z^{-1} e^{-\beta H_0}. \quad (10.69)$$

With this agreed upon, a few simple steps give

$$\begin{aligned} \langle \tilde{A}(0) \tilde{B}(t) \rangle_0 &= Z^{-1} \text{Tr} \left(e^{-\beta H_0} A e^{i \frac{H_0 t}{\hbar}} B e^{-i \frac{H_0 t}{\hbar}} \right) = Z^{-1} \text{Tr} \left(A e^{i \frac{H_0 t}{\hbar}} B e^{-i \frac{H_0(t - i\hbar\beta)}{\hbar}} \right) \\ &= Z^{-1} \text{Tr} \left(e^{i \frac{H_0 t}{\hbar}} B e^{-i \frac{H_0(t - i\hbar\beta)}{\hbar}} A \right) \\ &= Z^{-1} \text{Tr} \left(e^{-\beta H_0} e^{i \frac{H_0(t - i\hbar\beta)}{\hbar}} B e^{-i \frac{H_0(t - i\hbar\beta)}{\hbar}} A \right) \\ &= \langle \tilde{B}(t - i\hbar\beta) \tilde{A}(0) \rangle_0. \end{aligned} \quad (10.70)$$

We therefore have

$$\begin{aligned} \chi''_{BA}(t) &= \frac{1}{2\hbar} \langle \tilde{B}(t) \tilde{A}(0) - \tilde{A}(0) \tilde{B}(t) \rangle_0 \\ &= \frac{1}{2\hbar} [\langle \tilde{B}(t) \tilde{A}(0) \rangle_0 - \langle \tilde{B}(t - i\hbar\beta) \tilde{A}(0) \rangle_0], \end{aligned} \quad (10.71)$$

and

$$\begin{aligned} \chi''_{BA}(\omega) &= \frac{1}{2\hbar} \int dt e^{i\omega t} [\langle \tilde{B}(t) \tilde{A}(0) \rangle_0 - \langle \tilde{B}(t - i\hbar\beta) \tilde{A}(0) \rangle_0] \\ &= \frac{1}{2\hbar} S_{BA}(\omega) - \frac{1}{2\hbar} \int dt e^{i\omega(t + i\hbar\beta)} \langle \tilde{B}(t) \tilde{A}(0) \rangle_0 \\ &= \frac{1}{2\hbar} (1 - e^{-\hbar\omega\beta}) S_{BA}(\omega). \end{aligned} \quad (10.72)$$

The step to the second line involves the change of the integration variable $t - i\hbar\beta \rightarrow t$. In other words, the path of the integral that used to be the real axis is moved up in the complex plane by $\hbar\beta$. This is, in fact, not allowed without introducing extra terms in the result if it so happens that the integrand $\langle \tilde{B}(t)\tilde{A}(0) \rangle_0$ has singularities between the old and the new path. We do not discuss this possibility further, other than note that such a problem, if it were to come up, say, as a result of heavy-handed modeling, would likely cause obvious trouble all the way to the final result.

In summary, we have the fluctuation-dissipation theorem

$$\chi''_{BA}(\omega) = \frac{1 - e^{-\hbar\omega\beta}}{2\hbar} S_{BA}(\omega). \quad (10.73)$$

Many a time it is used to connect results from a basically classical analysis of the response to predict thermal fluctuations. In such a case one takes the classical limit $\hbar \rightarrow 0$ of the fluctuation-dissipation theorem, which is perfectly well defined and says

$$\chi''_{BA}(\omega) = \frac{\omega}{2kT} S_{BA}(\omega). \quad (10.74)$$

Example on Brownian motion

As an illustration of the machinery having to do with response and fluctuations we discuss Brownian motion of a test particle in a viscous liquid. The model is classical, and without any essential restriction also one-dimensional, so that the equation of motion for the position of the particle reads

$$\ddot{x} + 2\gamma\dot{x} = F(t)/m. \quad (10.75)$$

To begin with, let us write a monochromatic driving force $F(t) = F_0 e^{-i\omega_0 t}$. The ansatz for the stationary response $x(t) = x_0 e^{-i\omega_0 t}$ succeeds, and gives a solution to the equation of motion

$$x(t) = -\frac{F_0/m}{\omega_0^2 + 2i\gamma\omega_0} e^{-i\omega_0 t}. \quad (10.76)$$

Next, as we have already mentioned, adding a position independent force corresponds to adding an interaction Hamiltonian

$$H_I = -xF(t). \quad (10.77)$$

By linear response theory, we then have

$$\langle x(\omega) \rangle = \chi_{xx}(\omega) F(\omega), \quad (10.78)$$

and, again as per our earlier discussion, a force of the form $F(t) = F_0 e^{-i\omega_0 t}$ produces the response $\langle x(t) \rangle = F_0 \chi_{xx}(\omega_0) e^{-i\omega_0 t}$. Comparison with (10.76) shows that the response function is

$$\chi_{xx}(\omega) = -\frac{1}{m\omega[\omega + 2i\gamma]}. \quad (10.79)$$

The absorptive part of the response reads

$$\chi''_{xx}(\omega) = \Im[\chi_{xx}(\omega)] = \frac{2\gamma}{m\omega[\omega^2 + \gamma^2]}. \quad (10.80)$$

Given that we have used a classical equation of motion to obtain the response, it is appropriate to use the classical version of the fluctuation-dissipation theorem as well. We find the position correlation function

$$S_{xx}(\omega) = \frac{2kT}{\omega} \chi''_{xx}(\omega) = \frac{4\gamma kT}{m\omega^2[\omega^2 + \gamma^2]}. \quad (10.81)$$

An attempt to transform back to the time variable produces an ill-defined divergent integral. This should not come as a surprise; in the spatially homogeneous system the fluctuations in position are infinite, as the particle can be anywhere.

Let us instead ask about velocity fluctuations. As usual, we have

$$\langle \dot{x}^2 \rangle = S_{\dot{x}\dot{x}}(t=0). \quad (10.82)$$

Time derivatives of the operators transform into time derivatives and multiplications with $\pm i\omega$ just like in the case of the response function χ'' . The velocity correlation function is therefore

$$S_{\dot{x}\dot{x}}(\omega) = \omega^2 S_{xx}(\omega). \quad (10.83)$$

It remains to put all of our arguments together:

$$\begin{aligned} \langle \dot{x}^2 \rangle &= S_{\dot{x}\dot{x}}(t=0) = \frac{1}{2\pi} \int d\omega S_{\dot{x}\dot{x}}(\omega) = \frac{1}{2\pi} \int d\omega \omega^2 S_{xx}(\omega) \\ &= \frac{kT}{m\pi} \int d\omega \frac{2\gamma}{\omega^2 + (2\gamma)^2} = \frac{kT}{2m}. \end{aligned} \quad (10.84)$$

The expectation value of the kinetic energy is $E_k = \frac{1}{2}m\langle \dot{x}^2 \rangle = \frac{1}{2}kT$, as expected.

We have used heavy ammo for a problem that is trivially solvable in many other ways, but now we also have access to aspects that are not so immediately obvious. For instance, suppose that for one reason or the other there is a small signal oscillating at the frequency ω_0 in the velocity of the particle(s), which we are looking for. Let us do so with the aid of an (idealized) bandpass filter that only lets through the frequencies in the interval $(\omega_0 - \frac{1}{2}\delta, \omega_0 + \frac{1}{2}\delta)$, with $\delta \ll \gamma$. The thermal fluctuations that the detection has to contend with are

$$\langle \dot{x}^2 \rangle_\delta = \frac{kT}{m\pi} \int_{\omega_0 - \frac{1}{2}\delta}^{\omega_0 + \frac{1}{2}\delta} d\omega \frac{2\gamma}{\omega^2 + (2\gamma)^2} \simeq \frac{4\gamma\delta}{\pi[\omega_0^2 + (2\gamma)^2]} \frac{kT}{2m}, \quad (10.85)$$

possibly *much* smaller than the full thermal noise.

10.4 Exercises

10.1. Let us study a system coupled to an external field $h(t)$, with the Hamiltonian $H = H_0 - h(t)A$.

(a) In the interaction picture, solve the equation of motion of the density operator to *second* order in the perturbation.

(b) Assume that the perturbation is monochromatic, $h(t) = h_0 \cos \omega t$, and take the unperturbed density operator to be thermal. With the aid of a suitable average over the period of the perturbation, show that the rate of change of the expectation value of the unperturbed Hamiltonian H_0 equals $\frac{1}{2} h_0^2 \omega \chi''_{AA}(\omega)$.

(c) So why is $\chi''_{AA}(\omega)$ is called the *dissipative* part of the response function $\chi_{AA}(\omega)$?

WARNING: This one is a mighty grind.

10.2.

(a) Given an eigenbasis of the Hamiltonian H_0 , the states $|n\rangle$ with the corresponding eigenfrequencies ω_n , and a thermal unperturbed state $\tilde{\rho}_0$, show that the Fourier transform of the correlation function $S_{AA}(t) = \langle \tilde{A}(t) \tilde{A}(0) \rangle_0$ equals

$$S_{AA}(\omega) = \frac{2\pi}{Z} \sum_{n,m} e^{-\beta \hbar \omega_n} \delta(\omega + \omega_n - \omega_m) |\langle n|A|m\rangle|^2.$$

This is nonnegative, and as such qualifies as a frequency distribution of fluctuations. Similarly, $\bar{S}_{AA}(t) = \langle \tilde{A}(0) \tilde{A}(t) \rangle_0$ transforms into

$$\bar{S}_{AA}(\omega) = \frac{2\pi}{Z} \sum_{n,m} e^{-\beta \hbar \omega_m} \delta(\omega + \omega_n - \omega_m) |\langle n|A|m\rangle|^2.$$

(b) Show that the response function $\chi''_{AA}(t) = [S_{AA}(t) - \bar{S}_{AA}(t)]/2\hbar$ satisfies the fluctuation-dissipation theorem

$$\chi''_{AA}(\omega) = \frac{1 - e^{-\beta \hbar \omega}}{2\hbar} S_{AA}(\omega).$$

10.3. Consider a damped harmonic oscillator, $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = F(t)/m$. Quantum mechanically, the driving force corresponds to adding a term to the Hamiltonian of the form $H_I = -x F(t)$, so that a comparison between the response of the classical oscillator to the force $F(t) = F_0 e^{-i\omega t}$ and the general linear response theory gives in the usual way the response function

$$\chi_{xx}(\omega) = \frac{1}{m[(\omega_0^2 - \omega^2) - 2i\gamma\omega]}.$$

(a) Show that in the classical limit when kT/\hbar is the largest frequency parameter in the problem, the root-mean square fluctuations of position are

$$(\Delta x)^2 = \int d\omega \frac{2\gamma kT}{\pi m[(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]}.$$

(b) The nasty integral evaluates to $(\Delta x)^2 = kT/(m\omega_0^2)$, even independently of the value of the damping constant γ . The result may be surprising, but it makes sense. Why?

10.4. In a gravitational-wave detector a mirror mount with mass $m = 1$ kg is hung on a 0.5 m wire, which makes a pendulum with oscillation frequency $\omega_0 = 4.4 \text{ s}^{-1}$. Let us call the coordinate of small displacements of the pendulum from equilibrium x . The mount is hung in high vacuum, so that the Q value of the pendulum is 10^6 and the damping constant of pendulum amplitude thus is $\gamma = 2.2 \times 10^{-6} \text{ s}^{-1}$. When a force $F(t)$ drives the pendulum, the corresponding “interaction energy” is $H' = -F(t)x$. On the other hand, the equation of motion for the driven pendulum is

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = F(t)/m. \quad (10.86)$$

(a) Argue from the theory of linear response that a force $F(t) = F_0 e^{-i\omega t}$ elicits the stationary response $x(t) = F_0 \chi_{xx}(\omega) e^{-i\omega t}$.

(b) By comparing the result of part (a) with the proper stationary solution of Eq. (10.86), find the response function $\chi_{xx}(\omega)$ of the mirror mount.

(c) Now suppose that we are after a burst of gravitational radiation at the frequency $\omega_g = 2\pi \times 500 \text{ Hz}$, which lasts for about 10 s. The radiation therefore appears in a band of width $\Delta\omega \simeq 10^{-1} \text{ s}^{-1}$. At room temperature, what is the amplitude of the thermal fluctuations of the mirror that the experiment must contend with?

10.5. In P. 9.3 we, in fact, found the low-frequency response function for the order parameter $\phi(x)$ in the form

$$\chi_{\phi(x)\phi(x')}(\omega = 0) = \frac{\xi}{2\sqrt{\gamma}} \exp\left(-\frac{|x - x'|}{\xi}\right); \quad \xi = \left(\frac{2\alpha_2}{\gamma}(T - T_c)\right)^{-1/2}.$$

This is real, $\chi_{\phi(x)\phi(x')}(\omega = 0) = \chi'_{\phi(x)\phi(x')}(\omega = 0)$, but there must be a corresponding imaginary part to the response function that by the Kramers-Kronig relation satisfies

$$\chi'_{\phi(x)\phi(x')}(\omega = 0) = \frac{1}{\pi} \int d\omega \frac{\chi''_{\phi(x)\phi(x')}(\omega)}{\omega}.$$

By the numerous symmetries that the response functions have $\chi''_{\phi(x)\phi(x')}(\omega) \rightarrow 0$ as $\omega \rightarrow 0$, so that no principal value integral is needed here.

(a) Show that in the classical limit the equal-time correlation function reads

$$S_{\phi(x)\phi(x')}(t=0) = kT \chi_{\phi(x)\phi(x')}(\omega=0).$$

This means that there are blobs in the order parameter with the size on the order ξ and maximum value of the square of the order parameter proportional to ξ . The present results are for $T > T_c$, but essentially the same with $T - T_c$ replaced by $|T - T_c|$ also obtains below the critical temperature, as may be seen by linearizing Landau theory around the spontaneous order parameter.

(b) Show that there is a temperature region close to T_c where the size of the fluctuations of the order parameter *exceeds* the mean-field value. In this *critical region* mean field theory is no good.