

A Quantum Approach to Thermodynamics

Von der Fakultät Mathematik und Physik der Universität Stuttgart
zur Erlangung der Würde eines Doktors der
Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

vorgelegt von
Jochen Gemmer
aus Stuttgart

Hauptberichter: Prof. Dr. G. Mahler
Mitberichter: Prof. Dr. U. Seifert

Tag der mündlichen Prüfung: 26. Februar 2003

Institut für Theoretische Physik I
Universität Stuttgart

2003

All numerical simulations in this thesis have been computed very carefully by P. Borowski.

Contents

1	Introduction	1
2	Review of the Field	3
2.1	Boltzmann Equation	4
2.2	Ergodicity Approach	4
2.3	Ensemble Approach	6
2.4	Macroscopic Cell Approach	7
2.5	The Problem of Entropy Invariance	8
2.6	Shannon’s Entropy	9
2.7	Time averaged Density Matrix Approach	10
2.8	Open System Approach	11
3	Structure of a Foundation of Thermodynamics	13
3.1	Checklist of Properties of thermodynamical Quantities	13
3.1.1	Additional necessary Considerations	15
4	Background of the Present Approach	17
4.1	Density Operator and Reduced Density Operator	17
4.2	Compound Systems, Entropy and Entanglement	18
4.3	Fundamental and Subjective Lack of Knowledge	20
4.4	The natural Cell Structure of Hilbertspace	20
5	Analysis of the Cell Structure of Compound Hilbertspaces	23
5.1	Representation of Hilbertspace and Hilbertspace Velocity	23
5.2	Purity and Notation of States	25
5.3	Partition of the Full System into considered System and Surrounding	27
5.4	Microcanonical Conditions	29
5.4.1	Microcanonical Interactions and the corresponding accessible Region	29
5.4.2	The “Landscape” of P^g in the accessible Region	30
5.4.3	The minimum Purity State and the Hilbertspace Average of P^g	30
5.5	Canonical Conditions	34
5.5.1	The accessible and the dominant Region	34
5.5.2	Identification of the dominant Region	34
5.5.3	Analysis of the Size of the dominant Region	35
5.5.4	The canonical equilibrium State	36

5.6	Single Energy Probabilities and Fluctuations	38
5.7	Local Equilibrium States and Ergodicity	40
5.8	Numerical Results	41
5.8.1	Numerical Results for microcanonical Conditions	42
5.8.2	Numerical Results for Canonical Conditions	45
5.8.3	Numerical Results for Probability Fluctuations	49
6	Typical Spectra of Large Systems	51
6.1	The Extensivity of Entropy	51
6.2	Spectra of Modular Systems	52
6.3	The Boltzmann Distribution	56
6.3.1	Beyond the Boltzmann Distribution?	57
7	Temperatures	59
7.1	Definition of spectral Temperature	60
7.2	The Equality of spectral Temperatures in Equilibrium	61
7.3	Spectral Temperature as the Derivative of Energy with Respect to Entropy	63
8	Pressure	67
8.1	On the Concept of adiabatic Processes	67
8.2	The Equality of parametric Pressures in Equilibrium	71
9	Thermodynamical Limit	73
9.1	Weak coupling Limit	73
9.2	Microcanonical Equilibrium	74
9.3	Energy-Exchange Equilibrium	74
9.4	Canonical Equilibrium	75
9.5	Spectral Temperature	75
9.6	Parametric Pressure	76
9.7	Extensivity of Entropy	76
10	Quantum mechanical and classical State Densities	77
10.1	Similarity of classical and quantum mechanical State Densities	78
10.1.1	Sommerfeld Quantization	78
10.1.2	Partition Function Approach	78
10.1.3	Minimum Uncertainty Wavepackage Approach	79
11	Ways to Equilibrium	85
11.1	Theories of Relaxation Behavior	85
11.1.1	Fermi's golden Rule for external Perturbations	85
11.1.2	Fermi's golden Rule for a coupled Environment	86
11.1.3	Weisskopf-Wigner Theory	86
11.1.4	Large Environment Approach	86
11.1.5	Numerical Results for the Relaxation Period	93
12	Summary and Conclusion	95

13 Zusammenfassung	99
13.1 Einleitung und historischer Hintergrund	99
13.2 Axiomatischer Aufbau der Thermodynamik	101
13.3 Grundprinzipien der Theorie	102
13.4 Die Zellstruktur des Hilbertraums	104
13.4.1 Mikrokanonische Bedingungen	105
13.4.2 Kanonische Bedingungen	106
13.5 Spektren modularer Systeme	106
13.6 Spektrale Temperatur	107
13.7 Parametrischer Druck	107
13.8 Klassische und quantenmechanische Zustandsdichte	108
13.9 Wege zum Gleichgewicht	109
13.10 Schlussbetrachtung und Ausblick	109
Appendix	113
A Minimization of the effective Interaction, eq.5.24	113
B Conserved Probabilities, eq.5.28	113
C Averages over Hyperspheres, sect. 5.4.3:	114
D From Multiplication to Gaussian, eq. 6.16	115
E Entropy of an ideal Gas, eq. 6.23	116
F Stabilized adiabatic Approximation, eq. 8.18	117
G Sizes of Hypersphereregions, sect. 5.6	117
H Hilbertspace Averages, sect. 11.1.4	119
Bibliography	123

1 Introduction

In the very beginning thermodynamics have been a purely phenomenological science. Early scientists (Galileo, Santorio, Celsius, Fahrenheit) tried to give definitions for quantities that were intuitively obvious to the observer, like pressure or temperature and measured the interconnections between them. The idea that those phenomena could be directly linked to other fields of physics, like classical mechanics, etc., was not common at these days. This connection was basically introduced when Joule calculated the heat equivalent in 1840, showing that heat was a form of energy, just like kinetic or potential energy in the theory of mechanics of those days.

At the end of the 19th century, when the atomic theory became popular, people began to think of a gas as a box with a lot of bouncing balls inside. With this picture in mind, it was tempting to try to reduce thermodynamics entirely to classical mechanics. This was exactly what Boltzmann tried to do in 1866 when he connected entropy, a quantity which was so far only described phenomenologically, to the volume of a certain region in phase space, an object defined within classical mechanics [1]. This was an enormous step, especially from a practical point of view. Taking this connection for granted one could now calculate all sorts of thermodynamical behavior of a system from its Hamilton function. This gave rise to modern thermodynamics, a theory whose validity is beyond any doubt today. Its results and predictions are a basic ingredient for the development of all sorts of technical apparatuses, ranging from refrigerators to superconductors.

Boltzmann himself, however, tried to prove the conjectured connection between the phenomenological and the theoretical entropy, but did not succeed without making other assumptions like the famous ergodicity or the “a priori postulate”. Other physicists (Gibbs, Birkhoff, Ehrenfest, Von Neumann [2–5]) later on then tried to prove those assumptions, but none of this work seems to have solved the situation entirely. It has even been mentioned that there are more properties of the entropy than its equivalence with the region in phase space to be explained to reduce thermodynamics to classical mechanics, thus the discussion is still ongoing [6]. The vast majority of the work done in this field is based on classical mechanics.

Meanwhile quantum theory, also initially triggered of by the atomic hypothesis has made huge progress during the last century and is today believed to be more fundamental than classical mechanics. At the beginning of the 21st century it seems highly unlikely that a box with balls inside could be anything more than a rough sketch of what a gas really is. Furthermore thermodynamical principles seem to be applicable to systems that cannot even be described in classical phase space. Those developments make it necessary to rethink the work done so far, whether it led to the desired result (e.g., demonstration of

ergodicity) or not.

Of course there have been suggestions how to approach the problem on the basis of quantum theory [7–13], but again, none of them seems to have established the emergence of thermodynamics from quantum mechanics, as an underlying theory, in a decisive way, for there is an ongoing discussion about those ideas also.

So this thesis is meant as a contribution to this whole debate about how thermodynamics can be possibly viewed, as emerging from some underlying theory.

It is organized as follows:

First of all a short and necessarily incomplete overview over the different historical attempts to connect thermodynamics to some underlying theory and their problems is given. Since from this consideration the impression arises that it is even unclear what needs to be shown and proved to establish such a connection in a satisfying manner, in the following the program of this task is defined, i.e., one possible list of conjectures, concerning properties of the entropy and other quantities, is given, that could, proven item by item, serve as a basis for the techniques of thermodynamics. Those proofs, or at least approaches to them, are given in the main part of this thesis. In the last part the connections to other theories are pointed out, e.g., it is explained why the classical theory works so well from a practical point of view, despite its being unsatisfactory in the theoretical context.

2 Review of the Field

Almost all approaches to a connection between thermodynamics and an underlying theory deal with the irreversibility that seems to be present in thermodynamical phenomena, but is most likely absent, in any underlying theory. So the biggest part of these approaches are in some sense proofs of the second law of thermodynamics, which states this irreversibility. They consist of attempts to formulate entropy as a function of quantities, whose dynamics can be calculated in a microscopic picture, in such a way, that the entropy increases during any evolution until it reaches a maximum that is proportional to the logarithm of the volume of the accessible phase space (energy shell). This is the wanted limit because this is the quantity that the entropy has to be identified with, in order to get state functions, that are, for the biggest part, in excellent agreement with experiment. It has not really been appreciated very much that there are other properties of the entropy that remain to be shown even if the above behavior is established (see **3.1**).

One problem of all approaches based on considerations concerning Hamiltonian mechanics is the applicability of classical mechanics itself. To illustrate this, let's consider a gas consisting of atoms or molecules. In principle such a system should, of course, be described by quantum mechanics. Nevertheless, for simplicity, one could possibly treat the system classically, if it started and remained in the Ehrenfest limit, i.e., if the spread of the wavepackages was small compared to the structure of the potentials which the particles encounter. Those are given by the particles themselves, which basically repel each other. If we take the size of those particles to be roughly some 10^{-10}m , we have to demand that the wavepackages should have a width smaller than 10^{-10}m in the beginning. Assuming particle masses between some and some hundred proton masses, and plugging numbers into the corresponding formulas [14], we find that the spread of such wave packages will be on the order of some meters to 100m after one second, which means the system leaves the Ehrenfest limit on a timescale much shorter than the one typical for thermodynamical phenomena. If we demand the packages to be smaller in the beginning, it gets even worse. Considering this, it is questionable whether any explanation based on considerations of Hamiltonian dynamics in Γ -space (Cartesian space, spanned by the $6N$ position and momentum coordinates of a N particle system) or μ -space (Cartesian space, spanned by the 6 position and momentum coordinates of any particle of the system) can ever be a valid foundation of thermodynamics at all. This insufficiency of the classical picture becomes manifest at very low temperatures (freezing out inner degrees of freedom), and it is entirely unclear why it should become valid at higher temperatures even if it produces good results.

Nevertheless a short, and necessarily incomplete overview, also and mainly including such ideas, shall be given here.

2.1 Boltzmann Equation

Boltzmann's work was probably one of the first scientific approaches to irreversibility (1866) [1]. It was basically meant to explain and quantify the observation that a gas which is at first located in one corner of a volume will always spread over the whole volume, whereas a gas uniformly distributed over the full volume never suddenly shrinks to be then concentrated in one corner. This seems to contradict Hamiltonian dynamics according to which any process that is possible forwards in time, should be also possible backwards in time.

Boltzmann considered a function $f(\vec{q}, \vec{v}, t)$ in μ -space. μ -space is the 6-dimensional space of the position and the velocity of one particle. f is a density in this space. If f is high at some point, this means that there are many particles in the corresponding volume-element of position space, flying in the the same direction with the same velocity. Making the assumption of "molecular chaos", i.e., the assumption that the velocities and positions of particles within cells of μ -space that are subject to collisions, are uncorrelated before the collision, he could come up with an evolution equation (" Boltzmann equation") for f , that tends to make f more and more uniform. This is quantified in the famous "H-theorem". Boltzmann showed that a function H defined as

$$H(t) = \int f \log f d^3v d^3q$$

could only decrease in time, given the validity of the Boltzmann equation. This way he somehow established a kind of macroscopic irreversibility but not without additional assumptions. Furthermore H , though being mathematically reminiscent of later formulations of the entropy, could not really be identified with the entropy, for it converges against the logarithm of the accessible volume in μ -space, rather than in Γ -space.

2.2 Ergodicity Approach

The basis of this approach, also pursued by Boltzmann, is the assumption that any possible macroscopic measurement takes a time which is almost infinitely long, compared to the timescale of molecular motion. Thus, the outcome of such a measurement has to be seen as the time average over many hypothetical instantaneous measurements. Thus, if it was true, that a trajectory ventured through all regions of the accessible volume in Γ -space, no matter where it started, the behavior of a system would be, as if it was at any point at the same time, regardless of its starting point. This way irreversibility could be introduced, entropy being somehow connected to the volume, that the trajectory already ventured through, during the observation time.

In order to state this idea in a clearer form, the so called "ergodic hypothesis" has been formulated:

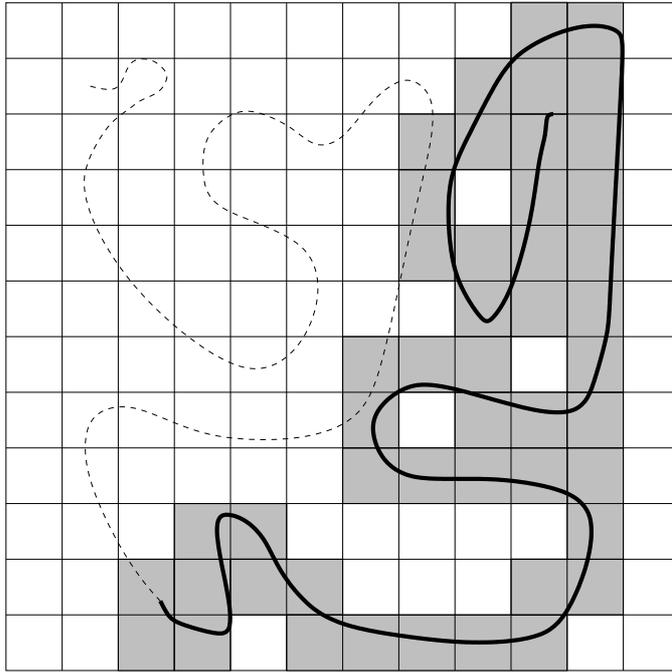


Fig. 2.1: Ergodicity approach. Entropy is defined by the volume in phase space which is occupied by the cells that the trajectory already ventured through (gray region). This obviously depends on the observation time (solid line) and the cell size.

The trajectory of a representative point of the system in Γ -space passes eventually through every point on the energy surface (accessible volume).

This statement taken for true, it can be shown that the amount of time that the trajectory spends in a given volume, is proportional to that volume [15]. This leads to another formulation of the ergodic hypothesis which is that the time average equals the ensemble average, the latter in this case being an average over all system states within the energy shell.

Unfortunately the ergodic hypothesis in this form is necessarily wrong for any system, since all points of the trajectory form a set of measure zero, which therefore cannot fill the energy shell whose measure is not zero [15]. To circumvent this, the quasi-ergodic hypothesis was introduced, which states that the representative point passes arbitrarily close to any given point in the accessible volume in phase-space. Birkhoff and Von Neumann actually demonstrated that there are systems that are quasi ergodic in this sense and that their representing points actually spend equal times in equal Γ -space cells [3,5]. This proof, however, is not generalizable to thermodynamical systems as a whole, and it remains unclear how exactly an entropy should be introduced. It has been suggested to grain phase space into finite size cells (coarse-graining) and simply count the cells that the trajectory passed through in a given time (see **Fig.2.1**), or to count the cells weighted with the time the representing point spent within the cell. Obviously a lot has to be introduced artificially, including the representation, for these considerations could be done the basis of other representing spaces than Γ -space, possibly leading to different results concerning ergodicity or entropy.

2.3 Ensemble Approach

The term “ensemble” was introduced by Gibbs about 1902 [2]. The idea was, that, in general, a macroscopic observation will be consistent with a very large number of microscopic configurations, all of which, represented by their corresponding points in Γ -space, form the ensemble. The ensemble therefore is basically represented by a density in Γ -space which is normalized and nonzero everywhere the system could possibly be found.

To describe the evolution of a system, one now considers the evolution of this density, rather than the evolution of a single representing point. The most important theorem to investigate the evolution of such a density is Liouville's theorem, asserting that the volume of any region in Γ -space is invariant under Hamiltonian evolution. This theorem has two important consequences. Firstly, if the system is described by a density which is uniform throughout all the accessible energy surface, it will be in a stationary state because this distribution cannot change in time. Thus such a state, that somehow fills the entire accessible region can be seen as an equilibrium state. Thus one could be tempted to connect the volume in which such a density is nonzero with the entropy. Unfortunately the second consequence is that such a volume cannot grow in time. This means that if a system does not start off in an equilibrium state, it can never reach one.

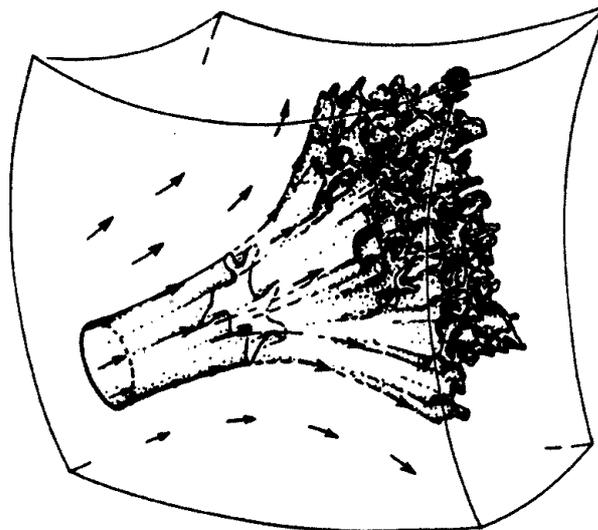


Fig. 2.2: Ensemble approach. The volume of a region in phase space can not grow during a Hamiltonian evolution, due to Liouville's law. Nevertheless it can be transformed to a complicated structure that may eventually be found in any cell, if some graining is introduced

In order to save this concept, Ehrenfest and others introduced coarse-graining also into this idea [4]. They claimed that entropy should not be connected to the volume in which the density is nonzero, but to the number of cells in which the density is nonzero somewhere. If a smooth-shaped region in which an initial density was nonzero would than be mapped by the Hamiltonian evolution onto a sponge like structure featuring the

same volume, but stretched over the whole energy shell (see **Fig.2.2**), such an entropy could grow up to the wanted limit, once the structures of the sponge region become small compared to the cell size. Such a behavior is called “mixing”, owing to a metaphor by Gibbs, who compared the whole scenario to the procedure of dripping a drop of ink into a glass of water, and then stirring it until a mixture results [2].

However, even if Γ -space plays a unique role here (Liouville's theorem is only true in Γ -space or any canonical transformation of it), the cell size has to be introduced artificially, and, of course, mixing has to be shown for any system under consideration. This has been done for some systems but again, there is no generalization to thermodynamical systems as a whole. Another objection raised against this idea in general was, that here entropy seems to be due to the observers inability to find out in which (micro)state exactly the system is. It has been argued that this would introduce an unacceptable amount of subjectivity into this whole field of physics.

2.4 Macroscopic Cell Approach

The idea of the “macroscopic cells” is also due to Ehrenfest who called them “stars”. Such a star is a region in Γ -space that only consists of points that are consistent with one macroscopic description of the system. E.g., if we wanted to describe a gas by the volume it occupies, V , and its total inner energy, E , all points in Γ -space corresponding to the gas occupying a certain volume and containing a certain amount of inner energy, would form the macroscopic cell labeled by those specific macroscopic variables, V and E . This way, Γ -space is not grained into equal sized Cartesian cells like in the former approaches, but into strangely shaped macroscopic cells, that may be extremely different in size from each other.

This difference in size is crucial here, for it is assumed that the “equilibrium cell”, i.e., the cell in which the gas occupies the biggest possible volume is by far the largest one, it is assumed to be large enough to almost fill the entire Γ -space. Technically this needs also to be proven for any thermodynamical system individually also, but it seems much more plausible than the assumption of ergodicity or mixing.

This plausibility is connected with the so called “law of large numbers”. It is usually established by considering some abstract space (basically meant as a model for μ -space), grained into a large number of equal sized cells, and divided into two halves, each one containing an equal number of cells. If now the set of all possible distributions of a large number of points into those cells is examined, it turns out that the vast majority of such distributions features the same amount of points in both halves. The larger the number of cells and the number of points are, the more drastic is this result. Transferring this result to Γ -space, it can be argued that almost all points in Γ -space, corresponding to distributions of points in μ -space, belong to one macroscopic state, specified by one macroscopic variable, that just measures the amount of points, say, in the left half [16]. Having established such a structure of Γ -space, one does not need the strict ergodicity hypothesis any longer, for if the trajectory wanders around in Γ -space without any further restrictions, it will most likely eventually spend almost all time in the biggest cell, even if it started in a small one (see **Fig. 2.3**).

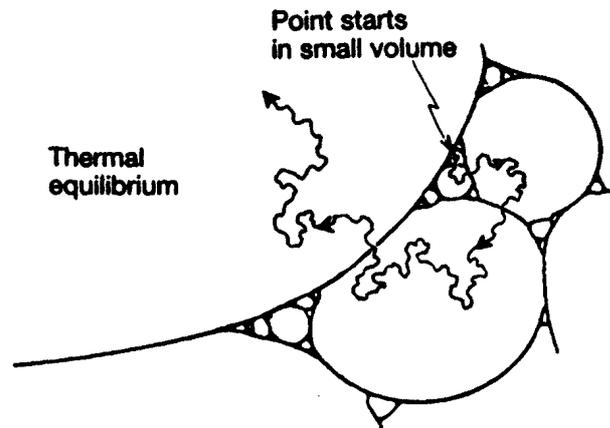


Fig. 2.3: Macroscopic cell approach. If phase space is divided into cells, one of which is extremely much larger than all the others, any trajectory that is not subject to any further restrictions, is likely to end up in the biggest cell.

In such an approach entropy is connected to the size of the cell the representing point is wandering through. But here new problems are to be faced: To decide whether or not a given micro state belongs to a macroscopic cell one has to assign a volume to the configuration of gas particles. This is more subtle than it might seem at first sight. One method is to coarse grain configuration space into standard Cartesian cells and count the cells that are occupied by at least one particle. But this yields only satisfactory results for a certain ratio of the cell size to the diluteness of the gas particles. Other approaches proceed by taking the convex cover or other measures defined for a set of points, thus there is a fair amount of arbitrariness.

Another problem in this idea arises, if one tries to examine situations in which the inner energy of a system is not rigidly fixed, but the system can exchange energy with another, in some sense bigger, system called a heat bath. In this case one finds empirically that the probability to find the system in a state of energy E is proportional to $\exp -\frac{E}{kT}$ (Boltzmann-distribution). If one enlarges the Γ -space now to contain both, the system and the bath, one can derive this behavior but then, again, only under the assumption of ergodicity, the very condition one tried to get rid of.

2.5 The Problem of Entropy Invariance

One issue that has attracted much less attention in the past than the second law, is the fact that entropy, even if it can be shown to increase up to desired equilibrium value, should be shown to be invariant under adiabatic reversible processes [17]. Otherwise the identification of negative pressure with the derivative of energy with respect to volume under constant entropy, a basic statement of the first law, would be without justification. The only well-known answer to this question on the basis of classical mechanics is the

“Law of the invariance of the phase space volume” [17]. This law states roughly the following: If one performs an infinitesimal step of an adiabatic process, all representative points that belonged to one accessible region in Γ -space before this step, will now, after the step, in general, belong to different, new accessible regions. The average of the volumes of all those new accessible regions equals the volume of the accessible region before the process step. In other words: The ensemble average of the volume of the accessible region is invariant under an infinitesimal adiabatic step.

The problem of this concept is the following: If one believes that the system truly is in one micro state, belonging to the ensemble corresponding to one macro state, even if this micro state cannot be known, then the system could end up at two different entropies (and energies), that is in two different macro states, after an adiabatic process, depending on the micro state that the system has initially been in. If one believes the existence of micro states, one could perform two identical processes on two identical systems in two identical macro states and they would end up in different macro states. Of course one can come up with arguments about how likely it is that they really do end up at different macro states, but this requires an individual analysis of any considered system individually, just like it is the case with ergodicity.

This problem is a problem that all approaches described so far are confronted with, for they all start from the idea of an precisely defined existing micro state underlying each macro state.

2.6 Shannon's Entropy

This consideration is neither based on classical mechanics nor on quantum theory, indeed its applicability is not even meant to be restricted to physics at all. The idea is that entropy is a measure of lack of information, thus the whole concept is introduced as a rational way of dealing with incomplete knowledge, wherever it may occur [18]. Thus, this approach somehow radicalizes the idea that has already been underlying some of the formerly described concepts, namely that thermodynamical behavior is due to the observers inability to measure precisely and in enough detail.

The technique proceeds as follows: Consider the quantity S'

$$S' = -c \sum_i p_i \ln p_i$$

where distinguishable states of the system, whatever they might be, are labeled by “ i ” and “ p_i ” is the probability to find the system in state i . Now to find the best guess about the probability distribution, if one knows some property of the system for sure, one has to compute the maximum of S' with respect to the ‘ p_i ’s under the restriction that the resulting description of the system has to feature the known property. This maximum of S' then is the entropy. (In the field of quantum mechanics, this scheme is mostly referred to as Jaynes' principle.) For thermodynamical systems one has to set $c = k$ and if, e.g., volume and energy of a system are known one has to maximize over all states featuring this energy and volume, which is all states from the accessible region. In this case a uniform distribution over all those states results, which is equivalent to

the “a priori” postulate by Boltzmann which states that one, if no other information is to be obtained, should assume equal probabilities for all possible states. In contrary to the other concepts, within this idea, the Boltzmann distribution of probabilities can be directly obtained, by claiming that if a system is in contact with a heat bath (canonical conditions), only the average energy of the distribution, rather than the energy of every individual state, should be taken as a known property. In the same way the situations of some other extensive quantities being allowed for exchange can be handled with success, i.e., yielding results which are in agreement with experiment.

Nevertheless, if this principle is applied to the situation of some other, less familiar, extensive quantities being allowed for exchange, the resulting states are not even stationary anymore, which is inconsistent for the resulting states should be equilibrium states.

The main problem of this concept is its basic idea: The fact that everything is based on a subjective lack of knowledge of the observer, that it is entirely unclear to what extent the observer could possibly get rid of this insufficiency, and in which way this would effect the behavior of physical thermodynamical quantities. Since within this approach it is not even possible to formulate entropy as a function of time it can hardly be seen as a justification of the second law.

2.7 Time averaged Density Matrix Approach

This concept is explicitly quantum mechanical. An entropy S was introduced into quantum mechanics by Von Neumann [19], who formulated it in terms of the density matrix $\hat{\rho}$ as

$$S = -k\text{Tr}\{\hat{\rho} \ln \hat{\rho}\} \quad (2.1)$$

Since any possible state of a quantum system $\hat{\rho}$ can be written as a weighted mixture of states that form an orthogonal set, within the basis formed by this set, Von Neumann’s entropy reduces to Shannon’s entropy with those orthogonal states taken as the distinguishable states, and their weights as the corresponding probabilities. The Von Neumann entropy is invariant under unitary transformations. This property has two consequences: It is independent of the chosen basis, and it is time independent, just like the Gibbs entropy in the ensemble approach.

Since one needs an entropy that can possibly change in time, it has been suggested to calculate S using a time averaged density matrix rather the actual instantaneous one.

The elements of the time dependent density matrix read

$$\langle n|\hat{\rho}(t)|m\rangle = e^{\frac{(E_n - E_m)t}{i\hbar}} \langle n|\hat{\rho}(0)|m\rangle \quad (2.2)$$

where $|n\rangle, |m\rangle$ are energy eigenstates and E_n, E_m the respective eigenenergies. Since all off diagonal elements are oscillating, they will vanish if one time averages the density matrix [20]. And it can be shown that the Von Neumann entropy indeed rises, if the off diagonal elements vanish.

The problem of this idea is the averaging time. If systems get big, typically the energy level spacing becomes very small, and the averaging time that is necessary to actually see

entropy rise significantly, may very well exceed typical relaxation times of thermodynamical systems. In the limit of degenerate energy eigenstates this averaging time becomes infinitely long. Thus, technically speaking, a system with a precisely given energy (micro-canonical conditions) and therefore only occupying degenerate energy eigenstates, would never exhibit an increasing entropy, regardless of the averaging time.

2.8 Open System Approach

This concept is based on the idea that any system is subject to some kind of interaction with some sort of environment, i.e., that isolation is an idealization [21]. Thus the correct description is given by a density matrix describing the system and its environment, rather than a density matrix, describing the only the system. A description of the system alone is then obtained in computing the so called reduced density matrix, which is basically the full density matrix traced over the environment (see (4.5)). Under some additional assumptions concerning the state of the environment, correlations between the system and the environment, and the timescale on which correlations within the bath decay, it is possible to derive an approximate autonomous evolution equation for the reduced density matrix, in which the influence of the environment appears more like a boundary condition for the considered system, the environment not having any degrees of freedom of its one any longer. This equation is called master equation. Practically this amounts to the addition of some anti-Hermitian terms to the Hamiltonian of the system, a modification that gives rise to a time dependence of the entropy, a behavior which could be never generated by a merely Hermitian Hamiltonian [23].

This method has been extremely successful in the field of quantum optics and gives usually rise to an evolution that reaches an equilibrium state of some maximum entropy. Nevertheless the master equation cannot be derived without some additional assumptions, furthermore the calculation of the exact form of the master equation gets extremely complicated if systems get bigger, such that is has to remain open whether or not an application of this technique to all sorts of thermodynamical phenomena can be performed or not.

3 Structure of a Foundation of Thermodynamics

For a foundation of thermodynamics it is tempting to give an ontological, intuitively appealing definition of entropy, like entropy is the number of micro realizations for some macro state, entropy is some volume in phase space, entropy is a lack of knowledge, etc. But then one is left with the task to show that the defined quantity indeed behaves in a way, that is empirically observable (see **2**).

Thus, often thermodynamics are simply directly introduced on the basis of some axiomatic structure, Telling how thermodynamical quantities are expected to behave. Usually some of the axioms deal with the notorious irreversibility of thermodynamics, others with the possibility of formulating the inner energy as a state function, in terms of the entropy (dS) and other extensive variables, rather than in terms of the amount of heat exchanged (dQ). Starting from first principles is it evident that a lot of the behavior of thermodynamical quantities (including their mere existence) needs explanation.

In the following a “checklist” of properties of thermodynamical quantities is given, meant as a set of rules by which thermodynamical behavior (as we observe it) is completely defined. Thus, if all those rules can be shown to result from an underlying theory, thermodynamics can be considered as emerging from this underlying theory, if the approach fails to demonstrate any of them, the task is not fully accomplished.

The set of rules given here is neither considered unique nor irreducible. If, e.g., ergodicity was assumed, the validity of the second law (**II**) and the equality of intensive variables (**V**) would follow naturally, thus one could replace those two items by demanding ergodicity; this, however, would not be suitable for the the approach at hand. So the choice of properties given here is only meant to be complete and approachable from the theory of quantum mechanics.

3.1 Checklist of Properties of thermodynamical Quantities

I Definition of thermodynamical Quantities:

All thermodynamical quantities (entropy, temperature, energy, pressure, etc.) should be precisely defined as functions of the variables of an underlying theory, such that this underlying theory describes the dynamics of those variables. If thermodynamical quantities are introduced this way, inevitably, they are always defined, even though they might not

always show the equilibrium properties listed below (**II-IV**) simply because equilibrium might not be reached yet, or, even more simple, the system does not belong to the class of thermodynamical systems (see **3.1.1**)

II Second law of thermodynamics:

This axiom establishes the irreversibility of some thermodynamical processes and states the existence of a stationary equilibrium state, into which thermodynamical systems will evolve eventually:

It should be shown that the system reaches a state for which the fluctuations of all well defined thermodynamical quantities are negligible. This state has to be controllable by macroscopic constraints. Since those constraints can be set in different ways, i.e., by keeping different sets of intensive and extensive variables fixed or controlled, at least two cases have to be distinguished:

a) Microcanonical conditions (energy kept fixed):

In this case the entropy should only increase during the evolution, and the final state should only depend on the energy of the initial state.

b) Canonical conditions (temperature kept fixed):

Since under these conditions the equilibrium state of the system is controlled by the contact with a heat bath, whose only specifying parameter is a temperature, its equilibrium state should only depend on this temperature, regardless of its initial state.

III First law of thermodynamics:

This Law is the one from which eventually connections between measurable, macroscopic intensive and extensive quantities are inferred. Thus it guarantees that for a certain class of processes that involve a change of those macroscopic variables, a detailed microscopic picture is dispensable and can be replaced by a simpler, macroscopic picture.

a) State function:

It should be shown that if the extensive variables, say, volume and entropy, take on certain values, energy necessarily takes on a corresponding value, regardless of the process by which the state was reached.

b) Temperature as a conjugate variable:

It should be shown that there are processes (heating, cooling) in which all extensive variables are kept fixed except for energy and entropy which then should be shown to depend on each other, connected as

$$\frac{\partial U}{\partial S}_{V=const.} = T \quad (3.1)$$

where of course the same definition of temperature as above has to be used.

c) Other conjugate variables:

It should be shown that there are processes (adiabatic, reversible) in which an extensive variable (e.g., the volume) changes, while all others, including especially entropy, remain constant. The analysis of such a process then has to yield

$$\frac{\partial U}{\partial V}_{S=const.} = -p \quad (3.2)$$

IV Extensivity:

It should be shown that thermodynamical variables that are claimed to be extensive, in particular the entropy S , are indeed extensive quantities

V Equality of intensive variables:

It should be shown that two systems which are allowed to exchange an extensive quantity, will end up having the same conjugate intensive variable.

Those properties of thermodynamical quantities and the various relations allow for an application of the standard techniques and methods of thermodynamics. Thus, if they are shown to result as claimed on the basis of quantum mechanics, the laws of thermodynamics are established, and the whole field is reduced to quantum mechanics. Such an axiomatic construction, theoretically satisfying as it might be, will eventually have to be judged by the results it produces, thus, in order to make it a physically meaningful theory, rather than just an abstract mathematical consideration, the limits of its applicability have to be examined just as much its connection to the standard classical theory.

3.1.1 Additional necessary Considerations**Thermodynamical limit:**

This point states the necessity to explain and clarify the relation between a possibly emerging theory and its underlying theory. If the emerging properties inevitably result from the underlying theory, one could discard the latter completely. In the case at hand this cannot be the case, for the underlying theory is supposed to be quantum mechanics, and it is obvious that not all systems that obey quantum mechanics can be described thermodynamically, since Schrödinger-type quantum theory is believed to underlie all sorts of non-relativistic systems. Thus, a fairly precise definition of a class of systems which can be expected to behave thermodynamically, should be given. This definition should not result in a tautology like “All systems that show the properties mentioned above are thermodynamical systems”, but in a criterion that can be checked with comparatively less effort.

Quantum versus classical entropy:

Despite its problematic foundation standard “classical” thermodynamics works pretty well for almost all practical purposes. If this is not just incidental, it should be possible to show, that entropy as a function of, say, volume and energy should be the same, no matter whether it is calculated based on a standard classical definition or the quantum mechanical one that can be shown to have the above properties. Here, this will eventually amount to showing that quantum state density and the volume of energy shells in classical Γ -space are proportional, for large classes of systems.

4 Background of the Present Approach

As already mentioned we want to demonstrate the properties of thermodynamical quantities on the basis of non-relativistic quantum mechanics, i.e., starting initially with a wavefunction to describe the system, whose evolution is generated by the Schrödinger equation. However, we are not going to solve any Schrödinger equation explicitly, since those dynamics only supply the background for our considerations just like Hamiltonian dynamics only supply the background for the considerations described in **2**, which essentially consist of the analysis of structures in Γ -space. Very similar our approach will essentially be based on the analysis of structures in Hilbertspace.

4.1 Density Operator and Reduced Density Operator

The most general way to note the information about the state of a quantum mechanical object is a density or state operator. It is defined as

$$\hat{\rho} = \sum_i W_i |\psi_i\rangle\langle\psi_i| \quad (4.1)$$

where the $|\psi_i\rangle$ are different wavefunctions, each of which has to be given the statistical weight W_i , to describe the state of the system. From this object the probability to find the system in a state $|\phi\rangle$ can be calculated as

$$W_\phi = \langle\phi|\hat{\rho}|\phi\rangle \quad (4.2)$$

Since any possible set of such probabilities is “codeable” in such a density operator, it represents the most general description. The evolution equation for the density operator (Von Neumann equation) can be directly calculated from the Schrödinger equation and is found to be:

$$\frac{d}{dt}\hat{\rho} = i\hbar [\hat{\rho}, \hat{H}] \quad (4.3)$$

If one chooses to consider the full Hilbertspace of a system to be composed as a product of two Hilbertspaces

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \quad (4.4)$$

and if one is only interested in observables or probabilities that are defined only within, say e.g., \mathcal{H}_1 , an object containing all this “local” information that is available, given the

density operator ρ of the full system which is defined on the entire Hilbertspace \mathcal{H}_1 , can be calculated as

$$\hat{\rho}_1 = \text{Tr}_2 \{ \hat{\rho} \} \quad (4.5)$$

and is called the reduced or local density operator. This local density operator is defined only on \mathcal{H}_1 and thus, naturally, does not contain any local information concerning \mathcal{H}_2 or information concerning correlations between the two parts of the system.

4.2 Compound Systems, Entropy and Entanglement

In all theoretical approaches to thermodynamics entropy plays a central role. The entropy we are going to consider is the Von Neumann entropy [19], which is defined for a system on the basis of its density operator.

$$S = -k \text{Tr} \{ \hat{\rho} \ln \hat{\rho} \} \quad (4.6)$$

In a diagonal representation of $\hat{\rho}$ this reads

$$S(\{W_n\}) := -k \sum_n W_n \ln W_n, \quad (4.7)$$

where the W_n 's are the eigenvalues of $\hat{\rho}$, which are routinely interpreted as the probabilities to find the system in the eigenstate $|n\rangle$. This entropy measures the ‘‘broadness’’ of the probability distribution onto the different eigenstates of $\hat{\rho}$. If the system is for sure in some pure state, which is a state that can be simply described by a wavefunction, this state will be an eigenstate of the density operator and its corresponding probability will be 1, while the probabilities for all other states vanish. Thus, for any such pure state, entropy will be zero. If within some finite subspace (or the full space of a finite system) of dimension N , any pure state is ‘‘equally likely’’, all probabilities will be $1/N$, which is the broadest possible distribution, and entropy will take on its maximum value $S = k \ln N$. Thus entropy can also in this context be viewed as a measure for a lack on knowledge.

As already explained **2.7** this entropy is invariant with respect to unitary transformations. Since the Von Neumann equation (4.3) gives rise to a unitary evolution for the density operator $\hat{\rho}$, this entropy can never change and is thus no good candidate for the thermodynamical entropy with all the required properties.

If one deals with compound systems, i.e., systems that can naturally be decomposed into two parts, for which a lot of the interesting observables can be defined separately (energy in subsystem 1, energy in subsystem 2, positions of the particles of subsystem 1, positions of the particles of subsystem 2, etc.), it is reasonable to organize the full Hilbertspace of the compound system as a product Hilbertspace as explained in (4.4). In this case a valid and irreducible description of, say, subsystem 1 is given by the reduced density operator $\hat{\rho}_1$, rather than by the density operator of the full system $\hat{\rho}$, whose entropy can never change. The entropy of subsystem 1 defined according to (4.6) on the basis of $\hat{\rho}_1$ (4.5) can very well change under unitary transformations of the compound system. It

does not change if the unitary transformation generated by the Von Neumann equation factorizes as $\hat{U}(t) = \hat{U}_1(t) \otimes \hat{U}_2(t)$ which, in turn, will be the case if the Hamiltonian of the full system does not contain any interactions between the subsystems. But if the Hamiltonian contains such interactions, regardless of how small they might be, the subsystem entropy is not a conserved quantity anymore and will, in general, change in time [24].

For our considerations it will be assumed that the state of the compound system is a pure state, i.e., a state of zero entropy. If for such a state the entropy of a subsystem vanishes also, the state must be a product state, i.e., a state that can be written as the direct product of two separate wavefunctions: $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$. If the subsystem entropy does not vanish the state cannot have this product form, which means it is entangled. Such pure entangled states, states that cannot be written in a product form, have been in the center of interest now for almost seventy years. They can have properties that seem to contradict intuition. If a local measurement on one subsystem is made, i.e., a projection of only one subsystem state is performed, the local state of the other subsystem can be severely affected, which raised the question whether quantum mechanics could be valid at all [25]. Regardless of their being puzzling these states can theoretically be shown to result from product states, if the subsystems are allowed to interact for a while, and on a small scale their existence has been demonstrated experimentally. However, it is a widespread belief that entanglement as a fundamentally quantum mechanical phenomenon could only exist between very small objects.

In our theory the Von Neumann entropy of the reduced density operator which describes the system under consideration is taken for the basic entropy of thermodynamics. Thus if this entropy rises this can only be due to increasing entanglement with another subsystem. This other subsystem is believed to be the environment of the thermodynamical system which is considered indispensable. So one of the basic ideas is, that there is no thermodynamical system without an environment and that both subsystems, the considered system and its environment have to be treated fully quantum mechanically, regardless of their size. If one, e.g., analyses a gas, the gas would not be a thermodynamical system if it was not confined to some volume by a container, which thus represents the environment subsystem. Hence, if a gas relaxes towards equilibrium, it does so, due to increasing entanglement with the container (or other coupled subsystems like the electromagnetic field, etc.).

It has often been argued that the influence of the environment should not play any crucial role, since entropy rises also (or especially) for the case of an isolated system. This, however, means in the context of thermodynamics only that the system is not allowed to exchange any extensive quantities like energy or particles etc. with the environment. It does not mean that there is no interaction between the system and its environment. This however is not to be confounded with a microscopically closed system. Gas particles within a thermally insulated container nevertheless interact with the particles that make up the container walls, otherwise they would not even stay inside the container. Quantum mechanically such an interaction, even if it does not allow for an exchange of energy with the environment, will, nevertheless, typically give rise to entanglement [24]. Thus even for a thermally closed system the existence of an environment and the interaction

with it is crucial.

4.3 Fundamental and Subjective Lack of Knowledge

It has often been argued that entropy should somehow be a measure for a lack of knowledge. But then the question arises whether the observer, by overcoming his deficiency to calculate or observe more precisely, i.e., by reducing his subjective lack of knowledge, could truly influence the entropy and the resulting thermodynamical behavior of real physical systems.

Within classical mechanics lack of knowledge may always be considered subjective: In principle, any observable can be known with unlimited precision. This is different in quantum mechanics. From the uncertainty principle we know that there are always observables that are undetermined. Nevertheless, in single system scenarios (no compound systems), at least one observable can, in principle, be known exactly at any time, if the initial state was a pure state, hence the fundamental lack of knowledge does not grow. But for compound systems there are states where all observables referring to a specific subsystem are unknown, even if some compound observable of the full system is exactly predictable, just like the position of a particle is necessarily unknown to anybody, if its momentum is exactly predictable, thus, the fundamental lack of knowledge is considerably larger here, than in the former case. Those states are exactly the entangled states mentioned above **4.2** [26, 27]. Thus, compound systems might evolve from states that contain exact knowledge about some observable of each subsystem (pure product states) into the above mentioned states, featuring this fundamental lack of knowledge about any observable [24].

So, in the quantum domain we have two possible sources of ignorance: one being due to our inability to identify the initial state and calculate the evolution exactly, the other being intrinsic to the momentary state and thus present, even for an “infinitely smart demon”.

Here we want to show that in typical thermodynamical situations the fundamental lack of knowledge by far dominates over the subjective lack of knowledge in the following sense: Almost all the possible evolutions (of which we are typically unable to predict the actual one), will eventually lead to states that are characterized by a maximum fundamental lack of knowledge about the considered subsystem; this lack is only limited by macroscopic constraints.

4.4 The natural Cell Structure of Hilbertspace

Within the context of classical mechanics, it has often been pointed out, that problems arise from the fact, that it is at least very difficult, if not impossible, to formulate entropy as an observable, i.e., as a function of a momentary micro state of a system. Within quantum mechanics it has also become impossible to formulate the Von Neuman entropy as an

observable in the sense of an entropy operator for which

$$S = \langle \psi | \hat{S} | \psi \rangle \quad \text{or} \quad S = \text{Tr} \{ \hat{S} \hat{\rho} \}. \quad (4.8)$$

But even if such a formulation is not feasible it is very well possible to formulate the the above mentioned entropy of the considered subsystem as a function of the momentary “micro state” $|\psi\rangle$ of the full system:

$$S = -k \text{Tr} \{ \text{Tr}_c \{ |\psi\rangle\langle\psi| \} \ln \text{Tr}_c \{ |\psi\rangle\langle\psi| \} \}. \quad (4.9)$$

(Here the index “ c ” denotes the environment subsystem.) This entropy is obviously exactly defined once a system-environment partition is established, and the state of the full compound system is known. No averaging time or cell size needs to be introduced artificially.

On the basis of this definition it is now possible to decompose the Hilbertspace of the full system into different cells, according to their local entropy. All states featuring, according to the above definition 4.9, the same entropy are grouped together to form one cell. This cell structure is thus uniquely and unambiguously specified.

Just like the representing point of a classical system wanders around in Γ -space, the state vector of a quantum mechanical system wanders around in Hilbertspace. And just like the representing point in Γ -space is confined to some energy shell, the state vector in Hilbertspace is also confined to some accessible region. This region is analogously set by the overall energy conservation, the space a system is allowed to occupy, but depends also on whether the system is allowed to exchange energy with the environment (canonical conditions), or not (microcanonical conditions), or other constraints that can be controlled macroscopically.

The crucial point to establish thermodynamical behavior will be to show, that those accessible regions lie almost entirely within the cell specified by the maximum entropy, that is consistent with the macroscopic constraints. Or stated the other way round: that almost all states within an accessible region feature the maximum entropy consistent with the macroscopic constraints.

The main idea here is that equilibrium is not reached because the motion of the state vector in Hilbertspace will cease at some point (in contrary it will be shown to move always with constant velocity), but is reached because the state vector will eventually almost always venture inside the cell with the maximum entropy, simply because this cell fills almost the entire accessible region.

From a topological point of view the picture here is pretty much the same as in the macroscopic cell approach **2.4**, for which Penrose argues [16]:

“We would seem now to have an explanation for the second law! For we may suppose that our phase space point does not move about in any particularly contrived way, and if it starts off in a tiny phase space volume, corresponding to a small entropy, then, as time progresses, it will indeed be overwhelmingly likely to move into successively larger and larger phase space volumes, corresponding to gradually increasing entropy values.” The assumption here is, that, if the maximum entropy compartment is overwhelmingly bigger than any other compartment within the accessible region, the trajectory, if it

started in a compartment of lower entropy, will eventually leave it, to enter the compartment of the maximum entropy. This is not to be confounded with the ergodic hypothesis or the quasi ergodic hypothesis! It is not at all assumed that the trajectory passes arbitrarily close to any point of the accessible region, or that it spends equal times in equal volumes. It is only assumed, that it does not stay within a extremely small volume, if it is not confined to it. The system is not treated as if it was at any point at the same time or with the same probability, it is always treated as being at some absolutely concrete point in Hilbertspace. Which point this exactly may be depends on the the precise knowledge of the initial state, the details of the Hamiltonian of the system, and could only be calculated by solving the Schrödinger equation of some 10^{23} interacting particles. Thus probably there will be a huge subjective lack of knowledge about where exactly the system might be found in Hilbertspace. But this is not what gives rise to the second law, and this subjective lack of knowledge is not to be confounded with the thermodynamical entropy. The entropy would not be any smaller if one knew the exact location in Hilbertspace. The fact that this precise knowledge is dispensable is exactly what creates the universal development towards equilibrium: Two systems under identical macroscopic constraints starting from two different initial states will wander on different trajectories through Hilbertspace forever, but if both trajectories eventually wander through the same compartment of maximum entropy, the systems will end up in the same equilibrium state.

Other than in the macroscopic cell approach, where the states that belong to the maximum entropy compartment are only macroscopically the same, here all states that belong to the maximum entropy compartment are locally identical. States belonging to a compartment of lower entropy may have to be described with different local reduced local density operators, but the density operator with the maximum entropy consistent with the macroscopic constraints is unique. Thus, although the trajectory of the full system might move quickly through the biggest compartment, the system is locally in a stationary state, its density operator does not vary in time. For the local considered system the maximum entropy state is an attractor.

5 Analysis of the Cell Structure of Compound Hilbertspaces

In this chapter we want to prove that it is indeed true that for appropriate compound Hilbertspaces almost all states that a system can possibly evolve into, under given macroscopic constraints, are states of the highest possible entropy, that is consistent with those constraints.

5.1 Representation of Hilbertspace and Hilbertspace Velocity

Other than real configuration space, Hilbertspace, the space on which quantum mechanical state vectors are defined, is neither three dimensional nor real, which makes it almost inaccessible to intuition. Thus there is no obvious way how to parametrize, i.e., specify by a set of numbers, quantum mechanical states. Probably one has to choose a basis $\{|i\rangle\}$ such that

$$|\psi\rangle = \sum_i \psi_i |i\rangle \quad (5.1)$$

It is, however, undetermined which basis one should choose and how one should represent the set of complex numbers $\{\psi_i\}$. This could be done in terms of real and imaginary parts, absolute values and phases or in many other ways. Eventually one will always have a set of real numbers that somehow specify the state. To decide now, how big a region in Hilbertspace really is, the only way is to calculate the size of the region that the corresponding specifying parameters occupy. Therefore one eventually has to organize the parameter space as a real cartesian space of some dimension. The problem now is that the size of this region will depend on the parametrization chosen, Thus, if one wants to compare such regions in size, one has to explain why one does this on the basis of the special chosen parametrization.

This question does not only arise in the context of quantum mechanics or Hilbertspaces, it needs to be answered for classical Γ -space considerations also. It is far from obvious that classical systems have to be parametrized in terms of their positions and momenta. In the Gibbs approach this parametrization is chosen to guarantee the validity of Liouville's law, other theories just assume it, because it is eventually the volume of the energy shell for a parametrization in terms of positions and momenta that has to be

identified with the entropy in order to get correct results. Especially in the macroscopic cell approach this parametrization remains without justification.

If one wants to use the relative sizes of compartments to guess in which compartment of a region, that is entirely accessible to a representing point, this point will most likely be found, the crucial thing is the effective velocity of this point. If one has, e.g., two compartments, one being extremely much bigger than the other, the guess that the representing point will be found in the bigger one might be bad, if the dynamics are such that the point moves with extremely high velocity in the big compartment and very slow in the small one. In the case at hand the effective velocity of the representing point, on his trajectory through the representing cartesian parameter space, depends on parametrization.

Most convenient would be a parametrization of the states such that the velocity of the representing point in the cartesian space of the specifying parameters (in the following simply called Hilbertspace) is constant throughout one accessible region. Fortunately this is feasible.

Consider a representation of a state in terms of the real and imaginary parts of its amplitudes in some basis $\{|i\rangle\}$:

$$|\psi\rangle = \sum_i (\psi_i + i\psi'_i)|i\rangle \quad (5.2)$$

If the ψ_i and ψ'_i are organized in an cartesian parameter space, a Hilbertspace with a real, regular, cartesian metric is defined. All vectors that represent physical states, i.e., that are normalized, lie on a hypersphere of radius one:

$$\langle\psi|\psi\rangle = \sum_i (\psi_i - i\psi'_i)(\psi_i + i\psi'_i) = \sum_i \psi_i^2 + \psi_i'^2 = 1 \quad (5.3)$$

this property is obviously independent of the choice of the basis $\{|i\rangle\}$

In this parametrization of Hilbertspace the effective Hilbertspace velocity, v is given by

$$v = \sqrt{\sum_i \left(\frac{d}{dt}\psi_i\right)^2 + \left(\frac{d}{dt}\psi'_i\right)^2}. \quad (5.4)$$

The square of this velocity may be written as

$$v^2 = \sum_i \frac{d}{dt}(\psi_i - i\psi'_i) \frac{d}{dt}(\psi_i + i\psi'_i) = \left\langle \frac{d}{dt}\psi \middle| \frac{d}{dt}\psi \right\rangle, \quad (5.5)$$

which means it is also independent of the chosen basis.

The velocity v can now be calculated from the Schrödinger equation in the following form:

$$i\hbar \frac{d}{dt}|\psi\rangle = (\hat{H} - c)|\psi\rangle \quad (5.6)$$

where c is a arbitrary constant that has no physical meaning since it results just in an overall phasefactor of $e^{\frac{ic}{\hbar}}$ for any solution. It has, however, an influence on the

Hilbertspace velocity, for it could, e.g., make a stationary state be represented by a moving point in Hilbertspace. One thus wants to choose c such as to make the Hilbertspace velocity as small as possible, since any motion that is due to c just reflects a changing overall phase which has no physical significance.

Inserting (5.6) into (5.5) we find for the square of the Hilbertspace velocity v^2 :

$$v^2 = \langle \psi | (\hat{H} - c) (\hat{H} - c) | \psi \rangle = \langle \psi | \hat{H}^2 | \psi \rangle - 2\langle \psi | \hat{H} | \psi \rangle c + c^2 \langle \psi | \psi \rangle \quad (5.7)$$

Obviously v is constant, since all terms in (5.7) that could possibly depend on time are expectation values of powers of \hat{H} and thus constants of motion.

Searching the minimum of v^2 with respect to c now yields

$$c = \langle \psi | \hat{H} | \psi \rangle = \langle \psi(0) | \hat{H} | \psi(0) \rangle \quad (5.8)$$

which, inserted into (5.7) makes the Hilbertspace velocity

$$v = \sqrt{\langle \psi | \hat{H}^2 | \psi \rangle - \langle \psi | \hat{H} | \psi \rangle^2} = \sqrt{\langle \psi(0) | \hat{H}^2 | \psi(0) \rangle - \langle \psi(0) | \hat{H} | \psi(0) \rangle^2}, \quad (5.9)$$

which is just the energy uncertainty, or the variance of the energy probability distribution of the corresponding state. Accordingly, stationary states, energy eigenstates, are now represented by non-moving points in Hilbertspace. Since all states that belong to one accessible region have the same energy probability distribution, and thus the same variance of this distribution, all states that venture, due to Schrödingerian dynamics, through one accessible region do so, with the same constant velocity.

5.2 Purity and Notation of States

As already explained one measure for the lack of knowledge of a (mixed) state is the entropy. Another (inverse) measure for the broadness of the probability distribution is the purity P [28]:

$$P(\{W_n\}) := \sum_n W_n^2 \quad \text{or} \quad P(\hat{\rho}) := \text{Tr} \{ \hat{\rho}^2 \}. \quad (5.10)$$

which is also invariant with respect to unitary transformations. In general, these two measures do not uniquely map onto each other, but for the limiting cases they do. Pure states have minimum entropy 0 and maximum purity 1. States featuring the maximum entropy of $S = k \ln N$ within some Hilbertspace of dimension, N have the minimum purity $1/N$, within the corresponding Hilbertspace.

Furthermore the state with the lowest possible purity, that is consistent with a given probability distribution of the system onto a set of orthogonal states (a given set of diagonal elements of the density matrix), is unique. This can be seen from splitting up $\hat{\rho}$ into two parts. One, \hat{D} that is a diagonal matrix, consisting of the mentioned set of probabilities, and the rest, $\hat{\epsilon}$, a matrix that does not have any diagonal elements, such that different states that are consistent with the given set of probabilities can only differ

in $\hat{\epsilon}$.

now with

$$\hat{\rho} = \hat{D} + \hat{\epsilon} \quad (5.11)$$

where $\hat{\epsilon}$ of course has to be hermitian, we find:

$$P(\hat{\rho}) = \text{Tr} \left\{ (\hat{D} + \hat{\epsilon})^2 \right\} = \text{Tr} \left\{ \hat{D}^2 \right\} + \text{Tr} \left\{ \hat{\epsilon}^2 \right\} + 2\text{Tr} \left\{ \hat{D}\hat{\epsilon} \right\}. \quad (5.12)$$

Due to the properties of \hat{D} and $\hat{\epsilon}$ the last term vanishes. Since

$$\text{Tr} \left\{ \hat{\epsilon}^2 \right\} \geq 0 \quad (5.13)$$

the smallest purity is reached for the equality in the above equation, which is only possible for

$$\hat{\epsilon} = 0. \quad (5.14)$$

Thus, the smallest possible purity state is unique and consists only of

$$\hat{\rho} = \hat{D}, \quad (5.15)$$

i.e, it does not have any diagonal elements. Thus, if one can show that for a certain region of Hilbert space, the purity of a subsystem takes on a minimum, one has established that any full state within this region yields the same local state featuring the same entropy, energy, etc.

Since, as will be seen below, purity can be expressed as a rather simple function of the full state, whose evaluation does not require the diagonalization of a matrix at some point as the calculation of local entropy would, we will mainly consider P , and exploit the above mentioned property, if P is found to be minimal.

In order to really calculate P , we now need to introduce a notation of states which is adequate for our purposes. Since we are going to deal with bi-partite systems, partitioned into the considered system, labeled by g ("gas") and the surrounding labeled by c ("container"), the pure state of the full system will be denoted as

$$|\psi\rangle = \sum_{A,B} \sum_{a,b} \psi_{ab}^{AB} |A, a\rangle \otimes |B, b\rangle, \quad (5.16)$$

where $|A, a\rangle(|B, b\rangle)$ denotes the a -th(b -th) energy eigenstate with the energy eigenvalue $E_A(E_B)$ of the gas (container) system, with $a = 1, 2, \dots, N_A^g, b = 1, 2, \dots, N_B^c$. $N_A^g(N_B^c)$ are the respective degrees of degeneracy.

From any such wavefunction the reduced density operator of the subsystem, g , is found by tracing over the container system:

$$\hat{\rho}^g := \sum_{A,A',B} \sum_{a,a',b} \psi_{ab}^{AB} (\psi_{a'b}^{A'B})^* |A, a\rangle \langle A', a'| \quad (5.17)$$

For this state the purity is:

$$P^g = \sum_{ABCD} \sum_{abcd} \psi_{ab}^{AB} (\psi_{cb}^{CB})^* \psi_{cd}^{CD} (\psi_{ad}^{AD})^* \quad (5.18)$$

Here and in the following a, c, A, C , label the gas, b, d, B, D , the container subsystem. (Note that $P^g = P^c$ as the total state is taken to be pure)

5.3 Partition of the Full System into considered System and Surrounding

As will be seen, for our approach it is crucial that any thermodynamical system has to allow for a partition of its full Hamiltonian in the following way:

$$\hat{H} =: \hat{H}^g + \hat{H}^c + \hat{I} \quad \text{with} \quad \sqrt{\langle \hat{I}^2 \rangle} \ll \langle \hat{H}^g \rangle, \langle \hat{H}^c \rangle \quad (5.19)$$

where \hat{H}^g and \hat{H}^c are the local Hamiltonians of the system and the environment, respectively, which means they have to act on two different parts of a product Hilbertspace, thus they have to fulfill $[\hat{H}^g, \hat{H}^c] = 0$. \hat{I} is some sort of interaction and the inequality for the expectation values has to hold for all states that the system can possibly evolve into, under given (macroscopic) constraints. If a partition according to this weak coupling scheme was impossible, the idea of system proper and surrounding would be meaningless. Furthermore, weak coupling is required to justify the concepts of extensive/intensive variables, a basic feature of thermodynamics [29] (cf. also [30])

Making such a partition possible, though, might involve a re-organization of the Hamiltonian.

Consider, just as an example, an ideal gas in a container. \hat{H}^g in this case is only the free particle Hamiltonian,

$$\hat{H}^g = \sum_i \frac{\hbar^2}{2m} (\hat{p}_i^g)^2 \quad (5.20)$$

This sum over the kinetic energies of all gas particles is the only part of the Hamiltonian that acts on the gas subspace alone. \hat{H}^c , the Hamiltonian of the container provides the environment that has at least to be present to make the gas particles a thermodynamical system. It reads

$$\hat{H}^c = \sum_i \frac{\hbar^2}{2m_i} (\hat{p}_i^c)^2 + \frac{1}{2} \sum_{ij} \hat{V}^c(\vec{x}_i^c, \vec{x}_j^c), \quad (5.21)$$

where $\hat{V}^c(\vec{x}_i^c, \vec{x}_j^c)$ are the interactions that bind the container particles to each other to form a solid, and acts exclusively in the container subspace. Thus, as required, \hat{H}^g and

\hat{H}^c commute. Now \hat{I} contains the interactions of all gas particles with all container particles and reads:

$$\hat{I} = \sum_{ij} \hat{V}^{gc}(\vec{x}_i^g, \vec{x}_j^c). \quad (5.22)$$

This part contains the repelling interactions of the gas particles with the container particles and establishes the container as a boundary for the gas particles they cannot escape of. Starting from first principles, the Hamiltonian has to be written in this way, especially the last part is indispensable. Unfortunately any stationary state of \hat{H}^g , and it is such a state we want to see the system evolve into, is unbound and thus not confined to any volume that might be given by the container. This means the expectation value of \hat{I} , for an energy eigenstate of the uncoupled problem, $\hat{H}^g + \hat{H}^c$ would definitely not be small, and thus the system would miss a fundamental prerequisite for a thermodynamical system, assessible from our method.

Fortunately this deficiency can be overcome by defining an effective local Hamiltonian and an effective interaction:

$$\hat{H}^{g'} := \hat{H}^g + \hat{V}^g \quad \hat{I}' := \hat{I} - \hat{V}^g \quad \hat{V}^g := \hat{V}^g(\{\vec{x}_i^g\}) \quad (5.23)$$

Here, \hat{V}^g is some potential for the gas particles alone. \hat{V}^g will be chosen to minimize $\sqrt{\langle \hat{I}'^2 \rangle}$. Substituting the real by the effective parts of the Hamiltonian obviously leaves the full Hamiltonian unchanged, but now there is a chance that the partition will fit into the above scheme. A good candidate for \hat{V}^g will be some sort of effective “box” potential, modeling the mean effect of all container particles together. This makes \hat{I}' , the deviation of the true “particle-by-particle” wall interaction from the “effective-box” wall interaction, that can indeed, expected to be small. The eigenstates of the gas system alone are then simply the wellknown bound eigenstates of particles in a box of corresponding size.

The exact mathematical minimalization of \hat{I}' for a given wavefunction of the full system $\psi = \psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})$, using the Euler-Lagrange method yields (see **A**):

$$\hat{V}^g(\{\vec{x}_i^g\}) = \frac{\int \hat{I}'(\{\vec{x}_i^g\}, \{\vec{x}_j^c\}) |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \prod_j d\vec{x}_j^c}{\int |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \prod_j d\vec{x}_j^c} \quad (5.24)$$

From this formula it can be seen that the ideal choice of \hat{V}^g is indeed some sort of box potential where the slope of the wall potential depends on the range of the repelling forces \hat{V}^{gc} , regardless of the details of the wavefunction, as long as the container particles stay tightly bound together. It is thus plausible that for a realistic model, (container particles tightly bound, short ranged repelling forces) the rearranged Hamiltonian will indeed fit the above scheme.

In general, however, even the effective interaction \hat{I}' cannot be made zero, and represents a coupling, i.e., a term that cannot be written as a sum of terms that act on the different subspaces separately. Such a coupling, however small it might be can, and in general will, produce entanglement, thus causing local entropy to increase. (For a detailed example of this partition scheme see [31]).

5.4 Microcanonical Conditions

5.4.1 Microcanonical Interactions and the corresponding accessible Region

It has often been claimed that a system under so-called microcanonical conditions would not interact with its environment. This, however, is typically not true (cf. [32, 33]). A thermally isolated gas in a container, e.g., definitely interacts with the container, otherwise the gas could not even have a well defined volume, as explained in the previous section. If a system is thermally isolated, it is not necessarily isolated in the microscopic sense, i.e, not interacting with any other system. The only constraint is, that the interaction with the environment should not give rise to any energy exchange. If the energies contained in the gas and the environment

$$E^g := \langle \hat{H}^g \rangle \quad E^c := \langle \hat{H}^c \rangle \quad (5.25)$$

are to be conserved, it follows from

$$\left[\hat{H}^g, \hat{H} \right] = 0 \quad \left[\hat{H}^c, \hat{H} \right] = 0 \quad (5.26)$$

that

$$\left[\hat{H}^{g'}, \hat{I}' \right] = 0 \quad \left[\hat{H}^c, \hat{I}' \right] = 0 \quad (5.27)$$

Except for these constraints we need not specify \hat{I}' in more detail. All interactions that fulfill this relation will create perfectly microcanonical situations, regardless of their strength or any other feature. And, as will be shown, there are a lot of possible interactions that, do fulfill this conditions and create entanglement.

Based on these commutator relations we find (see **B**), that for any energy eigenspace specified by A, B

$$\sum_{a,b} |\psi_{ab}^{AB}(t)|^2 = \sum_{a,b} |\psi_{ab}^{AB}(0)|^2 = W_{AB} \quad (5.28)$$

is a conserved quantity, set by the initial state. W_{AB} is the probability to find the gas system at the energy E_A^g and the container system at the energy E_B^c . Thus any state that features this same energy probability as the initial state, belongs to the accessible region and could possibly be reached during microcanonical dynamics.

For the case of initial product states, states that have zero local entropy in the beginning, that we mainly want to consider here, we get

$$\sum_{a,b} |\psi_{ab}^{AB}(0)|^2 = \sum_{a,b} |\psi_a^A(0)|^2 |\psi_b^B(0)|^2 = W_A^g W_B^c \quad (5.29)$$

where W_A^g (W_B^c) are now the separate probabilities of finding the gas-(container-) system somewhere in the possibly highly degenerate subspace, characterized by the energy eigenvalues E_A^g (E_B^c). This is the only constraint that microcanonical conditions impose on the accessible region of Hilbertspace.

5.4.2 The “Landscape” of P^g in the accessible Region

To demonstrate that the accessible region really has the cell structure mentioned in 2.4, namely that the biggest part of it is filled with states of almost minimum purity, we proceed as follows:

First we compute the (unique) state with the lowest possible purity, $\hat{\rho}_{min}^g$ (with purity $P = P_{min}^g$), that is at all consistent with a given initial state and microcanonical conditions, i.e., with a given energy probability distribution $\{W_A^g\}$. Then we compute the average of P^g over the total accessible Hilbertspace region, using the kind of parametrization described in 5.1. We will show that this average purity is very close to the purity of the lowest possible purity state $\hat{\rho}_{min}^g$ for a large class of systems. Considering only these systems, that then define the class of thermodynamical systems, we can conclude that $P^g \approx P_{min}^g$, for almost all states within this region. This argument holds since any distribution with a boundary at some side, that has a mean value close to that boundary, has to be sharply peaked. The closer the mean value lays to this boundary, the sharper the peak has to be. Since all states from the accessible region feature the same energy probability distribution $\{W_A^g\}$ and the minimum purity state consistent with this distribution, $\hat{\rho}_{min}^g$, is unique, all states within the accessible region that feature $P^g \approx P_{min}^g$, must yield reduced local states, that are very close to $\hat{\rho}_{min}^g$. Thus, as long as the trajectory keeps wandering through the compartment filled with those states, the gas system is locally in a stationary state.

5.4.3 The minimum Purity State and the Hilbertspace Average of P^g

The minimum purity state consistent with the microcanonical conditions (5.28, 8.11) and its corresponding purity are:

$$\hat{\rho}_{min}^g = \sum_{A,a} \frac{W_A^g}{N_A^g} |A, a\rangle \langle A, a| \quad P_{min}^g = \sum_A \frac{(W_A^g)^2}{N_A^g} \quad (5.30)$$

which means that this state is implemented, if all states that belong to the same energy eigenspace are equally likely ($\langle A, a | \hat{\rho}_{min}^g | A, a \rangle$ is independent of a). To check that this is indeed the state with the smallest purity consistent with a given energy probability distribution $\{W_A^g\}$, we introduce a deviation $\hat{\delta}$ such that the resulting state is still consistent with $\{W_A^g\}$ and compute its purity. For this deviation we can restrict ourselves to diagonal contributions since all off diagonal contributions will only make purity bigger, anyway (see (5.15)). For

$$\hat{\delta} = \sum_{A,a} \delta_{A,a} |A, a\rangle \langle A, a| \quad \text{with} \quad \sum_a \delta_{A,a} = 0 \quad (5.31)$$

we find

$$P(\hat{\rho}_{min}^g + \hat{\delta}) = P_{min}^g + \sum_{A,a} \delta_{A,a}^2 \quad (5.32)$$

which means that $\hat{\rho}_{min}^g$ is indeed the minimum purity state since the last term of (5.32) will always be positive and thus only make purity larger.

To calculate the Hilbertspace average of P^g , denoted as $\langle P^g \rangle$, we need a parametrization for $|\psi\rangle = \{\psi_{ab}^{AB}, \psi_{ab}^{AB'}\}$ confined to the allowed region (5.28, 8.11). The Hilbertspace average can then be written as

$$\langle P^g \rangle = \frac{\int P^g(\{\phi_n\}) \det \mathcal{F} \prod_n d\phi_n}{\int \det \mathcal{F} \prod_n d\phi_n}, \quad (5.33)$$

where ϕ_n is the respective set of parameters and \mathcal{F} is the corresponding functional matrix.

According to (5.28) and (8.11) the real and imaginary parts of the amplitudes that correspond to a degeneracy subspace AB must be parametrized to lie on a hypersphere of radius $R = \sqrt{W_A^g W_B^c}$. Thus there is a corresponding set of parameters $\{\phi_n^{AB}\}$ for each degeneracy subspace, on which the amplitudes of this and only this subspace depend. This means that the functional matrix \mathcal{F} , has block form and its determinant factorizes such that

$$\langle P^g \rangle = \frac{\int P^g \prod_{AB} \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}}{\prod_{AB} \int \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}} \quad (5.34)$$

As a consequence, the average over each term of P^g according to (5.18) factorizes and reduces to a product of averages over the degeneracy subspaces, that the amplitudes in that very term correspond to.

The averages of the terms of P^g can be grouped to form ‘‘classes’’: It turns out that the averages of the terms of P^g belonging to the same class have the same functional dependence on the degree of degeneracy $N_{AB} := N_A^g N_B^c$ of the subspaces they correspond to and on the probability of those subspaces to be occupied $W_{AB} := W_A^g W_B^c$. Thus, we first have to evaluate the average of each term within a given class, add all terms and finally sum over all classes.

There are six classes to be considered (For the details of their evaluation, see **C**):

I: All four amplitudes correspond to different states ($(A \neq C) \vee (a \neq c) \wedge ((B \neq D) \vee (b \neq d))$). Then each average is a product of averages of products of different cartesian coordinates over full hyperspheres. As can be seen from the appendix (C.8) those averages vanish.

II: The ‘‘gas indices’’ (indices referring to the gas subsystem) of all amplitudes corresponds to the same gas state, but the container indices to two different container subspaces ($(A = C) \wedge (a = c) \wedge (B \neq D)$). They factorize into:

$$T^{II} = \left(\frac{\int |\psi_{ab}^{AB}|^2 \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}}{\int \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}} \right) \left(\frac{\int |\psi_{ad}^{AD}|^2 \det \mathcal{F}^{AD} \prod_n d\phi_n^{AD}}{\int \det \mathcal{F}^{AD} \prod_n d\phi_n^{AD}} \right) \quad (5.35)$$

The application of (C.8) yields

$$T^{II} = \frac{W_{AB} W_{AD}}{N_{AB} N_{AD}} \quad (5.36)$$

There are $N_A^g N_B N_D^c$ terms belonging to that class and subspace-combination. Since subspaces that have $B = D$ are excluded we find, summing over subspaces:

$$\sum_{ABD} N_A^g (N_B^c N_D^c - \delta_{BD} (N_B^c)^2) T^{II} = \sum_A \frac{(W_A^g)^2}{N_A^g} \left(1 - \sum_B (W_B^c)^2 \right) \quad (5.37)$$

III: The container indices of all amplitudes correspond to the same container state, but the gas indices to two different gas subspaces ($A \neq C$) \wedge ($(B = D) \wedge (b = d)$). By repeating the above procedure (II) we get:

$$\sum T^{III} = \sum_A \frac{(W_B^c)^2}{N_B} \left(1 - \sum_A (W_A^g)^2 \right) \quad (5.38)$$

IV: The gas indices of all amplitudes correspond to the same gas state and the container indices to the same container subspace but to different container states within this subspace ($(A = C) \wedge (a = c) \wedge ((B = D) \wedge (b \neq d))$). Those terms are of the form

$$T^{IV} = \frac{\int |\psi_{ab}^{AB}|^2 |\psi_{ad}^{AB}|^2 \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}}{\int \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}} \quad (5.39)$$

To evaluate this equation one needs (C.9) and gets:

$$T^{IV} = \frac{(W_A^g)^2 (W_B^c)^2 \Gamma(N_{AB})}{\Gamma(N_{AB} + 2)} = \frac{(W_A^g)^2 (W_B^c)^2}{N_{AB} (N_{AB} + 1)} \quad (5.40)$$

For each subspace-combination there are $N_A^g N_B^c (N_B^c - 1)$ terms in that class. Thus, summing over subspaces we find:

$$\sum_{AB} N_{AB} N_B^c (N_B^c - 1) T^{IV} = \sum_{AB} \frac{(W_A^g)^2 (W_B^c)^2 (N_B^c - 1)}{N_{AB} + 1} \quad (5.41)$$

V: The container indices of all amplitudes correspond to the same container state and the gas indices to the same gas subspace but to different gas states within this subspace ($(A = C) \wedge (a \neq c) \wedge ((B = D) \wedge (b = d))$). For those terms the above calculation (IV) has to be repeated yielding:

$$\sum T^V = \sum_{AB} \frac{(W_A^g)^2 (W_B^c)^2 (N_A^g - 1)}{N_{AB} + 1} \quad (5.42)$$

VI: All four amplitudes correspond to the same state. These terms read:

$$T^{VI} = \frac{\int |\psi_{ab}^{AB}|^4 \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}}{\int \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB}} \quad (5.43)$$

Using (C.9) one gets:

$$T^{VI} = \frac{2(W_A^g)^2 (W_B^c)^2 \Gamma(N_{AB})}{\Gamma(N_{AB} + 2)} = \frac{2(W_A^g)^2 (W_B^c)^2}{N_{AB} (N_{AB} + 1)} \quad (5.44)$$

Since there are $N_A N_B$ terms in this class for each subspace, one gets:

$$\sum_{AB} T^{IV} N_{AB} = \sum_{AB} \frac{2(W_A^g)^2 (W_B^c)^2}{N_{AB} + 1} \quad (5.45)$$

If we now finally sum up all the contributions from the six classes we get:

$$\begin{aligned} \langle P^g \rangle = & \quad (5.46) \\ & \sum_A \frac{(W_A^g)^2}{N_A^g} \left(1 - \sum_B (W_B^c)^2 \right) + \sum_B \frac{(W_B^c)^2}{N_B^c} \left(1 - \sum_A (W_A^g)^2 \right) \\ & + \sum_{A,B} \frac{(W_A^g)^2 (W_B^c)^2 (N_A^g + N_B^c)}{N_A^g N_B^c + 1} \end{aligned}$$

The Hilbert space average is thus a unique function of the invariants W_A^g, W_B^c , specified by the initial product state, and the degeneracies N_A^g, N_B^c .

If we, just to get in contact with previous results, ask for the average purity of an arbitrary state with no constraints at all we can enlarge the accessible region over all Hilbertspace by formally taking both subsystems as completely degenerate. Doing so, we recover

$$\langle P^g \rangle = \frac{N^g + N^c}{N^g N^c + 1} \quad (5.47)$$

as a special case [34].

If the degeneracy of the occupied energy levels is large enough that $N_A^g N_B^c + 1 \approx N_A^g N_B^c$, which should hold true for typical thermodynamical systems, (5.46) reduces to

$$\langle P^g \rangle \approx \sum_A \frac{(W_A^g)^2}{N_A^g} + \sum_B \frac{(W_B^c)^2}{N_B^c} \quad (5.48)$$

The first sum in this expression is obviously exactly P_{min}^g (5.30), so that for systems and initial conditions, in which the second sum is very small, the allowed region almost entirely consists of states for which $P^g \approx P_{min}^g$. The second sum will be small if the container system occupies highly degenerate states, typical for thermodynamical systems, in which the surrounding is much larger than the considered system. This is the set of cases mentioned already in 5.4.2. Since the density operator that has $P^g = P_{min}^g$ and $S^g = S_{max}^g$ and is consistent with the microcanonical conditions is unique, the density operators with $P^g \approx P_{min}^g$ should not deviate much from this one and should therefore also have $S^g \approx S_{max}^g$, the latter being

$$S_{max}^g = -k \sum_A W_A^g \ln \frac{W_A^g}{N_A^g} \quad (5.49)$$

which reduces for sharp energy probability distribution $W_A^g = \delta U A$ to

$$S_{max}^g = k \ln N_U^g \quad (5.50)$$

For some more concrete illustration of the implications that the rather abstractly derived principles in this chapter bare on the dynamics of adequate systems, see 5.8.1.

5.5 Canonical Conditions

5.5.1 The accessible and the dominant Region

Our approach to the canonical conditions will be based on similar techniques as before. The possibility of a partition according to (5.19) is still assumed. But now there is no further constraint on the the interaction \hat{I} , since energy is allowed to flow from one subsystem to the other. The only constraint for the accessible region therefore derives from the initial state of the full system, and the fact that the probability to find the total system at some energy E

$$W_E := \sum_{A,B/E} W_{AB} = \sum_{A,B/E} \sum_{a,b} |\psi_{ab}^{AB}|^2 \quad (5.51)$$

should be conserved. (Here $A, B/E$ stands for: all A, B such that $E_A + E_B = E$) One could try to repeat the above calculation under this constraint, but now it turns out that the actual minimum purity is no longer close to the average purity over the accessible region. Thus, one has to proceed in a slightly different way.

Contrary to the microcanonical case, the probability to find the gas (container) subsystem at some given energy is no longer a constant of motion here. But we are going to prove that there is a predominant distribution, $\{W_{AB}^d\}$, which almost all states within the allowed region have in common. The subregion formed by these states will be called the “dominant region”. Once the existence of such a dominant region has been established, we can use the results from the microcanonical conditions to argue that almost all states within this dominant region feature the maximum local entropy that is consistent with the predominant distribution.

Therefore, just like in the previous case, our subjective lack of knowledge about where to find the system within the accessible region should be irrelevant. The reduced local state $\hat{\rho}^g(t)$ as a function of the full state $|\psi(t)\rangle$ should always evolve into a state with a fixed probability distribution W_A^g , and an almost time invariant entropy, which is the maximum entropy that is consistent with this (canonical) distribution. Nevertheless, the state of the full system continues to move with the constant velocity (5.7) in Hilbertspace.

5.5.2 Identification of the dominant Region

First, we calculate the size of a region in Hilbertspace that is associated with a certain energy probability distribution $\{W_{AB}\}$. This size will then be maximized with respect to the W_{AB} 's, under the condition of the energy probability distribution of the whole system $\{W_E\}$ being kept fixed, in order to find the predominant distribution $\{W_{AB}^d\}$. According to (5.34) the size of the region associated with the energy distribution $\{W_{AB}\}$ is:

$$A(\{W_{AB}\}) := \prod_{AB} \int \det \mathcal{F}^{AB} \prod_n d\phi_n^{AB} \quad (5.52)$$

Those integrals are just the surfaces of hyperspheres and can be done using the techniques described in the appendix:

$$A(\{W_{AB}\}) := \prod_{AB} W_{AB}^{N_{AB}-1/2} O(N_{AB}) \quad (5.53)$$

Here $O(N_{AB})$ is the surface area of a $2N_{AB}$ -dimensional hypersphere of radius $R = 1$. Instead of maximizing A directly, we choose to maximize $\ln A$; this being equivalent, since the logarithm is a monotonous function. Furthermore we drop all terms, that do not depend on $\{W_{AB}\}$, since they are of no relevance for the maximum and set $N_{AB} - 1/2 \approx N_{AB}$, an approximation that is not necessary but simplifies the calculation and is definitely valid for large degrees of degeneracy. Introducing the Lagrange multipliers $\{\lambda_E\}$, the function we want to maximize with respect to the W_{AB} reads:

$$\ln \tilde{A} = \sum_{A,B} N_{AB} \ln W_{AB} - \sum_E \lambda_E \left(\sum_{A,B/E} W_{AB} - W_E \right) \quad (5.54)$$

This maximization is routinely done by solving the following set of equations:

$$\frac{\partial \ln \tilde{A}}{\partial W_{AB}} = 0 \quad (5.55)$$

and yields

$$W_{AB}^d = \frac{N_{AB}}{\lambda_{E=E_A+E_B}} \quad (5.56)$$

Finally, using (5.51) we find for the Lagrange multipliers

$$\lambda_E = \frac{N_E}{W_E} \quad N_E = \sum_{A,B/E} N_{AB} \quad (5.57)$$

We have thus identified the energy probability distribution, which most of the states within the accessible region exhibit, i.e., the energy probability distribution of the dominant region, $\{W_{AB}^d\}$.

5.5.3 Analysis of the Size of the dominant Region

So far we have only shown that among the regions with given energy probability distribution $\{W_{AB}\}$, there is a biggest (or smallest) one. But for our argument we need to show that this region \tilde{A}^d is, indeed, extremely larger than all the others, that it really fills almost the entire accessible region. To examine the size of this region we need to know, how the size of a region depends on the corresponding distribution $\{W_{AB}\}$, if this distribution does not deviate much from the dominant distribution $\{W_{AB}^d\}$. Therefore we consider $W_{AB} =: W_{AB}^d + \epsilon_{AB}$, where the ϵ_{AB} 's are supposed to be small. For $\ln \tilde{A}$ we then find

$$\ln \tilde{A} = \sum_E \sum_{A,B/E} N_{AB} \ln(W_{AB}^d + \epsilon_{AB}) \quad (5.58)$$

with

$$\sum_{A,B/E} \epsilon_{AB} = 0 \quad (5.59)$$

The latter condition guarantees that $W_{AB} = W_{AB}^d + \epsilon_{AB}$ still belongs to the accessible region.

Expanding the logarithm to second order we get:

$$\ln \tilde{A} \approx \sum_E \sum_{A,B/E} N_{AB} \left(\ln(W_{AB}^d) + \frac{\epsilon_{AB}}{W_{AB}^d} - \frac{1}{2} \left(\frac{\epsilon_{AB}}{W_{AB}^d} \right)^2 \right) \quad (5.60)$$

Since the expansion is around an extremum the linear term should vanish. Indeed, using (5.59) the second summation over this term yields:

$$\sum_{A,B/E} N_{AB} \frac{\epsilon_{AB}}{W_{AB}^d} = \sum_{A,B/E} \lambda_E \epsilon_{AB} = 0 \quad (5.61)$$

so that, using (5.56) and (5.57), we finally find:

$$\tilde{A} \approx \tilde{A}^d \prod_{AB} \exp - \frac{\left(\sum_{A,B/E} N_{AB} \right)^2}{2N_{AB}W_E} \epsilon_{AB}^2 \quad (5.62)$$

i.e., regions, \tilde{A} , that correspond to energy probability distributions that deviate from the dominant one are smaller than the dominant region, \tilde{A}^d . Since the smallest factor that can appear in the exponent of (5.62) for given N_E is $\frac{N_E}{2W_E}$, the regions \tilde{A} will be extremely smaller already for very small deviations, if the corresponding N_E 's are large. So this is another prerequisite for a system to be thermodynamical.

5.5.4 The canonical equilibrium State

Finally, to find the marginal, dominant energy probability distribution W_A^d of the gas system individually, one has to sum the compound probabilities W_{AB}^d over the irrelevant container system to obtain, using (5.57):

$$W_A^d = \sum_B W_{AB}^d = \sum_B \frac{N_{AB}}{\lambda_E} = N_A^g \sum_B \frac{N_B^c W_{E/A,B}}{N_{E/A,B}} \quad (5.63)$$

or, stated more clearly by writing $W_A^d =: W^d(E_A^g)$, $W_E =: W(E)$, $N_A^g =: N^g(E_A^g)$, $N_B^c =: N^c(E_B^c)$, $N_E =: N(E)$ and switching from a summation over B to a summation over E

$$W^d(E_A^g) = N^g(E_A^g) \sum_E \frac{N^c(E - E_A^g) W(E)}{N(E)} \quad (5.64)$$

This is the energy probability distribution that one will, for a thermodynamical system, overwhelmingly likely find for the gas system. Simply by exchanging the indices (up to here everything is symmetric with respect to an exchange of the subsystems) we find the marginal dominant energy probability distribution for the container system, which will be of interest later (see 7.1).

$$W^d(E_B^c) = N^c(E_B^c) \sum_E \frac{N^g(E - E_B^c)W(E)}{N(E)} \quad (5.65)$$

As can be seen from (5.64) the dominant energy probability distribution W_A^d is not independent of the initial state since different energy probability distributions of the local initial state may result in different overall energy probability distributions $W(E)$, and those clearly enter (5.64). This contradicts the idea that under canonical conditions a system should reach some equilibrium state, that is entirely independent of its initial state. Such a behavior, however, results, if a special form of $N^c(E_B^c)$ is assumed. If

$$N^c(E_B^c) = N_0^c e^{\alpha E_B^c}, \quad (5.66)$$

where N_0^c, α are some constants, (5.64) reduces to

$$W^d(E_A^g) = N^g(E_A^g) e^{-\alpha E_A^g} \sum_E \frac{N_0^c e^{\alpha E} W(E)}{N(E)}. \quad (5.67)$$

Due to overall probability conservation, the sum over E has to be independent of $W(E)$ and thus, in this case, W_A^d is independent of the initial state. Whether or not (5.66) may be a reasonable assumption, is discussed in **6.2** and **6.3**.

So far we have only established the energy probability distributions that almost all states from the accessible region feature, but nothing has been said about entropy, purity, etc., thus the equilibrium state is still undetermined. But since all states within the dominant region feature the same energy probability distribution, motion within the dominant region will never give rise to any energy exchange between the subsystems. The situation now is as if the system was controlled by microcanonical conditions, and an initial state from the dominant region, as long as we assume that the trajectory will practically never leave the dominant region.

Thus we can take the arguments from **5.4** to identify the equilibrium state. Following this, we can conclude that if

$$\sum_B \frac{W_c^d(E_B^c)^2}{N^c(E_B^c)} \ll \sum_A \frac{W_g^d(E_A^g)^2}{N^g(E_A^g)}, \quad (5.68)$$

almost all states from the dominant region will yield approximately the same local gas state. This equilibrium state $\hat{\rho}_{eq}^g$ is again the state of minimum purity (maximum entropy) that is consistent with the dominant energy distribution:

$$\hat{\rho}_{eq}^g \approx \sum_{A,a} \frac{W_g^d(E_A^g)}{N^g(E_A^g)} |A, a\rangle \langle A, a| \quad (5.69)$$

For some more concrete illustration of the implications, that the rather abstractly derived principles in this chapter bare on the dynamics of adequate systems, see **5.8.2**.

5.6 Single Energy Probabilities and Fluctuations

Unfortunately the term “fluctuations” in itself is not really precisely defined in the field of physics. In the context of thermodynamics, one speaks of thermal fluctuations, meaning that some extensive variable, e.g. energy, might not be exactly sharp, if the intensive variable, e.g. temperature, is kept fixed. Rather than a sharp energy, one gets an energy probability distribution peaked at some value, but having a certain width. This width is referred to as fluctuations, but the distribution itself is constant in time, i.e. does not fluctuate.

In the context of quantum mechanics, fluctuations also refer to the width of (quantum mechanical) probability distributions. The so called “vacuum fluctuations” mention the fact that the probability to measure some finite electromagnetic field, different from zero, does not vanish in vacuum. But again the probability distribution itself, is constant in time.

The fluctuations we want to discuss in this section are of another kind. Here all probabilities are functions of time. Since the “probability landscape” is not entirely flat, probabilities will vary in time, while the states vector wanders around in Hilbertspace. It are those fluctuations in time we are going to analyze here, by analyzing the flatness of the probability landscapes.

So far we have only shown that there is a dominant region that will be the bigger, the higher the state density of the full system is. But that does, technically speaking, not necessarily mean that $W^d(E_A^g)$, being the probability to find the gas system in energy level A , is the probability that most of the states within the accessible region feature. This is due to the fact the the dominant region corresponds to the full set of dominant energy probabilities, the whole distribution, rather than to a single energy probability. Fortunately it is, however, possible to calculate the sizes of regions that are associated with different, individual $W(E_A^g)$ precisely, this being equivalent to the probability to find a state featuring a certain $W(E_A^g)$, if states from the accessible region are picked at random. We will call a function describing the sizes of those regions as a function of the concrete probabilities $f_A^g(W)$. Wherever $f_A^g(W)$ is peaked will then truly be the most likely value for $W(E_A^g)$, and the sharper this peak is, the smaller will fluctuations typically be.

To examine this, we restrict ourselves for the moment to cases for which the energy distribution of the full system is sharp, i.e.,

$$W(E) = \delta_{E,U} \quad \text{with} \quad E = E_A^g + E_B^c. \quad (5.70)$$

Given this initial condition, all states that belong to the accessible region can be written as

$$|\psi_{acc}\rangle = \sum_A \sum_{a,b} \psi_{ab}^{AB(A)} |A, a\rangle \otimes |B(A), b\rangle, \quad (5.71)$$

where $B(A)$ has to be chosen such that (5.70) is fulfilled. Since the states have to be normalized, the accessible region is now characterized by:

$$\sum_A \sum_{a,b} (\psi_{ab}^{AB(A)})^2 + (\psi_{ab}^{\prime,AB(A)})^2 = 1, \quad (5.72)$$

where $\psi, (\psi'')$ now denote the real(imaginary) parts of the corresponding amplitudes. If the parametrization described in 5.1 is chosen, which is again assumed here, the accessible region obviously consists of a hypersphere of radius 1, since the real and imaginary parts of the amplitudes are just cartesian coordinates. The dimension, d of this hypersphere is

$$d = 2 \sum_A N_A N_{B(A)}. \quad (5.73)$$

From the topological point of view all those dimensions are equivalent. It is only the underlying partition scheme that makes them distinguishable. Thus if some partition is established it may only be some subset of those amplitudes that determines quantities like, e.g., the probability of the gas system to be found at some energy E_C^g . This quantity in terms of the amplitudes reads:

$$W(E_C^g) = \sum_{c,b} (\psi_{cb}^{CB(C)})^2 + (\psi_{cb}''^{CB(C)})^2. \quad (5.74)$$

So the number of coordinates that enter this quantity, d_C is

$$d_C = 2N_C N_{B(C)} \quad (5.75)$$

Since it is just a question of notation, we may, for simplicity, think of those coordinates as being the first ones. $W(E_C^g)$ now is obviously the sum of the squares of those first d_C coordinates. If we now want to know how big the compartment of the accessible region associated with some concrete value of $W(E_C^g)$ is, we have to ask how big the zone on a surface of a hypersphere of radius 1 and dimension d is, that only consists of points for which the sum of the squares of the first d_C coordinates equals this concrete value W . The sizes of those zones can be calculated by parametrizing the surface in such a way that eventually $W(E_C^g)$ is one of the parameters. If now the surface integral is computed, and all other parameters are integrated out except for $W(E_C^g)$, the remaining integrand will describe the size of the zones as a function of $W(E_C^g)$. This can then be converted to the relative frequency of states featuring $W(E_C^g) = W$, among all states from the accessible region by adequate normalization, yielding (see 5.6 (G.8))

$$f_C^g(W) = \frac{1}{K} \sqrt{W}^{d_C-2} \sqrt{1-W}^{d-d_C-2}, \quad (5.76)$$

K just being some normalization constant. The place of the maximum, W_{max} and the mean value, $\langle W \rangle$ read (G.10,G.11):

$$W_{max} = \frac{d_C - 2}{d - 4}, \quad \langle W \rangle = \frac{d_C}{d} (= W^d(E_C^g)) \quad (5.77)$$

Thus, for large d, d_C i.e., for large full systems, the most frequent W as well as the mean value turn out to be identical with $W^d(E_C^g)$, thus making the energy probabilities of the dominant distribution indeed the most likely ones, also for each single probability.

Now, of course, the broadness of this relative frequency distribution, $f_C^g(W)$ can be calculated as well. Computing its variance, $\Delta_C^g W$, yields (see (G.12)):

$$\Delta_C^g W = \sqrt{\frac{2d_C(d - d_C)}{d^2(d + 2)}} \quad (5.78)$$

If the trajectories in Hilbertspace were ergodic, $\Delta_C^g W$ would be a direct measure of the fluctuations. Since we know that the trajectories are not ergodic, such an unbiased connection is not necessarily true. Nevertheless, if $\Delta_C^g W$ is small, almost all states within the accessible region feature a $W(E_C^g)$ very near to $W^d(E_C^g)$, such that fluctuations can be expected to be small. In that sense $\Delta_C^g W$ can be considered a measure of the fluctuations. If the environment gets bigger, both d and d_C , will scale with some factor, say, A . If A is large $\Delta_C^g W$ will eventually scale like (see (G.13))

$$\Delta_C^g W \propto \sqrt{\frac{1}{A}} \quad (5.79)$$

and thus, typically, fluctuations as well.

Coming, just for completeness, back to the case of the whole system having some energy probability distribution $W(E)$ that is not exactly sharp, we can assign to any energy subspace E its own dimension d^E and the dimension of its subset that corresponds to system g , being at energy E_C^g , d_C^E . Thus each energy subspace has its own relative frequency $f_C^E(W)$ (we now suppress the gas index g for simplicity) of the probability $W^E(E_C^g)$. This leads to a set of variances $\Delta_C^E W$. Eventually its the variance of the relative frequency of $W(E_C^g)$, the overall probability to find system g at E_C^g , that one is interested in. Since this variable reads

$$W(E_C^g) = \sum_E W(E)W^E(E_C^g) \quad (5.80)$$

and the energy subspaces are entirely independent of each other, we finally get:

$$\Delta_C^g W = \sqrt{\sum_E (W(E)\Delta_C^E W)^2} \quad (5.81)$$

For some more concrete illustration of the implications that the rather abstractly derived principles in this chapter bare on the fluctuations in adequate systems, see **5.8.3**.

5.7 Local Equilibrium States and Ergodicity

As explained in 4.4, no ergodicity is needed in our approach. Nevertheless concepts like ergodicity or the “a priori postulate” have been in the center of the discussion in the past. Here we want to explain, at least partially, the connection.

Given that a compound system meets the requirements of a thermodynamical system(see

9), the local considered system will, as explained, under microcanonical conditions, most likely be found in the momentary equilibrium state:

$$\hat{\rho}_{eq}^g \approx \sum_{A,a} \frac{W^g(E_A^g)}{N^g(E_A^g)} |A, a\rangle \langle A, a| \quad (5.82)$$

Obviously all states $|A, a\rangle$, which all belong to one “energy shell” E_A^g , have got the same probability $\frac{W^g(E_A^g)}{N^g(E_A^g)}$. Which means that this state is the same as the one one would have found, if one had applied the a priori postulate. It is also the same state that one would have got, if one had assumed perfect ergodicity for the single, microscopically isolated system, and time averaged over a long enough period.

It might be worth mentioning that, in order to obtain ρ_{eq}^g it is even irrelevant in which space one assumes this ergodicity: One could only take the orthogonal energy basis states as the possible states of the system, in each of which the system will be found for the same time, or one could allow for superpositions and take the Hilbertspace of the local system in the described parametrization for the space in which ergodicity is assumed, in both cases one would find ρ_{eq}^g as a result.

To sum it up, one can say that the effect of the adequate, microcanonically coupled environment is, that a considered system locally behaves, as if it was perfectly ergodic, within arbitrarily short time periods. Or one could describe the result by an ensemble, many copies of the same system, the same amount of copies being in every pure state that is accessible to the system. But in a larger picture the one considered system is simply highly entangled with its environment.

For a thermodynamical system under canonical conditions, the momentary equilibrium state is:

$$\hat{\rho}_{eq}^g \approx \sum_{A,a} \sum_E \frac{N^c(E - E_A^g) W(E)}{N(E)} |A, a\rangle \langle A, a|. \quad (5.83)$$

Like in the microcanonical case, this local state is the same that one would have gotten, if one had given all states that belong to the same energy shell of the full system, E , the same probability. In this case, the local equilibrium gas state is, as if the full system was perfectly ergodic, although it is not.

For other analyzations of weakly coupled system-surrounding scenarios, using entirely different techniques, but yielding to some extend results, similar to those described in this chapter, see [35–37]

5.8 Numerical Results

To illustrate the principles that have been analyzed analytically and in a rather abstract way in the last chapters, (5.4, 5.5, 5.6), we now want to present some numerical data in the following sections (5.8.1, 5.8.2, 5.8.3). These data refer to a certain type of models. Those models are still rather abstract, and may thus be viewed as models for a multitude of concrete physical situations. Those models always consist of two subsystems, the considered system, (gas) and the environment, (container). Those subsystems

are only specified by their spectra, or rather those parts of their spectra that play any role at all under the condition of energy conservation. What sort of physical structure gives rise to those spectra is not taken into account here, since it turns out to be irrelevant to some extent. Since we want to analyse the models on the level of discrete, closed, quantum mechanical Hamilton models, for which a discrete, finite dimensional Schrödinger equation has to be solved, those spectra are translated into discrete, diagonal, matrix-Hamiltonians, that describe the decoupled bi-partite system. One can choose them diagonal without any loss of generality, since any such system may be analysed in the energy eigenbasis of its decoupled parts. The concrete form of the interaction now depends on the concrete physical subsystems and their concrete interactions. But since the “guess” is, that for the quantities considered here (entropy, occupation probabilities, etc.) the concrete form of the interaction does not matter, the interaction is chosen to be some random matrix to avoid any specialization. It only has to be “small” in some sense. Other than that the interaction is not specified. All that is done to get a concrete model that is nevertheless as typical for thermodynamical situations in general as possible

5.8.1 Numerical Results for microcanonical Conditions

All data in this example refers to a situation depicted in **Fig.5.1**. The gas (the system under consideration) consists of a two-level-system, both levels nondegenerate, the container (the environment) consists of only one energy level with a degree of degeneracy of $N^c = 50$. Since the container consists of only one energy level, this is necessarily a microcanonical situation regardless of the interaction \hat{I} . Since the container cannot absorb any energy, energy cannot be exchanged between the systems. In this situation the probabilities to find the gas system in the ground(excited) state $W_0^g(W_1^g)$ are conserved quantities and in this example chosen as

$$W_0^g = 0.15 \quad W_1^g = 0.85 \quad (5.84)$$

Since

$$N_0^g = N_1^g = 1 \quad N^c = 50 \quad (5.85)$$

we have

$$N_0^g N^c + 1 = 51 \approx N_0^g N^c = 50 \quad N_1^g N^c + 1 = 51 \approx N_1^g N^c = 50 \quad (5.86)$$

and thus (5.48) applies. Using (5.48) and (5.30) we find

$$\langle P^g \rangle = 0,765 \quad P_{min}^g = 0,745 \quad (5.87)$$

which means that

$$\langle P^g \rangle \approx P_{min}^g \quad (5.88)$$

and we thus have, as explained in the last chapter, a situation in which almost the entire accessible region (see (5.28)) should be filled with the compartment containing

only states of almost maximum local entropy. To examine this, a set of random states which are uniformly distributed over the accessible region has been generated. Then the local entropies of all those states have been calculated and they have been sorted into a histogram, according to their local entropies. Since those states are distributed uniformly over the accessible region, the number of states in one “entropy channel” corresponds to the relative size of the respective Hilbertspace compartment. The histogram is shown in **Fig.5.2**. The maximum local entropy in this case is $S_{max} = 0.423k$. Obviously almost all states have local entropies close to S_{max} . Thus compartments corresponding to entropies of, say, $S > 0.4k$ fill almost the entire accessible region, just like theory predicts. Locally pure states ($S = 0$) are practically of the measure zero.

In order to examine dynamics a coupling \hat{I} has been introduced. To keep the concrete example as general a possible, \hat{I} has been chosen as a random matrix in the basis of the energy eigenstates of the uncoupled system, with Gaussian distributed real and imaginary parts of the matrixelements, with a standard deviation of

$$\Delta I = 0.01\Delta E \quad (5.89)$$

thus in this sense the coupling is weak, compared to the Hamiltonian of the uncoupled system. Now the Schrödinger equation for this system, including the interaction, has been solved for initial states consistent with (5.84). Then the local entropy at each time has been calculated, thus getting a picture of the entropy evolution. The result is shown in **Fig.5.3**. Obviously the entropy reaches almost S_{max} within a reasonable time, regardless of the concrete initial state (The different lines in **Fig.5.3** refer to different initial states). Thus the tendency towards equilibrium is obvious. The concrete form of the interaction \hat{I} has only influence on the details of this evolution, the equilibrium value is the same. If the interaction is generally made weaker, only the timescale on which equilibrium is reached gets longer, but eventually maximum entropy will always be reached.

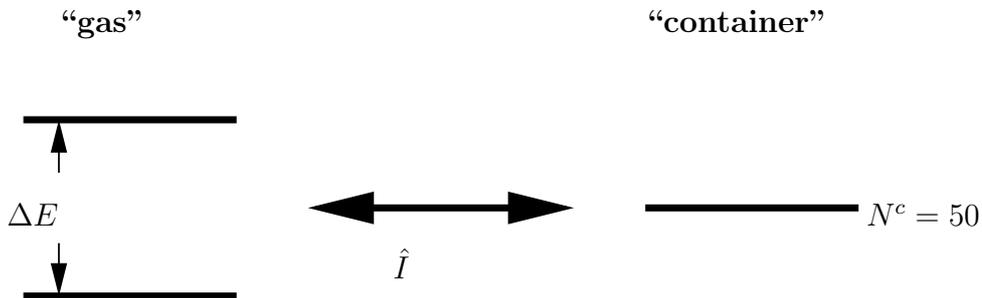


Fig. 5.1: Microcanonical situation. A nondegenerate two-level-system (gas) is weakly coupled to a system with one energylevel of degeneracy $N^c = 50$. This is a model for a system in contact with a much larger environment, such that no energy can be exchanged between the system and its environment.

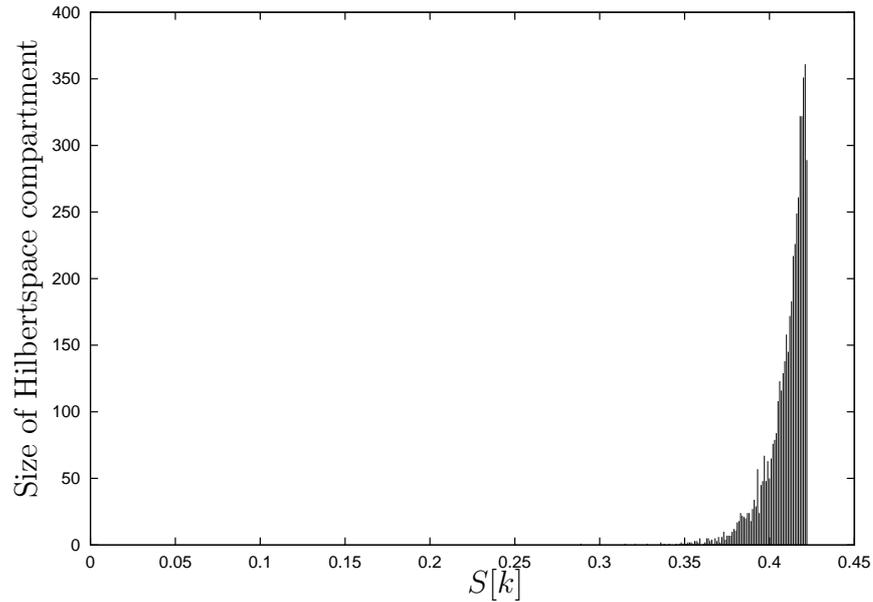


Fig. 5.2: Relative sizes of Hilbertspace compartments. This histogram shows the relative frequency of states with a given local entropy S , among all states from the accessible region. In this case the maximum possible entropy is $S_{max} = 0.423k$. obviously almost all states feature entropies close to that maximum one.

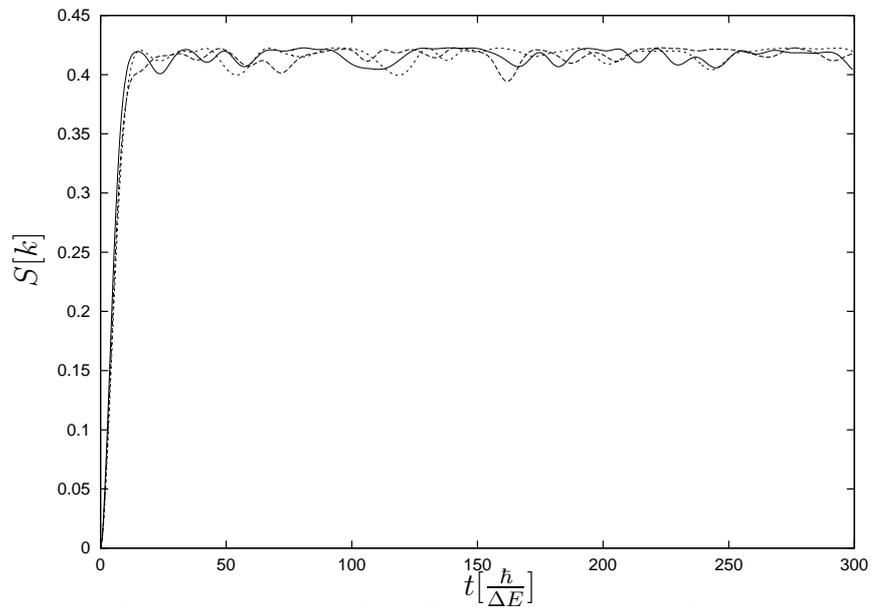


Fig. 5.3: Evolution of the local entropy for different initial states. A universal state of maximum entropy (equilibrium) is reached, independent of the initial state.

5.8.2 Numerical Results for Canonical Conditions

Just like in 5.8.1 we now want to present some numerical data to back up the principles derived in 5.5. The first model which was analyzed numerically is depicted in Fig. 5.4. The considered (gas) system, again, consists only of a nondegenerate two-level-system. The environment (container) system in this case consists of a three-level-system with an exponential “statedensity”: $N_B^c = 50 \cdot 2^B$. This has been chosen since theory predicts for such a degeneracy scheme of the environment an equilibrium state of the gas system which is independent of its initial state (see (5.67)). If we restrict ourselves to initial states featuring arbitrary states for the gas system but container states that only occupy the intermediate energy level, no other container levels except for the given ones could be reached, even if they were present, due to energy conservation. (This being true in the limit of weak interactions \hat{I} .) Thus in this case, the model can also be seen as a model for a situation with a much larger environment. For this case we find from (5.67)

$$W(E_0^g) = \frac{2}{3} \quad W(E_1^g) = \frac{1}{3} \quad (5.90)$$

To keep the situation as general as possible \hat{I} was, like in 5.8.1 chosen to be a matrix with random Gaussian distributed entries in the basis of the eigenstates of the uncoupled system. For this system the Schrödinger equation has been solved and the evolution of the probability to find the gas system in its ground state, $W(E_0^g)$ is displayed in Fig. 5.5. The different lines correspond to different interaction strengths, given by the standard deviation of the distribution of the matrix elements of \hat{I} , ΔI :

$$\Delta I_{solid,dashed} = 0.0075\Delta E \quad \Delta I_{dotted} = 0.002\Delta E \quad (5.91)$$

Obviously the equilibrium value of $W(E_0^g) = \frac{2}{3}$ is reached independent of the of the concrete interaction \hat{I} , the interaction strength, within the weak coupling limit, only influences the timescale on which equilibrium is reached.

Fig. 5.6 displays the evolution of the same probability, $W(E_0^g)$, but now for different initial states, featuring different probabilities for the groundstate, as can be seen in the figure at $t = 0$. The equilibrium value is reached for any evolution, regardless of the special initial state, thus we see the effective attractor behavior, which is a typical for thermodynamics.

Fig. 5.7 displays the evolution of the local entropy of the gas system for the same three initial states used for Fig. 5.6. The maximum entropy that is consistent with the equilibrium value of the energy probabilities is $S_{max} = 0.637k$. This is also the value one finds, if one maximizes entropy while keeping the mean energy fixed (Jayne’s principle). Obviously this value is reached for any initial state during the concrete dynamics of this model. This supports the validity of (5.69), which states that the density matrix of the equilibrium state is diagonal in the basis of the local energy eigenstates.

To analyze the formation of a full Boltzmann distribution, we investigated the system depicted in Fig. 5.8. Here the gas system is a nondegenerate, equidistant five-level-system, the container system is a five-level-system as well, but its degrees of degeneracy are: $N_B^c = 6 \cdot 2^B$, which should lead to a Boltzmann distribution. We restrict ourselves to

initial states where for both subsystems only the intermediate energy level is occupied in the initial state (symbolized by the black dots in **Fig. 5.8**). In this case, due to energy conservation, other states of the container system would not play any role even if they were present, just like in the previous model. **Fig. 5.9** Shows the probabilities of the different energy levels to be occupied $W(E_A^g)$. Initially, as already mentioned, the gas system starts in the intermediate (third) energy level. But soon a Boltzmann distribution develops. The probability that reaches the highest value, corresponds to the lowest energy level, respectively for the other probabilities. Obviously each probability is twice as high as the one for the level above. This is exactly what theory predicts (see (5.64) for the environment degeneracy scheme in this model.

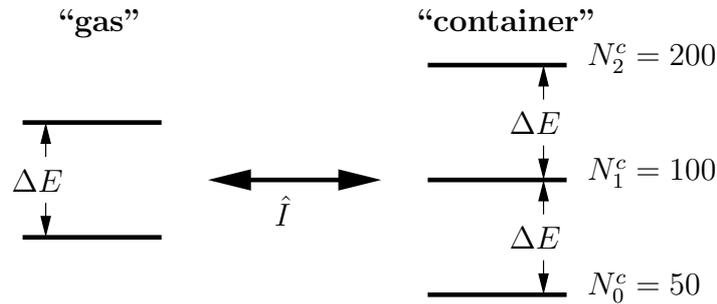


Fig. 5.4: Canonical situation. A two-level gas system is weakly coupled to a three level environment, such that energy can be exchanged. The exponential degeneracy scheme of the container system guarantees a full independence of the equilibrium state from the initial state.

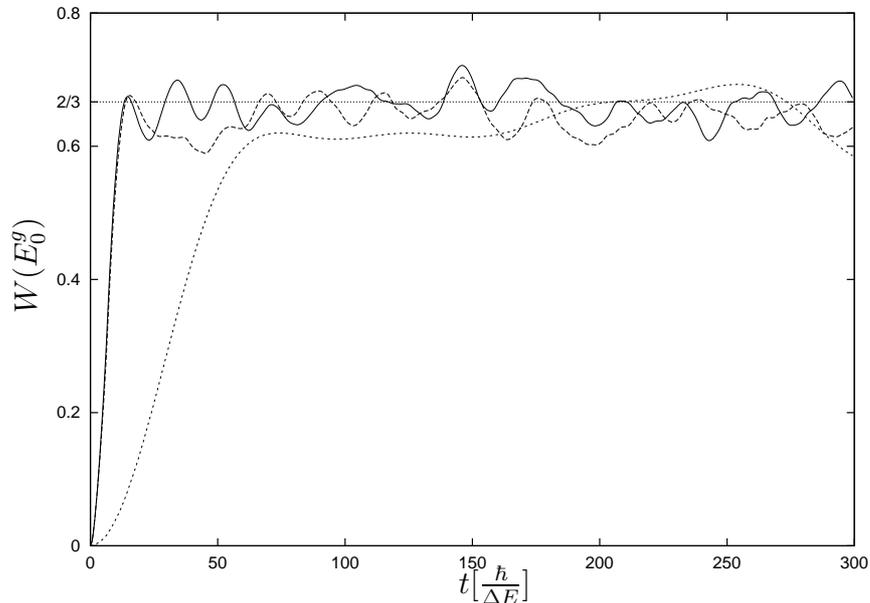


Fig. 5.5: Evolution of the ground level occupation probability for different interactions. The theoretically predicted equilibrium value is reached for different random interactions. The dotted line corresponds to a weaker interaction. Even in this case the same equilibrium value is reached, only on a longer timescale

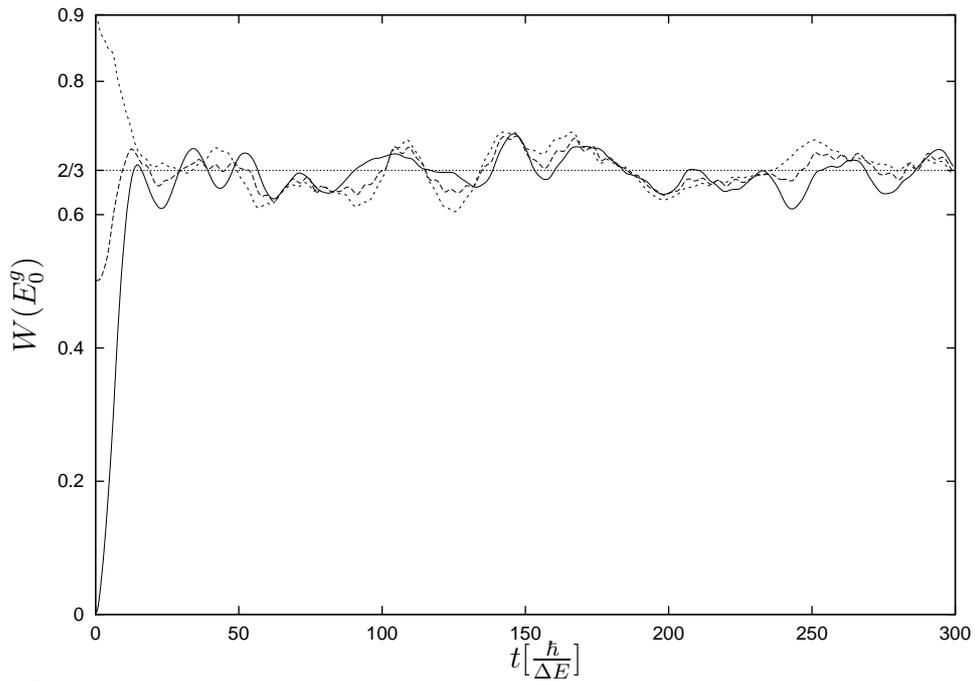


Fig. 5.6: Evolution of the ground level occupation probability for different initial states. The theoretically predicted equilibrium value is reached, independent of the initial states, which is typical for canonical conditions

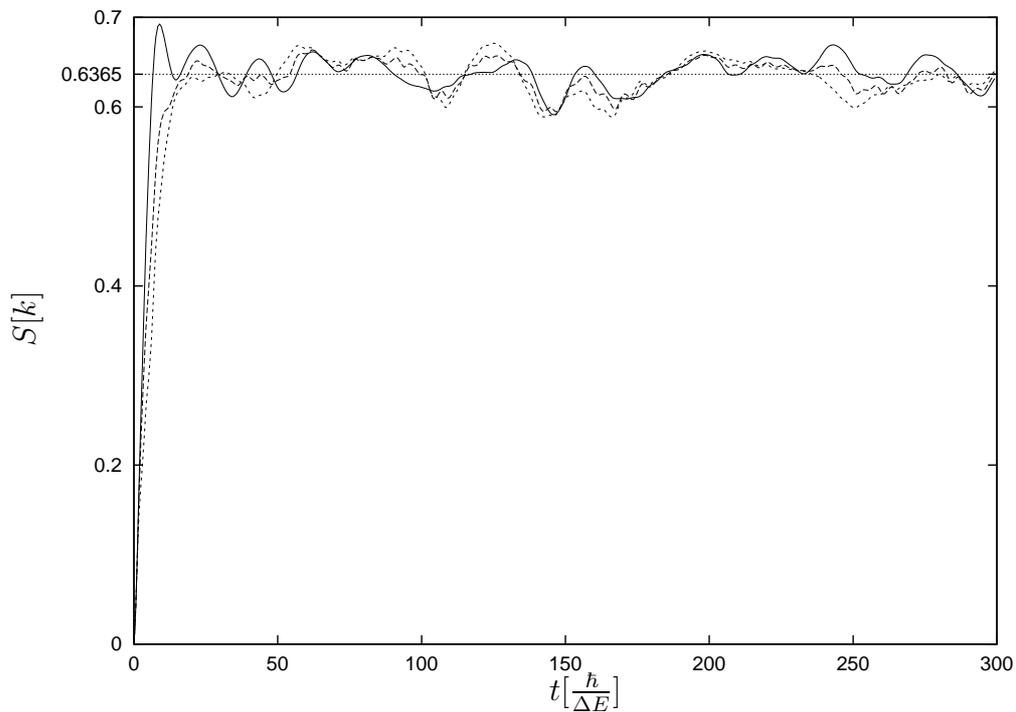


Fig. 5.7: Evolution of the local entropy for different initial states. $S = 0.637k$ is the maximum entropy that is consistent with the equilibrium energy probabilities. Thus, during the evolution a maximum entropy state is reached.

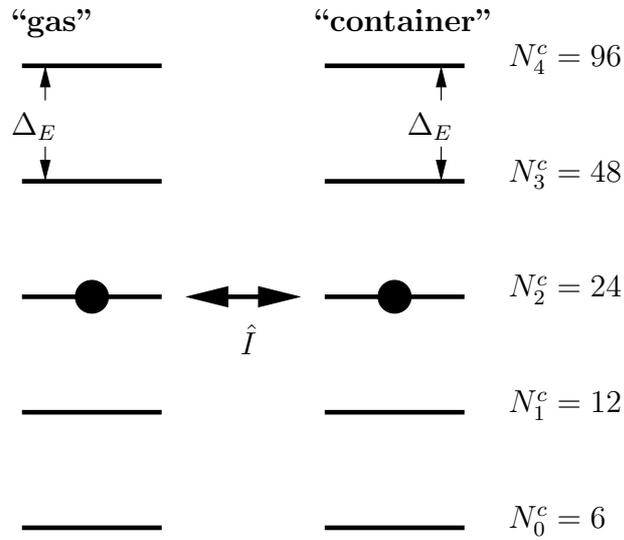


Fig. 5.8: Canonical multi-level situation. A five-level gas system is weakly coupled to a five level container system with an exponential degeneracy scheme, such that energy may be exchanged. This set-up should lead to a Boltzmann distribution

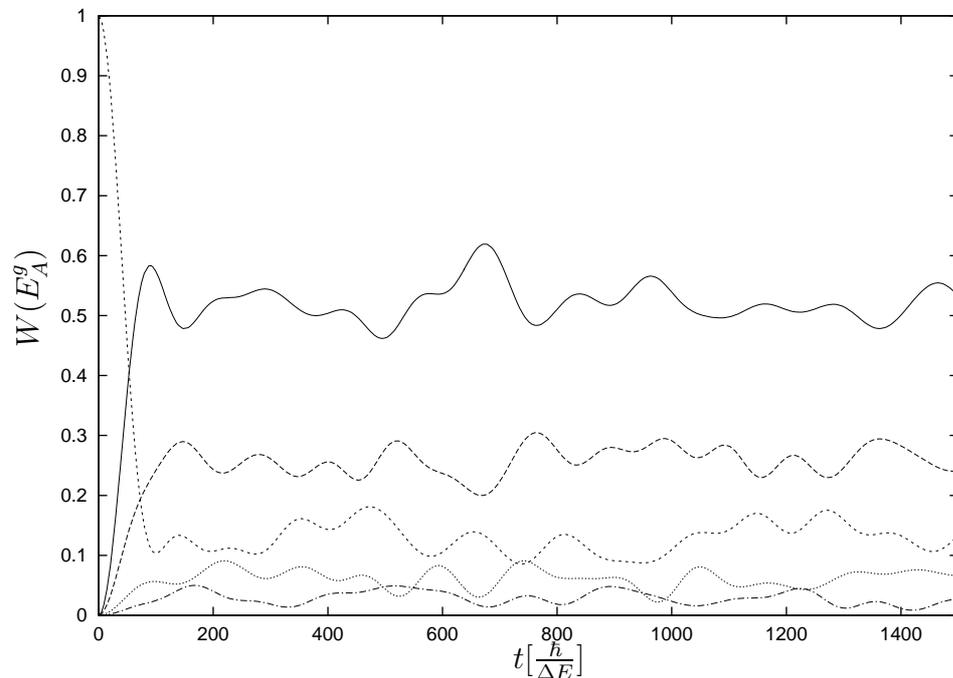


Fig. 5.9: Evolution of the energy probabilities. After some relaxation time a Boltzmann distribution is reached. Each probability is as high as the one for the next higher energy level, as theory predicts.

5.8.3 Numerical Results for Probability Fluctuations

To back up the previous section 5.6 by some numerical data, a system almost like the one depicted in **Fig. 5.4** is analyzed. The only difference is the different degeneracy scheme that is now:

$$N_B^c = \frac{N_1^c}{2} \cdot 2^B \quad (5.92)$$

that means, the proportions between the degrees of degeneracy of the different container levels are the same as for the system sketched in **Fig. 5.4** but the overall size of the container system is tuneable by N_1^c . For this system with various different N_1^c , the Schrödinger equations have been solved numerically, and the following measure of the fluctuations of the occupation probability of the ground level of the gas system has been computed:

$$\Delta_t^2 W_0^g := \frac{1}{t_f - t_i} \left(\int_{t_i}^{t_f} (W_0^g(t))^2 dt - \left(\int_{t_i}^{t_f} W_0^g(t) dt \right)^2 \right) \quad (5.93)$$

For initial states with

$$W_0^g(0) = 0.2 \quad W_1^g(0) = 0.8 \quad (5.94)$$

Fig. 5.10 now shows the dependence of the size of these fluctuations on the container system size N_1^c . The little crosses are the computed data points, the dashed line is a least square fit to a function proportional to $\frac{1}{N_B^c}$. Obviously the fit fits very well, which confirms that fluctuations should vanish like the squareroot of the system size (see (5.79)). The fit reads

$$\Delta_t W_0^g(N_1^c) = \sqrt{\frac{0.053}{N_1^c}} \quad (5.95)$$

Computing the variance of the relative frequency of states, featuring some W_0^g among the states of the accessible region, for the situation described above, we get, using the techniques of the previous section, 5.6, especially (5.81),

$$\Delta_0^g W(N_1^c) = \sqrt{\frac{0.053}{N_1^c}}, \quad (5.96)$$

which means that $\Delta_t W_0^g(N_1^c)$ equals to very good accuracy $\Delta_0^g W(N_1^c)$, although the trajectories are not ergodic.

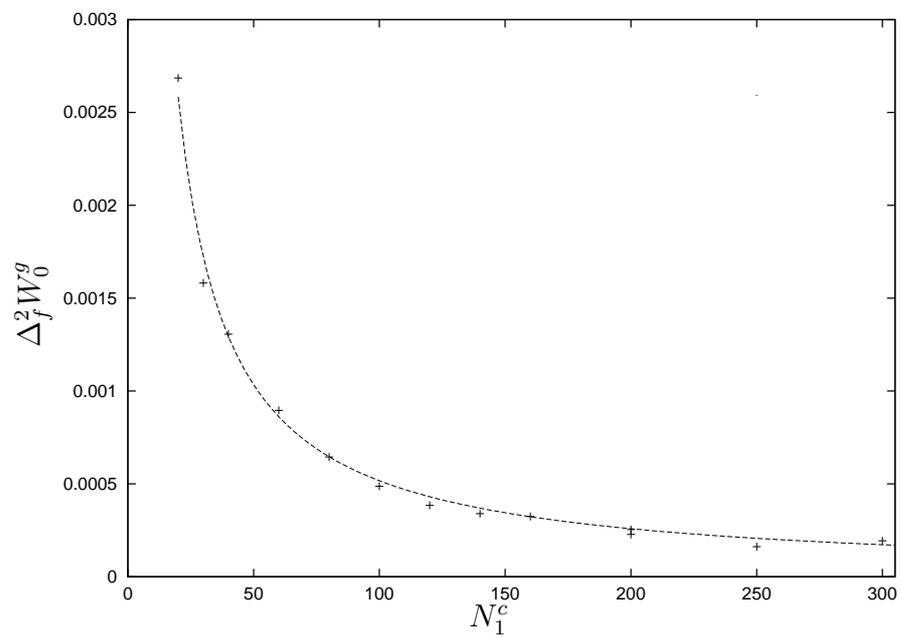


Fig. 5.10: Fluctuations of the probability of of the considered system to be in the ground state $\Delta_f W_0^g$, in dependence on the number of eigenstates of the environment system N_1^c . Obviously the fluctuations vanish, with the environment getting bigger.

6 Typical Spectra of Large Systems

6.1 The Extensivity of Entropy

If a set of axioms as a basis of thermodynamics is given, one often reads that entropy has to be an extensive quantity. This basically means that if one brings to identical systems in contact, such as to form a system of twice the size of the original systems, the entropy of the joint system should be twice as high as the entropy of the original systems. Formulated more mathematically this means that that entropy should be a homogeneous function of the first grade, or that is should be possible to write it as a function of the other extensive variables ,say, energy U , volume V and particle number N as:

$$S = Ns\left(\frac{U}{N}, \frac{V}{N}\right) \quad (6.1)$$

This is obviously an important property, since it guarantees, e.g., that temperature, defined in the usual way

$$T = \frac{\partial U}{\partial S} \quad (6.2)$$

remains the same, if the system is doubled in the above described way, i.e., that temperature is an intensive quantity.

But this basic claim already seems to cause problems for the standard definition of entropy. The standard classical definition of entropy for the microcanonical ensemble reads

$$S = k \ln g(U) \quad (6.3)$$

where $g(U)$ is the number of micro states that is consistent with the energy U , i.e., the volume of the corresponding energy shell in Γ -space, divided by the volume of some elementary cell. In our approach the same formula holds (for a sharp energy probability distribution) for the equilibrium entropy (see (5.50)), except for here $g(U)$ is the quantum mechanical energy statedensity at the energy U .

In both cases, if one assumes that the thermal contact that unifies two identical systems to one bigger system allows for energy exchange between the systems, but contains only negligible amounts of energy in itself, the entropy of the double system at the double energy should be calculated from the state density by a convolution:

$$S_2 = k \ln \int_0^{2U} g(E)g(2U - E)dE. \quad (6.4)$$

It is obvious that this, in general, cannot be twice the entropy of one of the separate systems, for

$$k \ln \int_0^{2U} g(E)g(2U - E)dE \neq 2k \ln g(U) \quad (6.5)$$

This could be true if the function $g(E)g(2U - E)$ would be extremely peaked at $E = U$ but, again, in general, there is no further reason to assume this, even if g is a rapidly growing function, as often claimed in standard textbooks. If g grows exponentially, as often assumed, the integrand of the convolution is flat, rather than peaked at all. The identity of (6.5) is often claimed by referring to the ideal gas, for which it happens to be approximately true, or by complicated considerations involving the canonical ensemble [39]. All this, however, is no straight, general derivation for the microcanonical case. So according to those definitions, one cannot claim without further considerations that entropy is an extensive quantity. This problem is not to be confused with Gibbs paradoxon that can be solved by using Boltzmann statistics of identical particles; here dividing the left hand side of (6.5) by some function of N will not fix the problem. It has often been claimed that Shannon's entropy

$$S^r = -k \sum_n W_n^r \ln W_n^r \quad (6.6)$$

was necessarily extensive, since $S^{12} = S^1 + S^2$ if $W_n^{12} = W_i^1 W_j^2$ which means that the probabilities to find the systems in their individual states are uncorrelated. This, however, is not the case in the microcanonical ensemble. If one system is found at the energy E , the other one necessarily has to be at the energy $U - E$.

What thus remains to be shown is if, and under what condition, S can be a homogeneous function of U .

6.2 Spectra of Modular Systems

Practically all of the matter we encounter in nature, seems to have some sort of modular structure. Gases are made of weakly interacting identical particles. Crystals are periodic structures of, maybe strongly interacting, identical units, even disordered matter like glass or polymers can be split up into fairly small parts without changing the properties of the parts essentially.

Let us, as an example, consider the successive built-up of some piece of metal: First, we have one atom with some energy spectrum. If we bring two atoms together the spectrum of the resulting molecule will be considerably different from the energy spectrum of the two separate atoms. The energy contained in the binding can be as large as the structures of the spectrum of the separate atoms. But the spectrum of the molecule will already be broader than the spectra of the separate atoms. If we now combine two 2-atom molecules to one 4-atom molecule the spectrum of the 4-atom molecule will again be considerably different from the spectrum of the two separate 2-atom molecules. If we continue this process, at some point, say if the separate parts contain a hundred atoms

or so each, the separate parts will already have wide energy spectra, containing, maybe, bands that stretch over a considerable energy region with a smooth statedensity. If we now combine these parts, the energy contained in the binding will be negligible compared to the structures of the energy spectrum of the two separate parts. Most of the atoms in one part do not even feel the force of the atoms in the other part anymore, simply because they are too far away. Thus, the energy distortion caused by the binding to the spectrum of the two separate parts, will be negligible. This is the limit beyond which the weak coupling limit, that is always assumed in thermodynamics, applies. For the contact between a system and its environment it is always assumed that the spectra of the separate systems are almost undistorted by the contact. So this principle should also apply to the different identical parts of one system above some size. Here we assume that there are a lot of this units to be combined above this limit, to form a macroscopic system, as it is the case in our example where there are a lot of parts containing a hundred atoms, to be combined to form a piece of metal, containing on the order of 10^{23} atoms.

The bottom line is that the spectrum or state density of any macroscopic system, can be viewed as the spectrum of a system consisting of very many almost interaction free parts, even if the basic particles might be strongly interacting. In the case of a gas, no such consideration is necessary for the spectrum of the gas can naturally be understood as the combined spectrum of all individual spectra of all individual gas particles.

Now we want to analyze the properties of spectra that result from very many identical non-interacting systems. Just like the state density of two non interacting systems should be the convolution of the two individual state densities, the state density of the modular system, $G(U)$, should be the convolution of all individual state densities, $g(U)$. Defining:

$$C_N^g(U') := \int g(E_1)g(E_2 - E_1)g(E_n - E_{n-1}) \cdots g(U' - E_{N-1}) \prod_{n=1}^{N-1} dE_n \quad (6.7)$$

as the convolution of N identical functions g we can write

$$G(U') = C_N^g(U') \quad (6.8)$$

To evaluate this convolution, we want to start by considering another convolution. Therefore we define:

$$r(E) := \frac{e^{-\alpha E} g(E)}{\int e^{-\alpha E} g(E) dE} \quad (6.9)$$

and the quantities

$$A := \int e^{-\alpha E} g(E) dE \quad m = \int E r(E) dE \quad \sigma^2 := \int E^2 r(E) dE - m^2. \quad (6.10)$$

If the increase of $g(E)$ with energy is not faster than exponential, which we have to assume here, than all defined quantities are finite and, since $r(E)$ is normalized, m is

the mean value of $r(E)$ and σ is the variance of $r(E)$. Now, consider the convolution of all $r(E)$, this can be written as

$$C_N^r(U') = \frac{e^{-\alpha U'} G(U')}{A^N} \quad (6.11)$$

To evaluate $C_N^r(U')$ we now gather properties of it that can be deferred from its being a convolution:

Since the integral over a convolution equals the product of the integrals of the convoluted functions, we find:

$$\int C_N^r(U') dU' = 1 \quad (6.12)$$

Since the mean value of a convolution of normalized functions is the sum of the mean values of the convoluted functions, we find:

$$M := \int U' C_N^r(U') dU' = Nm \quad (6.13)$$

Since the square of the variance of a convolution of normalized functions is the sum of the squares of the convoluted functions, we find

$$\Sigma^2 := \int U'^2 C_N^r(U') dU' - N^2 m^2 = N\sigma^2 \quad (6.14)$$

An important property of a convolution is that the Fourier transform of two convoluted functions equals the product of the Fourier transforms of the convoluted functions. Thus, if we for simplicity, define the Fourier transform of a function $f(E)$ as $f^F(\omega)$, we find:

$$C_N^{r^F}(\omega) = (r^F(\omega))^N. \quad (6.15)$$

If the the function $r^F(\omega)$ has one global maximum somewhere, (which is very likely, since if $r(E)$ is integrable, $r^F(\omega)$ is integrable as well) this maximum should become by far higher than any other point of the function, if the function is multiplied very many times with itself, regardless of how much higher than the rest of the function the maximum originally was. This means that the function $C_N^{r^F}(\omega)$ should be extremely sharp peaked at some point, if N is large enough. One can show (see **D**) that this peak, that should contain almost all of the area under the curve, is essentially a Gaussian. One can now split $C_N^{r^F}(\omega)$ up into two parts, one that is truly Gaussian and the rest. As already explained the area under the square of the rest should be much smaller than the area under the square of the Gaussian. Since a Fourier transform is additive, leaves the area under the square of the curve invariant, and transforms a Gaussian into a Gaussian, $C_N^r(U')$ should consist of a large part that is Gaussian, and a small part that cannot be determined, but gets smaller and smaller as N gets larger. In the region in which the Gaussian is peaked, $C_N^{r^F}(\omega)$ should be almost entirely dominated by the Gaussian part, at the edges, where the Gaussian vanishes, the small rest may dominate. If we assume that the integral, the mean value and the variance of $C_N^{r^F}(\omega)$ are entirely dominated by

its Gaussian part, we can, using (6.12),(6.13) and (6.14), give a good approximation for $C_N^{r,F}(\omega)$ that should be valid at the peak, i.e., around $U' = Nm$:

$$C_N^r(U') \approx \frac{e^{-\frac{(U'-Nm)^2}{2N\sigma^2}}}{\sqrt{2\pi N}\sigma}. \quad (6.16)$$

Solving (6.11) for $G(U')$ and inserting (6.16), evaluated at the peak we find:

$$G(Nm) \approx \frac{A^N e^{\alpha Nm}}{\sqrt{2\pi N}\sigma} \quad (6.17)$$

m , A and σ are all functions of α . Thus, we have got now G as a function of α . Since we want G as a function of the inner energy, U we define

$$U := Nm(\alpha) \quad \text{or} \quad \frac{U}{N} = m(\alpha). \quad (6.18)$$

Solving formally for α we get

$$\alpha = m^{-1}\left(\frac{U}{N}\right) = \alpha\left(\frac{U}{N}\right). \quad (6.19)$$

Now A , σ and α are all functions of the argument $\left(\frac{U}{N}\right)$ and we can rewrite (6.17)

$$G(U) \approx \frac{A\left(\frac{U}{N}\right)^N e^{\alpha\left(\frac{U}{N}\right)U}}{\sqrt{2\pi N}\sigma\left(\frac{U}{N}\right)}, \quad (6.20)$$

or taking the logarithm

$$\ln G(U) \approx N \left(\ln A\left(\frac{U}{N}\right) + \left(\frac{U}{N}\right) \alpha\left(\frac{U}{N}\right) \right) - \frac{1}{2} \ln(2\pi N) - \ln \sigma\left(\frac{U}{N}\right). \quad (6.21)$$

If we keep $\left(\frac{U}{N}\right)$ fixed, but let N be extremely large compared to one, $N \gg \gg 1$, which amounts to a simple upscaling of the system, we can neglect everything, except for the first part on the right hand side of (6.21) to get:

$$\ln G(U) \approx N \left(\ln A\left(\frac{U}{N}\right) + \left(\frac{U}{N}\right) \alpha\left(\frac{U}{N}\right) \right) \quad (6.22)$$

This now is obviously a homogeneous function of the first grade and thus an extensive quantity.

The joint spectrum of a few non or weakly interacting systems does not give rise to an extensive entropy, according to the standard definition of entropy, but the spectrum of very many non or weakly interacting subsystems always will, regardless of the form of the spectrum of the individual subsystem, that the joint system is made of.

Simply to check (6.22) we can apply it to an ideal gas, just taking the spectrum of a free

particle in one dimension as the function to be convoluted. This procedure yields (see **E**):

$$\ln G(U) = N \left(3 \ln \frac{L}{h} + \frac{3}{2} \ln \left(\frac{4m\pi U}{3N} \right) + \frac{3}{2} \right) \quad (6.23)$$

where m is the mass of a single particle, which is exactly the same form one finds in standard textbooks (before the application of the corrected Boltzmann statistics, see, e.g., [40], where it is usually calculated by evaluating the surface area of hyperspheres and using the Stirling formula.

6.3 The Boltzmann Distribution

In **5.5.4** we found an equilibrium energy probability distribution for a canonical situation, i.e., a system in contact that allows for energy exchange with a large surrounding. This distribution is given by (5.64) and reads, written in terms of state densities

$$W^d(E^g) = G^g(E^g) \int \frac{G^c(E - E^g)W(E)}{G(E)} dE \quad (6.24)$$

Where $G^g(E^g)$ and $G^c(E^c)$ are the state densities of the gas and the container system, respectively, whereas $G(E)$ is the state density of the full combined system. This is obviously not the familiar Boltzmann distribution. Instead of the Boltzmann factor $e^{-\frac{E_g}{kT}}$, one has here a factor depending on the state density of the environment. We are going now to analyze this factor under the assumption, that the spectrum of the environment has the structure that we established as typical in **6.2**.

According to (6.22), if the environment is a large system, it should be possible to write the logarithm of its state density as

$$\ln G^c(E^c) = N s^c \left(\frac{E^c}{N} \right) \quad (6.25)$$

where N is the number of some basic units of the environment. If one looks at the graph of such a homogeneous function for different N , it is plain to be seen that increasing N , just amounts to an upscaling of the whole picture. That means that the graph becomes smoother and smoother within finite energy intervals. This can be stated in a more mathematical form by looking at an expansion of $\ln G^c(E^c)$ around some point of fixed energy per unit $\left(\frac{E^c}{N}\right) = \epsilon$

$$\ln G^c(E^c) \approx N s^c|_{\epsilon} + \left. \frac{ds^c}{dE^c} \right|_{\epsilon} (E^c - N\epsilon) + \frac{1}{2N} \left. \frac{d^2 s^c}{d^2 E^c} \right|_{\epsilon} (E^c - N\epsilon)^2 + \dots \quad (6.26)$$

Evidently already the second order term scales with N^{-1} , terms of order n scale with N^{1-n} . Therefore, for large N , a truncation of the expansion, only keeping the linear term will be a valid approximation over wide energy ranges, the range of validity will be the wider, the larger N is. Without $\ln G^c(E^c)$ being a homogenous function of the

first grade, such a truncation remains without justification, although it is often routinely done [39].

In (6.24) the function $\frac{W(E)}{G(E)}$ is multiplied with the environment state density under the integral. If the range in which this function is peaked, or takes on values that differ significantly from zero, is smaller than the range over which a linearization of $\ln G^c(E^c)$ is valid, we might replace the state density of the environment in (6.24, by an exponential that the linearization gives rise to. The expansion then has to be around the energy where $\frac{W(E)}{G(E)}$ is peaked, i.e., $N\epsilon$ has to be chosen to lie in the center of the peak. This replacement yields:

$$W^d(E^g) \approx G^g(E^g) \int \frac{e^{N s^c|_{\epsilon} + \frac{ds^c}{dE^c}|_{\epsilon}((E-E^g)-N\epsilon)} W(E)}{G(E)} dE \quad (6.27)$$

or if we take functions of E^g out of the integral

$$W^d(E^g) \approx G^g(E^g) e^{-\frac{ds^c}{dE^c}|_{\epsilon} E^g} \int \frac{e^{N s^c|_{\epsilon} + \frac{ds^c}{dE^c}|_{\epsilon}(E-N\epsilon)} W(E)}{G(E)} dE. \quad (6.28)$$

To simplify this even further, we define $\beta = \frac{ds^c}{dE^c}|_{\epsilon}$ and rewrite (6.28) as

$$W^d(E^g) \propto G^g(E^g) e^{-\beta E^g} \quad (6.29)$$

which is exactly the wellknown Boltzmann distribution.

6.3.1 Beyond the Boltzmann Distribution?

Whithin the last years the idea of overcoming the standard limits of thermodynamics by exploiting the laws of quantum mechanics has attracted considerable attention [30,41,42]. Thus, it should be pointed out here, that in the frame of the ideas presented here, the Boltzmann distribution does not follow naturally from some basic principles like it does from the maximum entropy principle in the context of Jayne's principle, but is due to the special structure of the spectra of the systems, that represent the environments. If a system is in contact with a system that is not built according to the scheme described in **6.2**, it could have a stable equilibrium energy probability distribution that differs from the Boltzmann distribution. In fact any distribution that is described by (5.64) is considered stable, as long as the statedensity of the container system is large enough. Thus if one could create a system with a high state density but not successively built up in a modular style, one could get a equilibrium distribution that differs significantly from a Boltzmann distribution. But creating such a system is probably very hard. It would either have to be impossible to split it up into identical parts, or, the parts would have to interact strongly over large distances. Furthermore, one would have to decouple this system entirely from any further system, including the electromagnetic field. Although all this seems pretty unfeasible, this effect might be seen in some future experiment.

7 Temperatures

If it is hard to define entropy as a function of the micro state on the basis of classical mechanics, it is even harder to define temperature as a function of the micro state. One could claim that temperature is only defined for equilibrium and thus there is no need to define it as a function of the micro state. therefore, it is usually simply defined as the inverse of the change of the logarithm of the energy shell with energy (first law), neglecting all dynamical aspects altogether. But this definition, strictly speaking, excludes all investigations of situations in which temperature appears a function of time or space, because those are non equilibrium situations. Obviously it would at least be convenient to be able to express temperature as a function of the micro state, thus there have been several attempts.

As already explained in **3**, a quantity like temperature is essentially determined by two properties: It should take on the same value for two systems in energy exchanging contact, and if the energy of a system is changed without changing its volume, it should be a measure for the energy change per entropy change.

Most definitions rely on the second property. Maxwell connected the mean kinetic energy of a classical particle with temperature. In the canonical ensemble (Boltzmann distribution) it is guaranteed that the energy change per entropy change equals temperature. And the ensemble mean of the kinetic energy of a particle equals kT in this case. Thus, if ergodicity is assumed, i.e., if the time average equals the ensemble average, temperature may be defined as the time averaged kinetic energy. Similar approaches have been done on the basis of the microcanonical ensemble [43,44]. However, temperature is eventually not really given by an observable, but by a time average over an observable, leaving the question of the averaging time and thus the question, on what minimum timescale temperature may be defined open. Furthermore, the whole definition is entirely based on ergodicity. Nevertheless it allows, at least to some extend, for a investigation of processes in which temperature appears as varying in time and/or space, since the definition is not necessarily restricted to full equilibrium.

Due to those problems of standard temperature definitions, we want to present yet another, entirely quantum mechanical definition here.

7.1 Definition of spectral Temperature

We define:

$$\frac{1}{kT} := - \left(1 - \frac{W_0 + W_M}{2} \right)^{-1} \sum_{n=1}^M \left(\frac{W_n + W_{n-1}}{2} \right) \frac{\ln W_n - \ln W_{n-1} - (\ln N_n - \ln N_{n-1})}{E_n - E_{n-1}} \quad (7.1)$$

where W_n is the probability to find the system at the energy E_n , M is the number of the highest energy level, while the lowest one is labeled E_0 . This formula is motivated from the following idea: For a two energy level system, it seems plausible to define temperature just from the energy probability distribution and the degrees of degeneracy as

$$\frac{W_1 N_0}{W_0 N_1} = e^{-\frac{E_1 - E_0}{kT}}, \quad (7.2)$$

(7.1) now results, if one groups the energy levels of a multi-level system into neighboring pairs, to each of which via the above formula a temperature can be assigned, and weighs those temperatures with the probabilities for each pair to be occupied. This definition obviously depends only on the energy probability distribution and the spectrum of a system. It thus cannot change in time for an isolated system, and it is always defined, independent of whether or not the system is in an equilibrium state. Thus there will be many systems or situations for which this temperature is defined, but will not exhibit the thermodynamical properties, one expects. Those will, as we will explain in the following, only show in equilibrium situations or close to those.

If the spectrum of a system is very dense, and it is possible to describe the energy probability distribution $\{W_n\}$, as well as the degrees of degeneracy $\{N_n\}$, by smooth continuous functions ($W(E), N(E)$) with a well defined derivative, (7.1) can be approximated by

$$\frac{1}{kT} \approx - \int_0^{E_{max}} W(E) \left(\frac{d \ln W}{dE} - \frac{d \ln N}{dE} \right) dE, \quad (7.3)$$

which can be further simplified to yield

$$\frac{1}{kT} \approx \int_0^{E_{max}} W(E) \frac{d \ln N}{dE} dE + W(0) - W(E_{max}). \quad (7.4)$$

Since typically for larger systems, neither the lowest nor the highest energy level is occupied with considerable probability (if the spectra are finite at all), it is the first term on the right hand side of (7.4), that basically matters. This term can be interpreted as the average over the standard, system based, rather than micro state based, definitions of inverse temperature.

7.2 The Equality of spectral Temperatures in Equilibrium

If two systems are in heat conducting contact we expect their energy probability distributions to be the ones corresponding to the dominant region (see **5.5.2**), i.e., the ones defined by (5.64) and (5.65). We want to check now if, and under what circumstances, those dominant energy probability distributions yield the same temperature according to the definition (7.1) or (7.4).

First we examine the case of a small discrete system “g”, coupled to a large continuous system “c”, that is assumed to have a spectrum which is typical for large, modular systems, as described in **6.2**. For such a joint system a function $f(E^c)$

$$f(E^c) := N^c(E^c)N^g(E - E^g) \quad (7.5)$$

will always be peaked near $E^c \approx E$, since N^c grows, by definition, much faster with energy than N^g does. Thus, calculating $W^d(E^c)$ from (5.65) and assuming the situation assumed in **6.3**, namely that $\frac{W(E)}{N(E)}$ only takes on considerable values in some finite energy region, we find that $W^d(E^c)$ will also only take on considerable values within and slightly below the very same energy region. This result corresponds to intuition as it means that most of the total energy is in the larger system. Since, again by definition, the state density of the large system can be well described by an exponential, i.e.,

$$N^d(E^c) \propto e^{\beta E^c} \quad (7.6)$$

in the region where $W^d(E^c)$ is peaked, we find, applying (7.4), for the inverse temperature of the large system:

$$\frac{1}{kT^c} \approx \beta. \quad (7.7)$$

For the same situation we find for the small system, as explained in **6.3** (6.29)

$$W^d(E_n^g) \approx \frac{N^g(E_n^g)e^{-\beta E_n^g}}{\sum_n N^g(E_n^g)e^{-\beta E_n^g}} \quad (7.8)$$

Inserting this into (7.1) yields

$$\frac{1}{kT^g} \approx \beta. \quad (7.9)$$

From which we obviously can conclude

$$T^c \approx T^g. \quad (7.10)$$

To sum all this up, we can say that for a large, modular, continuous system in contact with a small, discrete system, and a reasonably peaked energy probability distribution on the combined system, we find, using our definition, the same local temperatures for

almost all states of the full system. This result is independent of whether or not the full system is in a pure or a mixed state, i.e., independent of whether there is a further surrounding or not. The temperatures are the same, although entropy is definitely not additive with respect to the chosen partition.

Now we want to examine the case of two large systems with continuous spectra in contact. In this case, as will be seen, we do not even need the assumption of the spectra being typical spectra of modular systems. Formulating (5.64) for a continuous spectrum yields:

$$W^d(E^g) = N^g(E^g) \int_0^\infty \frac{N^c(E - E^g)W(E)}{N(E)} dE \quad (7.11)$$

Applying (7.4) to this distribution yields:

$$\frac{1}{kT^g} = \int_0^{E_{max}} \int_0^\infty \frac{\frac{dN^g}{dE^g} N^c(E - E^g)W(E)}{N(E)} dE dE^g + W^d(0) - W^d(E_{max}) \quad (7.12)$$

The smallest energy value for which $N^c(E^c)$ takes on nonzero values at all, is $E^c = 0$. Thus we can, after reversing the order of integrations, replace E_{max} as a boundary by E . Furthermore we assume, that the probability densities to find the system in the ground state $W^d(0)$, or at infinitely highest possible energy (if there even is one) $W^d(E_{max})$ both vanish. We can then rewrite (7.12) as

$$\frac{1}{kT^g} = \int_0^\infty \frac{W(E)}{N(E)} \int_0^E \frac{dN^g}{dE^g} N^c(E - E^g) dE^g dE \quad (7.13)$$

If N^g either has no upper limit, or vanishes continuously, we can apply product integration to the inner integral, finding

$$\frac{1}{kT^g} = \int_0^\infty \frac{W(E)}{N(E)} \left(N^g(E)N^c(0) - N^g(0)N^c(E) - \int_0^E N^g(E^g) \frac{dN^c(E - E^g)}{dE^g} dE^g \right) dE. \quad (7.14)$$

Since state densities have to vanish at zero energy, we get $N^g(0) = N^c(0) = 0$. Substituting $E - E^g = E^c$ and reversing the boundaries of the integration yields

$$\frac{1}{kT^g} = \int_0^\infty \frac{W(E)}{N(E)} \int_0^E \frac{dN^c}{dE^c} N^g(E - E^c) dE^c dE \quad (7.15)$$

This form is exactly the one, one would have gotten if one had applied (7.4) to the container system. This may be seen from comparing it with (7.13), obviously only the subsystem indices are reversed.

If two large systems with continuous spectra are in heat conducting contact almost all micro states that the full system can be found in under the restriction of energy conservation of the full system, yield the same local spectral temperatures for the subsystems, regardless of whether the spectra are typical for modular systems, or how wide the energy probability distribution of the full system is.

7.3 Spectral Temperature as the Derivative of Energy with Respect to Entropy

As already explained, we do not only expect temperature to take on the same values for systems in contact, but also to be a measure for the energy change per entropy change if all other extensive variables are kept fixed, since this is basically what the first law states. Evidently there are situations in which temperature as defined by (7.1) will not show this behavior. If one, e.g., considers an isolated system controlled by a time dependent Hamiltonian, one finds that energy may very well change while entropy is strictly conserved. Nevertheless, one could compute a finite temperature for this system which would obviously not be in agreement with the temperature appearing in the first law. But this situation is probably not the one one has in mind, if trying to apply the first law. Here we want to distinguish two processes for which the first law should be applicable. Firstly, we investigate the process of bringing energy into an arbitrarily small system, by bringing it into contact with a, according to our definition, hotter environment, and secondly the case of slowly depositing energy in a large system by any kind of procedure (running current through it, stirring it, etc.).

In the first case we consider a discrete system whose entropy in equilibrium is given by

$$S = -k \sum_n W_n \ln W_n \quad (7.16)$$

where W_n is now the probability to find the system in some energy eigenstate, not at some energy level. The inner energy of the system is now given by

$$U = \sum_n W_n E_n. \quad (7.17)$$

As explained the energy probability distribution of the system in contact with a larger system reads (6.29)

$$W_n = \frac{e^{-\frac{E_n}{kT}}}{\sum_m e^{-\frac{E_m}{kT}}}, \quad (7.18)$$

where T is the temperature as well for the surrounding as for the system. If the surrounding gets hotter T increases, and S as well as U change. Thus we want to compute

$$\frac{\partial U}{\partial S} = \frac{\frac{\partial U}{\partial T}}{\frac{\partial S}{\partial T}}. \quad (7.19)$$

For the first derivative we get

$$\frac{\partial U}{\partial T} = \sum_n \frac{\partial W_n}{\partial T} E_n. \quad (7.20)$$

Computing the second yields

$$\frac{\partial S}{\partial T} = -k \sum_n \frac{\partial W_n}{\partial T} \ln W_n + \frac{\partial W_n}{\partial T}. \quad (7.21)$$

Since on the right hand side of (7.21) the sum can be exchanged with the derivative and $\sum_n W_n = 1$, the last term vanishes. We thus get

$$\frac{\partial S}{\partial T} = -k \sum_n \frac{\partial W_n}{\partial T} \left(-\frac{E_n}{kT} - \ln \sum_m e^{-\frac{E_m}{kT}} \right). \quad (7.22)$$

Since the second term in brackets does not carry the index n , the same argument as before applies and it vanishes. We find

$$\frac{\partial S}{\partial T} = \frac{1}{T} \sum_n \frac{\partial W_n}{\partial T} E_n \quad (7.23)$$

Inserting (7.20) and (7.23) into (7.19) eventually yields

$$\frac{\partial U}{\partial S} = T \quad (7.24)$$

which means that for this kind of process our temperature exhibits the desired behavior. Now we consider a large system in isolating contact with an environment into which energy is deposited by any kind of process. The inner energy of such a system reads:

$$U = \int W(E) E dE, \quad (7.25)$$

where now $W(E)$ is again the probability to find the system at some energy, not in a single energy eigenstate. The entropy of such a system in microcanonical equilibrium is (5.49)

$$S = -k \int W(E) \ln \frac{W(E)}{N(E)} dE. \quad (7.26)$$

Now we assume that the energy spread of the system is small enough, that the state density $N(E)$ is well described by some exponential, within the region where $W(E)$ takes on considerable values. As already explained this region can be fairly wide if the system is large. If this is the case we can replace

$$N(E) \approx N(U) e^{\beta(E-U)}. \quad (7.27)$$

Doing so we find

$$S \approx -k \int W(E) (\ln W(E) - \ln N(U) + \beta(E-U)) dE = k \ln N(U) - k \int W(E) \ln W(E) dE \quad (7.28)$$

or, if the state density $N(U)$ is very large, we can neglect the second term and then get

$$S \approx -k \ln N(U) \quad (7.29)$$

which means that in the above defined case the details of $W(E)$ do not matter, for the entropy it is almost entirely determined by the energy expectation value.

With the definition of inverse temperature (7.4) we get

$$\frac{1}{T} = \int W(E) \frac{d \ln N}{dE} dE \quad (7.30)$$

or, assuming the same case as above and approximating the state density in the relevant region,

$$\frac{1}{T} = \int W(E) \frac{d}{dE} \left(\ln N|_U + \frac{d \ln N}{dE}|_U (E - U) \right) = \frac{d \ln N}{dU}. \quad (7.31)$$

Thus in this case we obviously get

$$\frac{dS}{dU} = \frac{1}{T}, \quad (7.32)$$

which means that the temperature according to our definition exhibits a behavior, which is in agreement with the first law.

8 Pressure

8.1 On the Concept of adiabatic Processes

Technically, one could introduce pressure within classical statistical mechanics as a an observable, i.e., as a function of the micro state. The momentary change of the momenta of all particles that occurs due to the interaction with some wall has to equal the force exerted onto that wall and could thus be converted to pressure. And indeed, there are simple models of ideal gases that can account for some of there properties that proceed that way [45]. In general, however, this is not the way pressure is calculated within statistical mechanics. No ensemble average over such a “pressure observable” is taken. Instead one calculates energy U as a function of entropy S and volume V and takes the derivative of energy with respect to volume, while keeping entropy constant and identifies this quantity with negative pressure:

$$\left. \frac{\partial U}{\partial V} \right|_{S=\text{const.}} := -p, \quad (8.1)$$

which amounts to identifying force with the change of energy per change of length, which is natural, but the claim is, that the change appears in such a way, that entropy does not change. The inner energy of some system could principally change in many ways but it is assumed that a manner of change that keeps entropy constant is selected. Without this assumption the above definition (8.1) would be meaningless.

Since pressure is in this way defined by an infinitesimal step of an adiabatic process, first of all it has to be examined if, and under what conditions, adiabatic processes occur at all. In the case of temperature it was pretty obvious that processes exist during which entropy changes while the volume is kept constant, in this case, however, it is far from obvious that processes exist during which the volume changes while entropy remains constant.

Within quantum theory the most common approach to this issue is the adiabatic approximation [46, 47] (for classical theory see **2.5**). This is a method of solving the time dependent Schrödinger equation with a time dependent Hamiltonian. If a Hamiltonian contains a parameter a like length or volume, that varies in time, it will have the following form:

$$\hat{H}(a(t)) = |n, a(t)\rangle E_n(a(t)) \langle n, a(t)| = |n, t\rangle E_n(t) \langle n, t|. \quad (8.2)$$

At each time t a momentary Hamiltonian with a momentary set of eigenvectors and eigenvalues is defined. If the wavefunction is expanded in terms of this momentary basis

with an adequate phase factor, i.e., with the definition

$$\psi_n = \langle n|\psi\rangle \exp\left(\frac{1}{i\hbar} \int_0^t E_n dt'\right) \quad (8.3)$$

the time dependent Schrödinger equation can be transformed to the form:

$$\frac{\partial \psi_k}{\partial t} = - \sum_n \psi_n \langle k|\frac{\partial}{\partial t}n\rangle \exp\left(\frac{1}{i\hbar} \int_0^t (E_n - E_k) dt'\right) \quad (8.4)$$

The bracket term on the right hand side of (8.4) scales with the velocity of the parameter change, $\frac{da}{dt}$, this term gets multiplied by a rotating phase factor, that rotates the faster the larger the energy distance $E_n - E_k$ is. That means that if the initial state is a momentary eigenstate of the Hamiltonian $|\psi(0)\rangle = |n, 0\rangle$, the transition rate to the other eigenstates will be extremely small if the velocity of the parameter change is low, and it will fall off like $\frac{1}{E_n - E_k}$ for transitions to eigenstates that are energetically further away. Thus in this case we have as an approximate solution:

$$|\psi(t)\rangle \approx |n, t\rangle \quad (8.5)$$

For such an evolution entropy obviously is conserved. This is what is called the adiabatic approximation or the adiabatic following. This behavior has been substantially discussed for the situation described above, in which the initial state is a single momentary energy eigenstate and not a superposition of many of those. In this case the contributions to the wavefunction at some time t would consist of peaks in the spectrum, as described by the adiabatic approximation, centered around the energy of the energy eigenstate that generated them. If the edges of those peaks overlap, the corresponding amplitudes ψ_n have to be added coherently (according to (8.4)), thus producing possibly a higher probability density at the overlap region than the simple addition of the probability densities of the corresponding peaks would produce. This means that a coherent superposition of energy eigenstates is less likely to be correctly described by the adiabatic approximation than a mixture.

Thus, it is problematic to explain the existence of adiabatic processes in general by the scheme of the adiabatic approximation, if there is no decohering mechanism. If one tries to describe macroscopic thermodynamical systems as entirely isolated, non interacting systems, one would have to admit that these are most likely in superpositions of extremely many energy eigenstates, even if they are all degenerate, and, in this case, there is no decohering mechanism.

In the context of our approach the process of a changing local system parameter, like, e.g., volume, can be described by a Hamiltonian of the following form:

$$\hat{H}(t) = \hat{H}^g(a(t)) + \hat{H}^c + \hat{I}^{gc}(t) \quad (8.6)$$

To implement an adiabatic process, one still wants to have a thermally insulating contact with the environment. The full energy of the gas system, however, cannot be a strictly conserved quantity anymore, since without a changing energy one cannot get a finite

pressure. But the change of energy is only to be induced by the parameter a , thus, if a stopped changing at any time, energy should not change anymore. Demanding this behavior we get as a condition for the Interaction $\hat{I}^{gc}(t)$

$$\left[\hat{H}^g(t), \hat{I}^{gc}(t) \right] = 0 \quad (8.7)$$

As described in 5.4, the effect of suitable coupled environment system is to reduce purity within the gas system down to the limit set by the conserved quantities deriving from (8.7). As explained, this amounts to making the off diagonal elements of $\hat{\rho}^g$, represented in the basis of the, in this case momentary, eigenvectors of $\hat{H}^g(t)$, vanish. In order to get a qualitative understanding of the type of evolution that a Hamiltonian like the one defined in (8.6) will typically give rise to, we will sample this continuous evolution into two different types of (infinitesimal) time steps. In one type of step we imagine the interaction being turned off and the system developing according to its local, Hamiltonian $\hat{H}^g(t)$, this evolution being described by the corresponding von Neumann equation. During the other type of step, we imagine the interaction being turned on, but being constant in time, as well as the the local Hamiltonian. During this period the evolution is described by the Schrödinger equation for the full system, and will result in a wiping out of the momentary off diagonal elements. These two types of steps are now supposed to take turns with each other. In the limit of the steps being infinitesimally short, the true, continuous evolution results.

For the first type the Von Neumann equation reads

$$i\hbar \frac{\partial}{\partial t} \hat{\rho}^g = \left[\hat{H}^g(t), \hat{\rho}^g \right] \quad (8.8)$$

The probability W_n of the system to be found in a momentary eigenstate $|n(t)\rangle$ of $\hat{H}^g(t)$, is

$$W_n = \langle n(t) | \hat{\rho}^g(t) | n(t) \rangle. \quad (8.9)$$

If those probabilities do not change, the adiabatic approximation holds exactly true. Therefore we calculate there derivatives with respect to time finding

$$\frac{\partial}{\partial t} W_n = \left\langle \frac{\partial}{\partial t} n | \hat{\rho}^g | n \right\rangle + \left\langle n | \frac{\partial}{\partial t} \hat{\rho}^g | n \right\rangle + \left\langle n | \hat{\rho}^g | \frac{\partial}{\partial t} n \right\rangle. \quad (8.10)$$

Splitting up $\hat{\rho}^g$ in a diagonal part and an off diagonal part $\hat{\epsilon}$

$$\hat{\rho}^g =: \sum_n |n\rangle W_n \langle n| + \hat{\epsilon} \quad (8.11)$$

and inserting (8.8) and (8.11) into (8.10) yields:

$$\frac{\partial}{\partial t} W_n = W_n \left(\left\langle \frac{\partial}{\partial t} n | n \right\rangle + \left\langle n | \frac{\partial}{\partial t} n \right\rangle \right) + \left\langle \frac{\partial}{\partial t} n | \hat{\epsilon} | n \right\rangle + \left\langle n | \left[\hat{H}^g, \hat{\epsilon} \right] | n \right\rangle + \left\langle n | \hat{\epsilon} | \frac{\partial}{\partial t} n \right\rangle. \quad (8.12)$$

The first part on the ride hand side of (8.12) vanishes since

$$\left\langle \frac{\partial}{\partial t} n | n \right\rangle + \left\langle n | \frac{\partial}{\partial t} n \right\rangle = \frac{\partial}{\partial t} \langle n | n \rangle = 0. \quad (8.13)$$

Thus (8.12) reduces to

$$\frac{\partial}{\partial t} W_n = \left\langle \frac{\partial}{\partial t} n | \hat{\epsilon} | n \right\rangle + \langle n | [\hat{H}^g, \hat{\epsilon}] | n \rangle + \langle n | \hat{\epsilon} | \frac{\partial}{\partial t} n \rangle \quad (8.14)$$

Obviously, this derivative vanishes, if $\hat{\epsilon}$ vanishes. That means, that if, during the intermediate step, in which the interaction is active, the off diagonal elements get completely wiped out, the change rate of the probability vanishes, at the beginning of each step of the Von Neumann type. It will take on nonzero values during this step, especially if the step is long and $\hat{\rho}^g(t)$ changes quickly. If we make the steps shorter, the interaction with the environment might not erase the off diagonal elements completely. Thus, this situation is controlled by a sort of antagonism: A rapidly changing $\hat{\rho}^g(t)$ tends to make the adiabatic approximation fail, but a contact with the environment that reduces the off diagonal elements quickly stabilizes such a behavior.

This principle can also be found from a different consideration. Instead of solving the full Schrödinger equation one can introduce a term into the Von Neumann equation of the local system, that models the effect of the environment the way it is found to result in 5.4. Such an equation reads:

$$i\hbar \frac{\partial}{\partial t} \hat{\rho}^g = [\hat{H}^g(t), \hat{\rho}^g] - i\hbar \sum_{n,n'} |n\rangle \langle n'| \langle n | \hat{\rho}^g | n' \rangle C_{n,n'} \quad (8.15)$$

with

$$C_{n,n'} = 0 \quad \text{for } n = n', \quad C_{n,n'} \geq 0 \quad \text{for } n \neq n'. \quad (8.16)$$

This equation obviously leaves $\text{Tr}\{\hat{\rho}^g\}$ invariant and reduces off diagonal elements. The bigger the $C_{n,n'}$'s are, the quicker this reduction will proceed. To analyze this equation we define:

$$\rho^g(t) = \sum_n W_n |n(t)\rangle \langle n(t)| + r(t) \hat{\epsilon}(t) \quad \text{Tr}\{\hat{\epsilon}^2\} = 1 \quad r \geq 0 \quad (8.17)$$

where now the W_n 's are time independent and $\hat{\epsilon}$ may contain diagonal elements, thus the second term on the right hand side of (8.17) contains all deviations from the adiabatic behavior. With this definition r is a measure for those deviations. Inserting (8.17) into (8.15) and solving for $\frac{\partial r}{\partial t}$ yields (see **F**):

$$\frac{\partial r}{\partial t} = -r \sum_{n,n'} |\langle n | \hat{\epsilon} | n' \rangle|^2 C_{n,n'} - \sum_n W_n \left(\langle n | \hat{\epsilon} | \frac{\partial}{\partial t} n \rangle + \langle \frac{\partial}{\partial t} n | \hat{\epsilon} | n \rangle \right) \quad (8.18)$$

If the right hand side of (8.18) consisted only of the first sum, r could only decrease in time, the decrease would be the faster the bigger the $C_{n,n'}$'s were. Only the second sum can cause r to deviate from its stable value $r = 0$ and this sum is big if $\hat{H}^g(t)$ changes quickly. Thus, a fast local decoherence will stabilize the adiabatic approximation even for rapidly changing Hamiltonians.

To conclude we can say that within the context of our approach adiabatic, i.e., entropy

conserving processes are very likely to happen, if the decoherence induced by the environment proceeds fast, compared with the change velocity of the local Hamiltonian. The evolution of the state will then have the following form:

$$\hat{\rho}^g(t) \approx \sum_n W_n |n(a(t))\rangle \langle n(a(t))|. \quad (8.19)$$

If we let the parameter be the volume $a = V$, we eventually find for the pressure

$$\frac{\partial U}{\partial V} = \frac{\partial \text{Tr}\{\hat{\rho}^g \hat{H}^g\}}{\partial V} = \sum_n W_n \frac{\partial E_n}{\partial V} = -p. \quad (8.20)$$

This way, pressure, or any other conjugate variable (except for temperature), is always defined, whenever a local Hamiltonian \hat{H}^g can be defined, such that the weak coupling limit applies (see **9**) and the change of the system could proceed in such a way, that, with a changing local Hamiltonian, the whole system remained within the weak coupling limit. If this is guaranteed, pressure is defined by (8.20) regardless of whether the system is really changing or not and regardless of whether the system is really thermally insulated or not. the infinitesimal process step is just a virtual one, it is rather a mathematical construction to define pressure in such a way that it fits the whole scheme of thermodynamics, rather than something that is really physically happening.

8.2 The Equality of parametric Pressures in Equilibrium

To examine what pressures systems in “volume exchanging contact” will reach in equilibrium, we consider the often discussed case of a container with a moving wall inside, dividing the gas into two different compartments. This situation is just meant as an example, the principles of the consideration can be transferred to other situations. This situation in general, i.e., including the case of a non-heat conducting, thermally insulating wall, is difficult to access from the concept of dominant regions, in the quantum mechanical as well as in the classical case. This is due to the fact that it is difficult to define an accessible region in an holonomic way: In principle the gas systems can exchange energy while the wall is moving. On the other hand we know from experience that the system will not move further, once it has reached a state where the pressures on both sides are equal, although there might be further development if the wall was heat conducting. Thus in the case no further energy exchange is possible. Obviously there might be energetic inhibitions that prevent the system from venturing through its full principally accessible region. This situation results in the impossibility to predict the equilibrium state in general: This state depends on details like the weight of the wall, etc. If, e.g., the wall was so heavy that it would never be accelerated beyond the speed at which processes become non adiabatic, strictly speaking no equilibrium could ever be reached since entropy could not increase. Nevertheless, it should be possible to show, that the system will, on the level of local systems, rest in its state if the pressures on both sides are equal, and undergo some evolution if they are different.

To get a fully quantum mechanical description of this situation we now have to include

at least three systems: The gas in the left compartment (labeled “ l ”), the gas in the right compartment (labeled “ r ”) and the wall itself (labeled “ w ”). In this case we may think of the wall as just having one translational degree of freedom, the fact that the wall necessarily consists of many particles does not matter here. If we assume that this whole tri-partite system is energetically isolated and a weak coupling scheme may be applied, the following has to hold true:

$$U_l(t) + U_r(t) + U_w(t) = \text{const.} \quad (8.21)$$

where the U 's are the energy expectation values of the respective subsystems. If the wall is at first at rest, there will be a first period in which it moves slowly enough, if it starts moving at all, that for the two gas systems the adiabatic approximation will be valid. The heavier the wall is, the longer this period will last. For this period we thus get:

$$U_l(x(t)) + U_r(-x(t)) + U_w(t) = \text{const.} \quad (8.22)$$

where x is the position of the wall and positive x values stand for the compression of the right compartment. Taking the derivative of (8.22) with respect to time yields

$$\left(\frac{\partial U_l}{\partial x} \Big|_s - \frac{\partial U_r}{\partial x} \Big|_s \right) \frac{dx}{dt} + \frac{dU_w}{dt} = 0 \quad (8.23)$$

with the definition (8.1) we get

$$(p_l - p_r) \frac{dx}{dt} = \frac{dU_w}{dt} \quad (8.24)$$

If now the pressure in the left compartment is higher than the pressure in the right compartment the wall can move to the right while its energy is rising. This process is energetically allowed, and thus likely to happen, no locally stationary situation will result. If the pressures in both compartments are the same, the wall would have to start moving without picking up inner energy. This is obviously energetically forbidden. If the wall starts moving its inner energy has to increase, no matter how slow the movement is, especially if the wall is heavy. Thus, in this case the system will remain at rest.

This way, the equality of pressures in equilibrium may be established, although not on the basis of some dominant region, like in the case of temperature, simply because, in general, no accessible can be defined in a holonomic way.

9 Thermodynamical Limit

Since this approach is entirely based on the Schrödinger equation and no additional abstract concept, like ergodicity or the “a priori postulate”, has been involved, it is obvious that the thermodynamical behavior, that has been explained to result in the previous chapters, can only result in some situations or physical set-ups, for not all systems that can be described by the Schrödinger equation are thermodynamical systems. Many of the numerous restrictions and conditions that are necessary for thermodynamical behavior to emerge, have already been mentioned in the text, but for clarity we want to gather and discuss them here again explicitly. It should be pointed out, that, unlike the conditions in many other approaches, the conditions mentioned here are simply well defined physical properties of the systems in quest, rather than abstract claims or postulates.

Also unlike in other approaches, thermodynamics are here somehow considered to emerge bit by bit here, rather than systems being either fully thermodynamical or not at all. Since we defined “thermodynamicity” as a multitude of properties (see **3**), it could be that systems show some aspects of thermodynamics while lacking others. We indeed found this to be the case.

9.1 Weak coupling Limit

The first prerequisite for any system to exhibit thermodynamical behavior whatsoever in the context of this approach, is the existence of some other system, some surrounding, to which it is weakly coupled. Weakly in this context basically means

$$\langle \sqrt{\hat{I}^2} \rangle \ll \langle \hat{H}^g \rangle, \langle \hat{H}^c \rangle \quad (9.1)$$

i.e., the energies contained in the system itself and in the surrounding have to be much larger than the energy contained in the interaction. This has to hold true for any state, that the full system can possibly evolve into. This also means, that the spectrum of the joint system almost results as the convolution of the individual spectra of the subsystems, that the interaction can be considered as a perturbation that only has a weak effect on the spectrum (cf. **5.3**).

So on one hand, the interaction is indispensable but on the other it has to be very weak. What exactly the “optimum” interaction strength is, is not easy to tell and depends on the spectra of the decoupled subsystems. If those systems have truly and exactly degenerate energy levels, all the formulas about equilibrium states ((5.82),(5.83) etc.) apply even for arbitrarily small interaction strengths. Relaxation times might become extremely long,

but eventually equilibrium will be reached. If the subsystems have a high statedensity, or, more precisely, a small but finite levelspacing, it depends on the interaction strength how many energy levels “count” as degenerate, in the sense of the formulas describing the equilibrium states. If there is something like an average absolute value of the elements of the interaction matrix in some energy region, the number of levels within an energy interval of the size of this average value, can be considered the corresponding degree of degeneracy. To some extent a stronger interaction might even produce “better” results, since it yields higher effective degrees of degeneracy which enhances thermodynamical behavior (see **11.1.4**). If the interaction becomes too strong, the system might even still reach some local equilibrium state, but this equilibrium state may not be predictable anymore from the properties of the decoupled parts of the system. Except for this overall strengths and principal structures of the interaction matrix, like non-energy exchanging (microcanonical) or energy exchanging (canonical) interaction matrices, the details of the interaction matrix are irrelevant. The outcomes of all simulations show no dependence on such details (see, i.e., **Fig. 5.5**)

9.2 Microcanonical Equilibrium

For a system that cannot exchange energy with its surrounding, an equilibrium state of maximum entropy (a state that is diagonal in the local energy eigenbasis)(5.82) will be reached if

$$\sum_B \frac{(W_B^c)^2}{N_B^c} \ll \sum_A \frac{(W_A^g)^2}{N_A^g}, \quad (9.2)$$

where $W_A^g(W_B^c)$ are the occupation probabilities of the system(environment) and $N_A^g(N_B^c)$ the corresponding degree of degeneracy (cf. (5.48)). That means that an equilibrium will be reached, if the environment system either occupies energy levels with much higher degrees of degeneracy than the considered system, or if its energy probability distribution is much broader, i.e., contains more energy levels. The equilibrium state is not entirely independent of the initial state, but depends on the initial energy probability distribution of the considered system as can be expected under microcanonical conditions. Thus a behavior as claimed in **3.1 II a** results.

9.3 Energy-Exchange Equilibrium

If a system is in contact with its environment in such a way that energy exchange is allowed, an equilibrium energy probability distribution, $W^d(E_A^g)$ (see (5.64)), will be reached if

$$N(E_A^g)N(E - E_A^g) \ll \left(\sum_A N(E_A^g)N(E - E_A^g) \right)^2 \quad (9.3)$$

(cf. (5.78)) which means that either the statedensity of the considered system or the statedensity of the surrounding has to be high. An state of the maximum entropy that is

consistent with the equilibrium energy probability distribution (a state that is diagonal in the local energy eigenbasis), however, is only reached if additionally

$$\sum_B \frac{W^d(E_B^c)^2}{N_B^c} \ll \sum_A \frac{W^d(E_A^g)^2}{N_A^g} \quad (9.4)$$

(5.68), which is essentially the condition from the microcanonical equilibrium, except for here the equilibrium probabilities enter the formula, rather than the initial probabilities. A full equilibrium that only depends on the initial energy probability distribution of the full, joint system, $W(E)$, is thus only reached if the the surrounding is much larger.

9.4 Canonical Equilibrium

A full canonical equilibrium state including complete independence of the initial state of the considered system, maximum entropy consistent with the mean energy, i.e., standard Boltzmann distribution is reached, if in addition to the requirements of the paragraph above, the environment system has a state density that is well described by some exponential

$$N(E^c) \propto e^{\beta E} \quad (9.5)$$

in the region where the function $\frac{W(E)}{N(E)}$ takes on non-negligible values (see **6.3**). This is very likely to be the case if the environment system consists of very many, identical, weakly interacting subsystems **6.2**. The more subsystems are contained in this modular environment, the larger will be the ranges in which the state density is well described by an exponential. Thus, in this case a behavior like it is claimed in **3.1 II b** results.

9.5 Spectral Temperature

It is possible to define a parameter, called the spectral temperature **7.1**, that is defined on the basis of the spectrum and momentary energy probability distribution of a system, that has the following properties:

If two systems reach the equilibrium distribution $W^d(E_A^g), W^d(E_B^c)$ they have the same spectral temperature if either one system has a state density that is typical for large modular systems, like described in **6.2** and one (small) system has a discrete spectrum, or if both systems are large and have spectra that may be described by continuous state densities **7.2**. This is the property of temperature demanded in **3.1 V**.

This equality of spectral temperatures is entirely independent of the question whether of the entropies of the two systems in quest are extensive, (additive) or not. The above principle applies for the case of the joint, full system being in a pure state, thus having the entropy 0. Since the local entropies of the two systems in equilibrium will be bigger than zero, entropies will definitely not be additive.

This spectral temperature has yet another feature. If a small system is first coupled to some large modular environment with some specific parameter β , like described in

9.4, and than coupled to some other (‘ hotter’) large modular environment with an infinitesimally smaller specific parameter, the amount of energy change divided by the amount of entropy change in this process is exactly given by the spectral temperature. If the energy probability distribution of a large modular system only takes on non-negligible values in a region where its state density is described by an exponential, and then energy is brought into the system (‘heating’) such that the ‘peak’ of the energy probability distribution shifts to another region, described by another exponential, the energy change divided by the entropy change in this process will be approximately the spectral temperature, the approximation will be the better the higher the state density in all regions is **7.3**. This is the property of temperature demanded in **3.1 III b**.

9.6 Parametric Pressure

If for a system a changing local Hamiltonian $\hat{H}^g(V)$ is defined whose change may be described by an continuous parameter V , a parametric pressure (8.20) can be defined on the basis of this Hamiltonian and the momentary energy probability distribution of this system. This parametric pressure has the following properties:

If the parameter is changed by a small (infinitesimal) amount, and this change is performed slowly compared to the rate at which off diagonal elements of the density matrix of the system vanish due to the weak coupling of the system to a larger environment (see **9.2**), the change of the inner energy divided by the change of the parameter is given by the momentary parametric pressure of the system **8.1**. This is the property of a conjugate variable like pressure which is demanded in **3.1 III c**

If, furthermore, two such systems are in contact in such a way that the parameter of the Hamiltonian in one system can only increase if the other decreases, and both systems are weakly microcanonically coupled to some larger environment, the whole system can only be at rest if both parametric pressures are equal **8.2**. This is the property of an intensive variable like pressure as demanded in **3.1 V**.

9.7 Extensivity of Entropy

If a system in full canonical contact with larger environment is being split up into two weakly interacting subsystems, their entropies will be additive in equilibrium. If, however, a system is in microcanonical contact with some larger environment, such that the requirements of **9.2** are met, the entropies of the subsystems will not necessarily be additive. This will in general only be the case if both systems feature a state density that is typical for large modular systems **6.2**. Thus, in this case entropy is an extensive quantity as claimed in **3.1 IV**. So if all systems that one deals with are in themselves systems that are large in that sense that they consist of many identical sub units, and if all those systems are at least microcanonically coupled to some even larger environment, the resulting situation will show all properties and features of standard equilibrium thermodynamics.

10 Quantum mechanical and classical State Densities

Regardless of its possibly problematic foundation (see **2**), Boltzmann’s “recipe” to calculate thermodynamical behavior from a Hamiltonfunction of a system works extremely well. This recipe essentially consists of his classical entropy definition, the first and second law. Using this recipe, not only the thermodynamical behavior of gasses, but also thermodynamical properties of much more complicated systems, like liquid crystals, polymers, etc., which are definitely quantum mechanical systems, may be computed to very good precision.

If now, like in this approach, another (fully quantum mechanical) entropy definition is suggested, the question arises whether this other definition produces equally good, or even better results. Since the best results are for the biggest part the same results, it has to be checked whether or not

$$S^c \approx S^q \quad \text{with} \quad S^c = k \ln G^c(U, V) \quad S^q = k \ln G^q(U, V) \quad (10.1)$$

and therefore

$$G^c(U, V) \approx G^q(U, V) \quad (10.2)$$

where G^c is, according to Boltzmann, the number of classical micro states that is consistent with the macro state specified by U, V , or, stated more mathematically: the volume of the region in Γ -space that contains all micro states of the system, which feature the energy U and are restricted to the (configuration space) volume V . This region is also referred to as the energy shell. G^q is the quantum mechanical density of energy eigenstates at the energy U , given that the whole system is contained within the volume V . With this definition S^q is the equilibrium entropy, we found for the case of microcanonical conditions and sharp energies (5.50). If the validity of (10.2) cannot be established, a theory relying on S^q remains highly problematic, from a practical point of view, regardless of its theoretical plausibility.

From an operational point of view, the validity of (10.2) is far from obvious, because both quantities are evidently computed in entirely different ways. And, of course, in general, G^q is discrete, while G^c is a smooth continuous function. There are indeed cases where the recipe based on G^q works better than the one, based on G^c . If one, e.g., changes from an ideal to a molecular gas, the deficiencies of the Boltzmann entropy definition, S^c , become visible at low temperatures. The behavior, e.g., the heat capacity deviates significantly from the predicted behavior. This effect is referred to as the “freezing out of

inner degrees of freedom". It is due to the fact that the quantum mechanical levelspacing of the spectrum arising from the inner (vibrational, rotational) degrees of freedom is much larger than that arising from the translational degrees of freedom. This behavior is described correctly from calculations based on S^q . Nevertheless, if one claims S^q to be the "correct" definition, the striking success of the Boltzmann entropy definition needs explanation. This can only be done by showing the validity of (10.2) for a reasonably large class of cases.

For some simple systems, for which both types of spectra can be calculated exactly, there is a striking similarity between $G^c(U, V)$ and $G^q(U, V)$: For a free particle they are the same. If the free particle is restricted to some volume $G^q(U, V)$ becomes discrete, but as long as V is large, the levelspacing is small and if the energy interval is chosen to contain many levels - it may still be extremely small compared to macroscopic energies - $G^c(U, V)$ and $G^q(U, V)$ are still almost identical. This eventually is the reason why both methods lead to almost identical thermodynamical state functions for an ideal gas. A very similar situation is found for the harmonic oscillator. The quantum energy spectrum G^q of the harmonic oscillator consists of an infinite number of equidistant energylevels. The volume of the classical energy shell, G^c , of an harmonic oscillator is constant with respect to energy. Thus, if the levelspacing is small, like it is the case for small frequencies, the number of levels that is contained within a given interval is almost independent of the energy U around which the interval is centered.

Thus, in the following we want to analyse whether G^c and G^q can be viewed as approximations for each other, at least for a large set of cases.

10.1 Similarity of classical and quantum mechanical State Densities

10.1.1 Sommerfeld Quantization

One hint in that direction comes from the Sommerfeld quantization [39]. This theory from the early days of quantum mechanics states that energy eigenstates correspond to closed trajectories in classical phase space, that enclose areas of the size $n\hbar$, n being integer. If this theory was right, the connection would have been established and the quantum mechanical spectrum could be calculated from G^c by

$$\int_0^{U_n} G^c(E) dE = n\hbar \quad (10.3)$$

with U_n being the energylevels. Unfortunately the Sommerfeld theory is not precisely right, and the above formula holds only exactly true for some special cases.

10.1.2 Partition Function Approach

Some more evidence for the similarity of G^c and G^q for the case of modular systems (see 6.2) can be obtained from a consideration which is usually done in the context

of the partition function [48]. The partition functions which, within standard classical mechanics, determine the thermodynamical properties of a system completely, read for the classical and the quantum mechanical case:

$$Z^q = \text{Tr} \left\{ e^{-\frac{\hat{H}}{kT}} \right\} = \int e^{-\frac{E}{kT}} G^q(E) dE, \quad Z^c = \int e^{-\frac{H(\vec{p}, \vec{x})}{kT}} dx dp = \int e^{-\frac{E}{kT}} G^c(E) dE \quad (10.4)$$

If one sets $\frac{1}{kT} = \alpha$, the partition function becomes equal to the function $A(\alpha)$ which, as described in **6.2**, is crucial for the spectrum of large, modular systems. In the literature one finds:

$$Z^q(\alpha) = A(\alpha) \approx Z^c(\alpha) - \frac{\hbar \alpha^3}{24m} \int e^{-\alpha H(\vec{p}, \vec{x})} \sum_i \left(\frac{\partial H}{\partial x_i} \right)^2 dx dp \quad (10.5)$$

where the correction term is basically the leading order term of an expansion in terms of powers of \hbar , but higher order terms will also involve higher orders of α , the gradient of the Hamiltonian, and inverse mass.

If $Z^q(\alpha)$ and $Z^c(\alpha)$ were exactly the same, G^c and G^q would have to be equal as well, since by taking derivatives with respect to α of $Z(\alpha)$ all momenta of $e^{-\alpha E} G(E)$ can be produced and if all momenta of two functions are the same, the two functions have to be the same. This, however, cannot be the case since one knows that G^q is discrete while G^c is not. But this difference may only show for very high momenta corresponding to very high order derivatives. Fortunately, as explained in **6.2** and stated in equation (6.20). For the spectra of compound systems, consisting of many identical, weakly interacting units, only $A(\alpha)$ and the first two momenta of $e^{-\alpha E} G(E)$ play a crucial role. That means, that for systems consisting of many weakly interacting subsystems for each of which the correction term in (10.5) is small, G^c and G^q will indeed be the same in that sense that the number of states, described by G^q , within a very small energy interval, will be proportional to the value of G^c at the corresponding energy.

This consideration, however, does not provide a necessary criterion for the equality of G^c and G^q . If one thinks, e.g., of a wide potential well, with a saw tooth shaped bottom, it is obvious that if one makes the tooth smaller and smaller, the spectrum should approach the spectrum of a flat bottom potential well, for which the correction term in (10.5) is zero. The correction for the saw tooth bottom potential well does not decrease with tooth getting smaller, it might indeed be arbitrarily big if the edges of the teeth are arbitrarily steep. I.e., there are systems for which the expansion in (10.5) does not even converge, nevertheless G^c and G^q of those systems may be very similar.

10.1.3 Minimum Uncertainty Wavepackage Approach

Due to the various insufficiencies of the above analysis of the relation between G^c and G^q , we want to present here yet another consideration which might help to clarify this question. The basic idea of this consideration is the following. Rather than analyzing the spectrum of the Hamiltonian directly, one can analyze the spectrum of a totally mixed state ($\hat{1}$ -state) subject to this Hamiltonian. Since a system in the totally mixed state occupies every state with the same probability, it can be found in a certain energy interval

with a probability proportional to the number of energy eigenstates within this interval. If the $\hat{1}$ -state is given as an incoherent mixture of many contributions, its spectrum will result as the sum of the individual spectra of the contributions. Here, the $\hat{1}$ -state will be given as a mixture of minimum momentum-position uncertainty wavepackages, thus each of them corresponds to a point in classical Γ -space. If it is then possible to show that to $G^q(U)$ only those wavepackages contribute, that correspond to points in Γ -space that feature the energy U , i.e., if the energy spread of those packages is small, a connection between G^c and G^q can be established.

We start off by writing the quantum mechanical state density as the trace of a special operator,

$$\text{Tr} \left\{ \hat{\Pi}(E) \right\} = G^q(E), \quad \hat{H} =: \sum_E E \hat{\Pi}(E) \quad (10.6)$$

where $\hat{\Pi}(E)$ is a projector, projecting out the energy eigenspace with energy E , that some Hamiltonian \hat{H} gives rise to. From this we find

$$\begin{aligned} \text{Tr} \left\{ \hat{\Pi}(E) \right\} &= \text{Tr} \left\{ \hat{\Pi}(E) \sum_{\gamma} |\gamma\rangle\langle\gamma| \right\}, \quad \text{this being true if:} \\ &\sum_{\gamma} |\gamma\rangle\langle\gamma| = \hat{1} \quad \text{even if} \quad \langle\gamma|\gamma'\rangle \neq 0. \end{aligned} \quad (10.7)$$

Carrying out the trace operation yields

$$\text{Tr} \left\{ \hat{\Pi}(E) \sum_{\gamma} |\gamma\rangle\langle\gamma| \right\} = \sum_{\gamma} \langle\gamma|\hat{\Pi}(E)|\gamma\rangle \quad (10.8)$$

defining

$$g(\gamma, E) := \langle\gamma|\hat{\Pi}(E)|\gamma\rangle \quad (10.9)$$

we get

$$G^q(E) = \sum_{\gamma} g(\gamma, E). \quad (10.10)$$

Since $g(\gamma, E)$ is the energy spectrum of a single contribution $|\gamma\rangle$ to the $\hat{1}$ -operator, we now have established that the full spectrum $G^q(E)$ results as a sum of the individual spectra of the contributions $g(\gamma, E)$.

Now we introduce our special representation of the $\hat{1}$ -operator and show that it has indeed the required properties. Its contributions in position representation read:

$$\langle x|\gamma\rangle := \langle x|q, p\rangle := \frac{1}{(2\pi)^{\frac{3}{4}} \sqrt{\hbar\Delta x}} e^{\frac{-(x-q)^2}{4\Delta^2 x} - i\frac{p}{\hbar}x} \sqrt{dqdp} \quad (10.11)$$

Obviously those are infinitesimal Gaussian (minimum position-momentum uncertainty) wavepackages, each of them corresponding to a point $\gamma = (q, p)$ in phase space. With this definition we get for the whole $\hat{1}$ -operator in position representation

$$\langle x| \int_{\gamma} |\gamma\rangle\langle\gamma| d\gamma |x'\rangle = \frac{1}{(2\pi)^{\frac{3}{2}} \hbar\Delta x} \int e^{\frac{-(x-q)^2 - (x'-q)^2}{4\Delta^2 x} - i\frac{p}{\hbar}(x-x')} dqdp, \quad (10.12)$$

where we changed from a sum to an integral since the contributions are infinitesimal. The integration over p yields a δ -function in $(x - x')$ that can be pulled out of the q integration. Since the δ -function is zero everywhere except for $x = x'$, therefore we can set $x = x'$ in the integrand and do the integration over the remaining Gaussian, finding

$$\langle x | \int_{\gamma} |\gamma\rangle \langle \gamma| d\gamma |x'\rangle = \delta(x - x') = \langle x|x'\rangle = \langle x|\hat{1}|x'\rangle, \quad (10.13)$$

which proves that the operator, we introduced as a $\hat{1}$ -operator, is indeed truly a $\hat{1}$ -operator. Now we want to analyze the “weight” of a single $|\gamma\rangle$ wavepackage to the full spectrum. Since such a wavepackage might bring contributions to many energy levels we have to sum over all energies:

$$\sum_E g(\gamma, E) = \sum_E \langle \gamma | \hat{\Pi}(E) | \gamma \rangle = \langle \gamma | \gamma \rangle \quad \text{since} \quad \sum_E \hat{\Pi}(E) = \hat{1}. \quad (10.14)$$

To evaluate $\langle \gamma | \gamma \rangle$ we use definition (10.11) finding

$$\langle \gamma | \gamma \rangle = \frac{1}{(2\pi)^{\frac{3}{2}} \hbar \Delta x} \int e^{-\frac{(x-q)^2}{2\Delta^2 x}} dx dq dp = \frac{dq dp}{2\pi \hbar} = \frac{dq dp}{h} \quad (10.15)$$

This means that while the spectrum is built up from contributions coming from γ -wavepackages, the contributions from wavepackages corresponding to a phase space volume of $dq dp = h$, will add the weight 1 to the spectrum. If this weight was now concentrated in one single energy level, one could say that the Sommerfeld quantization was exactly right and that an volume of h would correspond to exactly one energy eigenstate. This, however, is typically not the case. Nevertheless, (10.15) means that, if the contributions of wavepackages corresponding to classical states up to a certain energy, fall entirely into some energy range of the quantum mechanical spectrum, and no other wavepackages lead to contributions within this range, the total number of states within this range has to equal the classical phase space volume up to the corresponding energy divided by h . And this situation might very well be to pretty good accuracy the case, as we will see. This is the reason why in some simple cases like the particle in a box or the harmonic oscillator this simple relation holds exactly true.

Generalizing this scheme to N particles, or rather $3N$ degrees of freedom, is straight forward. $|\gamma\rangle$ has then to be chosen as a product of the above described wavepackages, yielding that the contributions corresponding to a Γ -space region of volume h^{3N} , will add the weight one to the spectrum. (Note that at this point this is just a generalization to a larger number of degrees of freedom, so far there is no modularity involved). Thus, in this case, dividing the classical Γ -space volume by h^{3N} , can be a reasonable approximation.

To decide now, whether or not this will indeed be a reasonable approximation, we analyze (10.10) more thoroughly. Since the exact form of $g(\gamma, E)$ can only be calculated with huge effort, we cannot evaluate this equation exactly. But some properties of $g(\gamma, E)$ can be estimated with pretty good precision. E.g., the mean value reads

$$\overline{g(\gamma)} := \frac{\sum_E E g(\gamma, E)}{\sum_E g(\gamma, E)} = \frac{\langle \gamma | \hat{H} | \gamma \rangle}{\langle \gamma | \gamma \rangle} \approx \frac{p^2 + \frac{\hbar}{\Delta x}^2}{2m} + V(q), \quad (10.16)$$

where V is the potential energy part of the Hamiltonian and the corresponding term an approximation, which is valid if Δx is small compared to the curvature of the potential. Also the energy spread, the variance of $g(\gamma, E)$ can be roughly estimated:

$$\Delta g(\gamma) := \frac{\sqrt{\langle \gamma | \hat{H}^2 | \gamma \rangle - \langle \gamma | \hat{H} | \gamma \rangle^2}}{\langle \gamma | \gamma \rangle} \approx \frac{p\hbar}{m\Delta x} + \frac{\partial H}{\partial x}|_{q,p} \Delta x. \quad (10.17)$$

We thus find that the mean of $g(\gamma, E)$ will be pretty near to the energy of the corresponding classical state, and that its width, at least for systems with large mass and small potential gradients, will be reasonably small. Or, stated in other words, a minimum momentum-position uncertainty wavepackage has a spectrum concentrated around the energy of its corresponding classical state. If we now assume that the spectra of those packages have a form, in which almost all of their weight is concentrated within a range of some variances $\Delta g(\gamma)$ from the mean, like for Gaussian or Poissonians, we can state that

$$\sum_{E=U-a\Delta g(\gamma)}^{E=U+a\Delta g(\gamma)} g(\gamma, E) \approx \frac{dqdp}{h} \quad (10.18)$$

where a is some small integer and U the classical energy $H(q, p)$. With this result we can now go back to (10.10) and write

$$\sum_{E=U-ab\Delta g(U)}^{E=U+ab\Delta g(U)} G^q(E) = \sum_{E=U-ab\Delta g(U)}^{E=U+ab\Delta g(U)} \sum_{\gamma} g(\gamma, E) \approx \int_{\gamma, H(\gamma)=U-ab\Delta g(U)}^{\gamma, H(\gamma)=U+ab\Delta g(U)} \frac{dqdp}{h}, \quad (10.19)$$

where b is meant to be some integer number and $\Delta g(U)$ is now supposed to be some $\Delta g(\gamma)$, typical for the corresponding energy region. The idea is that, if b is large enough, almost all $g(\gamma, E)$ from this energy region, fall entirely into that part of the spectrum that is summed over and thus to those, (10.18) applies. Within an interval of $a\Delta g(U)$ at the edges of the energy region, (10.18) can obviously not be applied, this is why (10.19) is an approximation. But if the gradient of the classical Hamiltonfunction does not vary much within the energy region, the relative error, ϵ , that might occur due to this approximation, can be estimated as the relation of the energy interval for which the $\Delta g(\gamma)$ fall entirely into the summed over part of the spectrum, to the energy interval for which they do not:

$$\epsilon \approx \frac{a\Delta g(U)}{ab\Delta g(U)} = \frac{1}{b}. \quad (10.20)$$

Keeping this in mind we evaluate the right hand side of (10.19):

$$\int_{\gamma, H(\gamma)=U-ab\Delta g(U)}^{\gamma, H(\gamma)=U+ab\Delta g(U)} \frac{dqdp}{h} = \frac{1}{h} \int_{U-ab\Delta g(U)}^{U+ab\Delta g(U)} G^c(E) dE. \quad (10.21)$$

Thus we eventually find:

$$\sum_{E=U-ab\Delta g(U)}^{E=U+ab\Delta g(U)} G^q(E) \approx \frac{1}{h} \int_{U-ab\Delta g(U)}^{U+ab\Delta g(U)} G^c(E) dE \quad (10.22)$$

with a relative error given by (10.20). So we find that the number of states within a certain energy region is equal to the phase space volume corresponding to that energy region, divided by \hbar . The energy region that has to be summed over is proportional to $b\Delta g(U)$. The relative error occurring is inversely proportional to b . Thus, this method is definitely not adequate to identify individual energy eigenstates with precision. But how many sates are within an interval of, say 100 times the energy spread of a minimum uncertainty wavepackage, can be predicted with pretty good accuracy.

Obviously this method is useful for systems for which the energy spread (10.17) of a minimum uncertainty wavepackage is much less than its mean energy (G.10). Those systems are obviously all spatially large systems with big masses. If the system of the earth revolving around the sun is considered quantum mechanically, within any energy interval appropriate to that scale, should be an enormous amount of energy eigenstates. Since the typical energy spread of a minimum uncertainty wavepackage for this system should still be extremely small, compared to this macroscopically small energy interval, the number of quantum sates within this interval will be given with extremely good relative precision by the size of the corresponding classical energy shell, divided by h^3 . Since the sun-earth system can be downscaled to a hydrogen atom, without changing the principal form of the spectrum, the result even applies to the spectrum of a hydrogen atom, at least near the ionization energy. Something very similar is true for the particle in a box.

But the more interesting class of systems, from a thermodynamical point of view, for which this method should yield good results is the class of systems consisting of many identical, almost interaction free subsystems. If they are assumed to be truly interaction free, the mean energy of a minimum uncertainty wavepackage of the full system is given by a sum over the mean values of the individual systems:

$$\overline{g_{tot}(\gamma_{tot})} = \sum_n^N \overline{g_n(\gamma_n)}. \quad (10.23)$$

The energy spread of the full system in this case reads

$$\Delta g_{tot}(\gamma_{tot}) = \sqrt{\sum_n^N \Delta^2 g_n(\gamma_n)}. \quad (10.24)$$

If we now assume that the individual mean values and the individual energy spreads of the subsystems are of comparable size for most γ_{tot} , we can roughly estimate the relation as

$$\frac{\Delta g_{tot}(\gamma_{tot})}{\overline{g_{tot}(\gamma_{tot})}} \approx \frac{1}{\sqrt{N}}, \quad (10.25)$$

which means that the described method should produce good results for energy intervals of U/\sqrt{N} , where U is the total energy of the full system.

Eventually we may conclude that there is a wide range of systems for which $G^q(U)$ is indeed proportional to $G^c(U)$ including large systems consisting of heavy masses and systems consisting of very many almost interaction free subsystems. Since thermodynamical systems typically belong to the latter class, it will be very difficult to decide whether a classical or a quantum theory of thermodynamics holds true from thermodynamical experiments, since the predictions for those will be very similar from both theories.

11 Ways to Equilibrium

Almost everything stated so far referred to situations of fully developed equilibrium. Means to predict the value that some variable, e.g. the probability of a system to be found at some energy level, will take on after some relaxation time, have been described. Nothing has been said so far about how long such a relaxation time will be, or what behavior can be expected during that relaxation. But this is an important aspect as well.

11.1 Theories of Relaxation Behavior

11.1.1 Fermi's golden Rule for external Perturbations

The most common relaxation scenario within quantum mechanics is an exponential decay. Such a exponential relaxation behavior can be obtained from various considerations. The simplest one is the one based on "Fermi's golden rule". In this approach transition probabilities are calculated in the following way: It is assumed that a system is in an energy eigenstate of its Hamiltonian. Then, an external perturbation is described by adding a time dependent part to the Hamiltonian, for a certain time. If the strength of this perturbation and the time during which it occurs fulfill some conditions, the Schrödinger equation can be solved for that period, yielding a final state with nonzero amplitudes for energy eigenstates of the systems own Hamiltonian, which have been zero before. From these amplitudes one than calculates transition probabilities, interpreted as the probability of the system to switch from some energy eigenstate to another in some given time. If many such transition processes are considered a rate equation yielding an exponential decay results. This scheme would be perfectly right if there was some decohering mechanism transforming the coherent wavefunctions into incoherent mixtures in between the perturbation (transition) periods. But without such a decohering mechanism, the interpretation of the wavefunction after a perturbation as a mere statistical mixture of energy eigenstates is wrong. Thus the true, longtime solution of the Schrödinger equation may lead to behavior that differs significantly from the one described by the solution of the rate equation.

11.1.2 Fermi's golden Rule for a coupled Environment

To account for spontaneous emission the above scheme is often modified by replacing the external time dependent perturbation by coupling the system weakly to the system of the electromagnetic field [14]. (This already appears very similar to the situations we discussed as thermodynamical.) But since in this picture, again, the Schrödinger equation can analytically only be solved for a short time step, one has to rely on the above described transition rate idea, with all its problems. Nevertheless, in the context presented in the work at hand, we can interpret the situation as a system in heat conducting contact with a much larger environment with temperature 0. The prediction that the system will finally have temperature 0, i.e., will end up in its ground state, is in accordance with the decay scheme.

11.1.3 Weisskopf-Wigner Theory

An approach that is based on a (approximate) continuous solution of the Schrödinger equation for long times, rather than on a transition probability scheme, is the Weisskopf-Wigner theory [49]. The situation analyzed in this theory is pretty much the same as above. Whether or not an approximate solution can be found depends strongly on a non-degenerate ground state of the environment and an, in the considered limit, infinitely high state density, combined with an infinitesimally weak coupling. Those conditions are met by the example of an atom coupled to the electromagnetic field in open space, but make it difficult to apply the method to other situations.

11.1.4 Large Environment Approach

Here we want to present yet another method to analyze the decay process that yields, under specific conditions, an exponential relaxation. This method, like the transition rate approach, relies on a stepwise solution of the Schrödinger equation, but it takes superpositions into account. The crucial approximating here is the replacing of some specific quantities by their Hilbertspace averages. But this approximation is backed up by the prove that the corresponding Hilbertspace landscape is almost flat, thus for almost all states the Hilbertspace average will be a good approximation.

The energy scheme of the situation we are going to analyze is depicted in **Fig. 11.1 A**

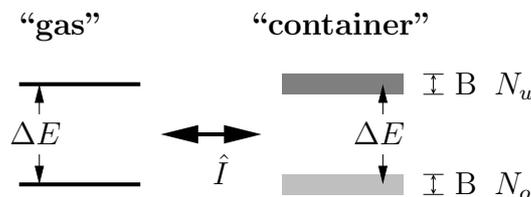


Fig. 11.1: Discrete two-level system coupled canonically to a quasi-continuous container system. This setup should, for a high statedensity in the container system, and an adequately tuned coupling, exhibit an exponential decay of an excitation in the container system.

two level system is in contact with an “many level” environment. Only the relevant parts of the spectrum of the environment enter the model. Those are in this case two “bands” of width B , containing $N_o(N_u)$ equidistant eigenstates in the lower(upper) band. Since we want to consider an evolution from an initial state, with the system in the excited state and the environment in the lower band, due to overall energy conservation the only other set of states that the full system can evolve into, is the set with the considered system in the ground state, and the environment in its upper band. Thus the Hamiltonian within the relevant subspace of the entire Hilbertspace may be visualized as follows:

$$\hat{H} = \left(\begin{array}{cc|cc} \cdots & & 0 & \\ & n\Delta E_o & & \hat{I} \\ \hline 0 & & \cdots & \\ & \hat{I} & & m\Delta E_u \\ \hline & & 0 & \cdots \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} |\psi_o\rangle \\ \\ \\ |\psi_u\rangle \end{array} \quad (11.1)$$

The Hamiltonian is displayed in the eigenbasis of the uncoupled system, for simplicity we assume for the moment that the coupling \hat{I} only adds terms to the Hamiltonian in the off diagonal blocks. In the following we are going to call that part of the wavevector that corresponds to the considered system being in the excited state $|\psi_o\rangle$ and the part that corresponds to the system being in the ground state $|\psi_u\rangle$, thus,

$$|\psi\rangle = |\psi_o\rangle + |\psi_u\rangle \quad (11.2)$$

To analyze this model we first of all switch to the Dirac picture, like it is routinely done by the following transformations:

$$\hat{S} := e^{-\frac{i}{\hbar}\hat{H}_0 t} \quad |\psi_D\rangle := \hat{S}^\dagger |\psi\rangle \quad \hat{I}_D := \hat{S}^\dagger \hat{I} \hat{S}, \quad (11.3)$$

Where \hat{H}_0 is the Hamiltonian of the uncoupled system. Those transformations lead to

$$i\hbar \frac{\partial}{\partial t} |\psi_D\rangle = \hat{I}_D |\psi_D\rangle \quad (11.4)$$

which is also known as the interaction picture.

The crucial quantities in this context are the probabilities to find the system in its excited(ground) state, $W_o(W_u)$. Due to the diagonality of \hat{H}_0 those quantities have the same representation in the Dirac and in the Schrödinger picture:

$$W_o = \langle \psi_{Do} | \psi_{Do} \rangle = \langle \psi_o | \psi_o \rangle \quad W_u = \langle \psi_{Du} | \psi_{Du} \rangle = \langle \psi_u | \psi_u \rangle \quad (11.5)$$

Thus, pretty much everything can be calculated directly within the Dirac picture. For simplicity we omit in the following the Dirac subscript D , but everything will refer to the Dirac representation. To approximate the evolution of the system for a short time step, we can truncate the corresponding Dyson series:

$$|\psi(\tau)\rangle \approx \left(1 - \frac{i}{\hbar} \hat{U}_1(\tau) - \frac{1}{\hbar^2} \hat{U}_2(\tau) \right) |\psi(0)\rangle \quad (11.6)$$

This is a truncation at second order, in which the \hat{U} 's are the time ordered integrals that occur in the Dyson series []:

$$\hat{U}_1(\tau) = \int_0^\tau \hat{I}(\tau') d\tau' \quad \hat{U}_2(\tau) = \int_0^\tau \hat{I}(\tau') \int_0^{\tau'} \hat{I}(\tau'') d\tau'' d\tau' \quad (11.7)$$

(To further simplify notation we do not write the τ dependence of the \hat{U} 's explicitly any longer. Furthermore we omit the time dependence of the wavefunction, if it refers to the initial state, i.e., $|\psi(0)\rangle := |\psi\rangle$). Using (11.6) we find for the probabilities

$$\begin{aligned} \langle \psi_o(\tau) | \psi_o(\tau) \rangle &= \langle \psi_o | \psi_o \rangle - \frac{i}{\hbar} \langle \psi_u | \hat{U}_1 | \psi_o \rangle + \frac{i}{\hbar} \langle \psi_o | \hat{U}_1 | \psi_u \rangle \\ &+ \frac{1}{\hbar^2} \langle \psi_u | \hat{U}_1^2 | \psi_u \rangle - \frac{1}{\hbar^2} \langle \psi_o | \hat{U}_2 | \psi_o \rangle \\ \langle \psi_u(\tau) | \psi_u(\tau) \rangle &= \langle \psi_u | \psi_u \rangle - \frac{i}{\hbar} \langle \psi_o | \hat{U}_1 | \psi_u \rangle + \frac{i}{\hbar} \langle \psi_u | \hat{U}_1 | \psi_o \rangle \\ &+ \frac{1}{\hbar^2} \langle \psi_o | \hat{U}_1^2 | \psi_o \rangle - \frac{1}{\hbar^2} \langle \psi_u | \hat{U}_2 | \psi_u \rangle \end{aligned} \quad (11.8)$$

At first sight it seems like there were some terms missing. This, however, is not the case. Some terms that should be present if \hat{I} could take on any form vanish due to the special off diagonal block form of \hat{I} (11.1) and the resulting form of the \hat{U} 's. The strict overall probability conservation requires

$$\langle \psi_o(\tau) | \psi_o(\tau) \rangle + \langle \psi_u(\tau) | \psi_u(\tau) \rangle = \langle \psi_o | \psi_o \rangle + \langle \psi_u | \psi_u \rangle = 1 \quad (11.9)$$

This has to hold true for every order separately in the expansion. Thus, exploiting (11.9), for the second order in (11.8) we find

$$\langle \psi_o | \hat{U}_2 | \psi_o \rangle = \langle \psi_o | \hat{U}_1^2 | \psi_o \rangle \quad \langle \psi_u | \hat{U}_2 | \psi_u \rangle = \langle \psi_u | \hat{U}_1^2 | \psi_u \rangle \quad (11.10)$$

Inserting this into (11.8) yields

$$\begin{aligned} \langle \psi_o(\tau) | \psi_o(\tau) \rangle &= \langle \psi_o | \psi_o \rangle - \frac{i}{\hbar} \langle \psi_u | \hat{U}_1 | \psi_o \rangle + \frac{i}{\hbar} \langle \psi_o | \hat{U}_1 | \psi_u \rangle \\ &+ \frac{1}{\hbar^2} \langle \psi_u | \hat{U}_1^2 | \psi_u \rangle - \frac{1}{\hbar^2} \langle \psi_o | \hat{U}_1^2 | \psi_o \rangle \\ \langle \psi_u(\tau) | \psi_u(\tau) \rangle &= \langle \psi_u | \psi_u \rangle - \frac{i}{\hbar} \langle \psi_o | \hat{U}_1 | \psi_u \rangle + \frac{i}{\hbar} \langle \psi_u | \hat{U}_1 | \psi_o \rangle \\ &+ \frac{1}{\hbar^2} \langle \psi_o | \hat{U}_1^2 | \psi_o \rangle - \frac{1}{\hbar^2} \langle \psi_u | \hat{U}_1^2 | \psi_u \rangle \end{aligned} \quad (11.11)$$

Now to evaluate the right hand side exactly one would need to know what $|\psi_o\rangle, |\psi_u\rangle$ are in detail. It could, however, be the case that the right hand side of (11.11) does not even depend significantly on what $|\psi_o\rangle, |\psi_u\rangle$ exactly are. Maybe the right hand side of (11.11) takes on almost the same value for almost all possible $|\psi_o\rangle, |\psi_u\rangle$. This would be the case if the landscape defined by the right hand side, over the region in Hilbert space consistent with given W_o, W_u , was essentially flat. Whether or not this is the case can only be decided by calculating the Hilbertspace variances $\Delta\langle \psi_o(\tau) | \psi_o(\tau) \rangle, \Delta\langle \psi_u(\tau) | \psi_u(\tau) \rangle$. But if

those are small, i.e., if the landscape truly is essentially flat, the right hand side of (11.11) can be replaced by its Hilbertspace average $\langle \langle \psi_o(\tau) | \psi_o(\tau) \rangle \rangle$, $\langle \langle \psi_u(\tau) | \psi_u(\tau) \rangle \rangle$ for a valid approximation. At the moment we will proceed doing so and come back to the Hilbertspace variances later to justify this replacement.

Before calculating the Hilbertspace averages, we introduce abbreviations for the terms occurring in (11.11):

$$\langle \psi_u | \hat{U}_1 | \psi_o \rangle := \alpha \quad \langle \psi_u | \hat{U}_1^2 | \psi_u \rangle := \beta \quad \langle \psi_o | \hat{U}_1^2 | \psi_o \rangle := \gamma. \quad (11.12)$$

For a detailed calculation of the Hilbertspace averages see **H**. Here we just want to give the results and point out there plausibility.

$$\langle \alpha \rangle = 0 \quad \langle \alpha^* \rangle = \langle \alpha \rangle^* = 0 \quad (11.13)$$

$|\psi_o\rangle$ and $|\psi_u\rangle$ lie on different, i.e, entirely independent hyperspheres (see (H.2)). Thus calculating this average eventually amounts to integrating a weighted sum of cartesian coordinates over the full hyperspheres. For symmetry reasons such an integral has to vanish. Consequently the first order contributions vanish.

$$\langle \beta \rangle = \frac{\langle \psi_o | \psi_o \rangle}{N_o} \text{Tr}^o \{ \hat{U}_1^2 \} \quad \langle \gamma \rangle = \frac{\langle \psi_u | \psi_u \rangle}{N_u} \text{Tr}^u \{ \hat{U}_1^2 \} \quad (11.14)$$

Here $\text{Tr}^{o(u)}$ is supposed to denote the trace only over the (upper)lower subspace of the operator. This can be understood by thinking of the average being calculated in the diagonal representation of the operator. Then the average is basically an integration of a sum of the squares of all coordinates weighted by the corresponding eigenvalues of the operator, over a full hypersphere. For symmetry reasons the average of the square of a coordinate has to equal the square of the radius divided by the dimension.

Plugging those results now into (11.11), we get for the above mentioned replacement

$$\begin{aligned} W_o(\tau) &= W_o(0) + \frac{W_u(0)}{\hbar^2 N_u} \text{Tr}^u \{ \hat{U}_1^2 \} - \frac{W_o(0)}{\hbar^2 N_o} \text{Tr}^o \{ \hat{U}_1^2 \} \\ W_u(\tau) &= W_u(0) + \frac{W_o(0)}{\hbar^2 N_o} \text{Tr}^o \{ \hat{U}_1^2 \} - \frac{W_u(0)}{\hbar^2 N_u} \text{Tr}^u \{ \hat{U}_1^2 \}. \end{aligned} \quad (11.15)$$

Now we have to analyze those traces in more detail. We will do this explicitly for the upper subspace but by simply exchanging the indices, the result will be valid for the lower subspace as well.

$$\text{Tr}^o \{ \hat{U}_1^2 \} = \sum_m^{N_o} \langle m | \hat{U}_1^2 | m \rangle = \sum_m^{N_o} \left| \hat{U}_1 | m \rangle \right|^2 \quad (11.16)$$

Here m runs over the eigenstates of \hat{H}_0 in the upper subspace. The object that is summed over here, is evaluated in the literature in the context of Fermi's golden rule [].

$$\left| \hat{U}_1 | m \rangle \right|^2 = \sum_n^{N_u} |\langle n | \hat{I} | m \rangle|^2 \frac{4 \sin^2(\frac{1}{2} \omega_{nm} \tau)}{\omega_{nm}^2} \quad (11.17)$$

with

$$\omega_{nm} = \frac{1}{\hbar} \left(n \frac{B}{N_u} - m \frac{B}{N_o} \right) \quad (11.18)$$

which is essentially an energy difference divided by \hbar . Now our arguments, including the conditions we have to impose on the model, follow closely the ones brought forth in the context of Fermi's golden rule.

The function $f(\omega)$

$$f(\omega) = \frac{\sin^2(\frac{1}{2}\omega\tau)}{\omega^2} \quad (11.19)$$

is basically a peak at $\omega = 0$, with the width $w = \frac{4\pi}{\tau}$ a height of $h = \frac{\tau^2}{4}$ the area under the peak is $A = \frac{\pi\tau}{2}$. This means the peak is getting higher and narrower as τ increases. Defining τ_1 by

$$w(\tau_1) = \frac{4\pi}{\tau_1} = \frac{B}{c\hbar} \quad \Rightarrow \quad \tau_1 = c \frac{4\pi\hbar}{B} \quad (11.20)$$

we can state that for times, longer than τ_1 , the summation in (11.17) stretches over a range that is c times wider, than the width of the function that is summed over. In this case and if, furthermore, the function f does not change much over many summation steps $\Delta\omega$, i.e., if

$$\Delta\omega = \frac{B}{N_u\hbar} \ll w(\tau_1) = \frac{B}{c\hbar} \quad \Rightarrow \quad N_u \gg c \quad (11.21)$$

the summation averages out the different elements of the \hat{I} -matrix in (11.17), the sum may be approximated by the integral times the average matrix element ϵ^2 ,

$$\epsilon^2 = \frac{1}{N_o N_u} \sum_{n,m}^{N_u, N_o} |\langle n | \hat{I} | m \rangle|^2 = \frac{1}{2N_o N_u} \text{Tr} \left\{ \hat{I}^2 \right\}. \quad (11.22)$$

Doing so we find

$$\left| \hat{U}_1 | m \rangle \right|^2 \approx \frac{\epsilon^2 4A}{\Delta\omega} = \frac{2\pi\epsilon^2 \hbar N_u \tau}{B}. \quad (11.23)$$

This approximation breaks down after some time τ_2 , when the peak gets too narrow, i.e.,

$$w(\tau_2) = \frac{4\pi}{\tau_2} = \Delta\omega = \frac{B}{N_u\hbar} \quad \Rightarrow \quad \tau_2 = \frac{4\pi\hbar N_u}{B} \quad (11.24)$$

Thus (11.23) is a valid approximation for $\tau_1 < \tau < \tau_2$. Hence, plugging (11.23) into (11.16) yields:

$$\text{Tr}^o \left\{ \hat{U}_1^2 \right\} = \frac{2\pi\epsilon^2 \hbar N_u N_o \tau}{B} \quad (11.25)$$

Since this expression is symmetric under exchange of the indices u, o , “virtually” redoing the whole procedure with exchanged indices, yields the same expression for the lower subspace:

$$\mathrm{Tr}^u \left\{ \hat{U}_1^2 \right\} = \frac{2\pi\epsilon^2 \hbar N_u N_o \tau}{B}. \quad (11.26)$$

Inserting (11.25) and (11.26) into (11.15) yields

$$\begin{aligned} W_o(\tau) &= W_o(0) + C\tau N_o W_u(0) - C\tau N_u W_o(0) \\ W_u(\tau) &= W_o(0) + C\tau N_u W_o(0) - C\tau N_o W_u(0), \end{aligned} \quad (11.27)$$

where we abbreviated

$$\frac{2\pi\epsilon^2}{B\hbar} := C. \quad (11.28)$$

(11.27) is an equation that describes, within the discussed limits, a short time step starting from any initial state, not necessarily an eigenstate of \hat{H} . Since it directly connects the probabilities W_o, W_u of the initial state with those of the state reached after time τ , we can now iterate this equations under some conditions. We have only considered terms up to second order, and we can only iterate after a time step of length τ_1 . Thus we have to make sure that the considered second order terms are still small compared to 1 after τ_1 , to justify the dropping of higher order terms.

$$\frac{8\pi^2 c}{N_u} \left(\frac{\epsilon}{\Delta E_u} \right)^2 \ll 1 \quad \frac{8\pi^2 c}{N_u} \left(\frac{\epsilon}{\Delta E_o} \right)^2 \ll 1 \quad \Delta E_u := \frac{B}{N_u} \quad \Delta E_o := \frac{B}{N_o} \quad (11.29)$$

If this is fulfilled the “linear regime” is reached while the truncation to second order is still a valid description, and we can iterate (11.27) after some time $\tau > \tau_1$. However, if we want to use the above scheme (11.27) we should make sure that we have to iterate before the linear regime is left again, i.e., before τ_2 . We have to iterate before τ_2 if the considered terms at τ_2 become big compared to 1. Thus, the case for which iterating (11.27) is the best description we can possibly get is:

$$8\pi^2 \left(\frac{\epsilon}{\Delta E_u} \right)^2 \gg 1 \quad 8\pi^2 \left(\frac{\epsilon}{\Delta E_o} \right)^2 \gg 1 \quad (11.30)$$

In this case we get iterating (11.27)

$$\begin{aligned} \frac{W_o((n+1)\tau) - W_o(n\tau)}{\tau} &= CN_o W_u(n\tau) - CN_u W_o(n\tau) \\ \frac{W_u((n+1)\tau) - W_u(n\tau)}{\tau} &= CN_u W_o(n\tau) - CN_o W_u(n\tau) \end{aligned} \quad (11.31)$$

Or, in the limit of τ being extremely small:

$$\begin{aligned} \dot{W}_o &= CN_o W_u - CN_u W_o \\ \dot{W}_u &= CN_u W_o - CN_o W_u \end{aligned} \quad (11.32)$$

This evolution equation for the probabilities obviously conserves the overall probability. Its solutions are simple exponential decays, with exactly the same decay rates one would have gotten from Fermi's golden rule. A solution for the considered system being initially entirely in the excited state reads:

$$W_o(t) = \frac{N_o}{N_u + N_o} + \frac{N_u}{N_o + N_u} e^{-C(N_o+N_u)t} \quad W_u(t) = \frac{N_u}{N_o + N_u} (1 - e^{-C(N_o+N_u)t}). \quad (11.33)$$

The equilibrium values reached after very long times are

$$W_o(\infty) = \frac{N_o}{N_u + N_o} \quad W_u(\infty) = \frac{N_u}{N_o + N_u}, \quad (11.34)$$

which are obviously exactly the same as the ones derived in 5.5 (5.64). For $N_o = 1$ and $N_u \gg 1$ the technique reproduces essentially the result from the Weisskopf-Wigner theory.

Before we compare those results with some numerical data, we want to come back to the Hilbertspace variances, mentioned above. There are basically three variances to consider. The first one refers to the linear part of (11.8)

$$\Delta^2 i(\alpha - \alpha^*) = \langle -\alpha^2 - \alpha^{*2}, +|\alpha|^2 \rangle - i \langle (\alpha - \alpha^*) \rangle^2. \quad (11.35)$$

The averages of α^2, α^{*2} vanish (see (H.9)), so do, as already mentioned, the averages of α, α^* . Thus we get:

$$\Delta^2 i(\alpha - \alpha^*) = \langle |\alpha|^2 \rangle = \langle \psi_u | \hat{U}_1 | \psi_o \rangle \langle \psi_o | \hat{U}_1 | \psi_u \rangle \quad (11.36)$$

Since $\langle |\psi_o \rangle \langle \psi_o| \rangle$ is basically a projector, (see (H.10)), we get

$$\Delta^2 i(\alpha - \alpha^*) = \frac{W_o}{N_o} \langle \psi_u | \hat{U}_1^2 | \psi_u \rangle = \frac{W_o}{N_o} \langle \gamma \rangle > \frac{W_o W_u}{N_o N_u} \text{Tr}^u \{ \hat{U}_1^2 \}, \quad (11.37)$$

where we used (11.14). Thus we find in the linear regime

$$\Delta^2 i(\alpha - \alpha^*) = \frac{2\pi\epsilon^2 \hbar W_o W_u \tau}{B}. \quad (11.38)$$

To judge whether this is big or small, we have to compare this variance, that is a possible error made by the replacement of the actual values in (11.8) by there Hilbertspace averages, to those Hilbertspace averages themselves. Thus we consider, e.g.,

$$\frac{\frac{1}{\hbar^2} \Delta^2 i(\alpha - \alpha^*)}{(\frac{1}{\hbar^2} \langle \beta \rangle)^2} = \frac{\hbar \Delta E_u W_u}{2\pi\epsilon^2 W_o N_u \tau} \quad (11.39)$$

If we now evaluate this at a time τ_i at which the corresponding contribution in (11.8) is still small, i.e., a typical iteration time, e.g., $\frac{1}{\hbar^2} \langle \beta(\tau_i) \rangle = \frac{1}{10}$ we find:

$$\tau_i = \frac{\Delta E_u \hbar}{20\pi\epsilon^2} \Rightarrow \frac{\frac{1}{\hbar^2} \Delta^2 i(\alpha - \alpha^*)}{(\frac{1}{\hbar^2} \langle \beta \rangle)^2} = \frac{10 W_u}{W_o N_u}, \quad (11.40)$$

which means that the square of the relative error possibly committed at each iteration step by replacing actual values by Hilbertspace averages, scales as $\frac{1}{N_u}$. The relative error with respect to the other contributions is simply found by exchanging the indices. The other variances that require consideration are

$$\Delta^2\beta = \langle \beta^2 \rangle - \langle \beta \rangle^2 \quad (11.41)$$

and the same thing for γ . Those are the Hilbertspace averages of expectation values of Hermitian operators. For any such object it can be shown that (see **H** (H.15))

$$\Delta^2\langle\psi|\hat{A}|\psi\rangle = \frac{1}{N+1} \left(\frac{\text{Tr}\{\hat{A}^2\}}{N} - \left(\frac{\text{Tr}\{\hat{A}\}}{N} \right)^2 \right) \quad (11.42)$$

where N is the dimension of the corresponding Hilbertspace. This is obviously the square of the spectral variance divided by $N+1$. In our case, we need to consider $\hat{A} = \hat{U}_1^2$. The diagonal elements of $\hat{A} = \hat{U}_1^2$ are given by (11.17) and are in the considered limit at times for which $w(\tau) \ll \frac{B}{\hbar}$ all the same, except for those within a region of $w(\tau)$ at the edges of the spectrum. The off diagonal elements tend to vanish, since there is no phase correlation between the terms of the corresponding sums. Thus $\hat{A} = \hat{U}_1^2$ will have a small spectral variance which is independent of N_o, N_u . Therefore the square of the Hilbertspace variance vanishes with the dimension of the corresponding space. $\frac{1}{N_o}, \frac{1}{N_u}$

11.1.5 Numerical Results for the Relaxation Period

To check the validity of the theory developed in the previous section **11.1.4**, a model of the type depicted in **Fig. 11.1**, with the Hamiltonian of the type described in (11.1) has been analyzed numerically. The interaction matrix \hat{I} has been filled with random Gaussian distributed entries such that

$$\left(\frac{\epsilon}{\Delta E_u} \right)^2 \approx 1. \quad (11.43)$$

Different container sizes have been analyzed, corresponding to $N_u = 50, 100, 200, 400, 800$ and $N_u = 2N_o$. For the different container sizes the level spacings $\Delta E_u, \Delta E_o$ have been kept fixed, such that for increasing container sizes the bandwidths increase. With those parameter settings the theoretical prediction for $W_u(t)$ from (11.32) is the same for all container sizes. The numerical results are displayed in **Fig. 11.2**. The solid line is the prediction from theory. Obviously the theoretical predictions from theory are not very accurate for “few-level” container environments. This is due to the fact, that the replacement of actual quantities by their Hilbertspace averages (11.15), is only a valid approximation for high dimensional Hilbertspaces. Furthermore, for the few level cases, the iteration step times that have to be longer then τ_1 (11.20) are pretty long, because already τ_1 is long. This means that the recursion (11.31) cannot really be replaced y

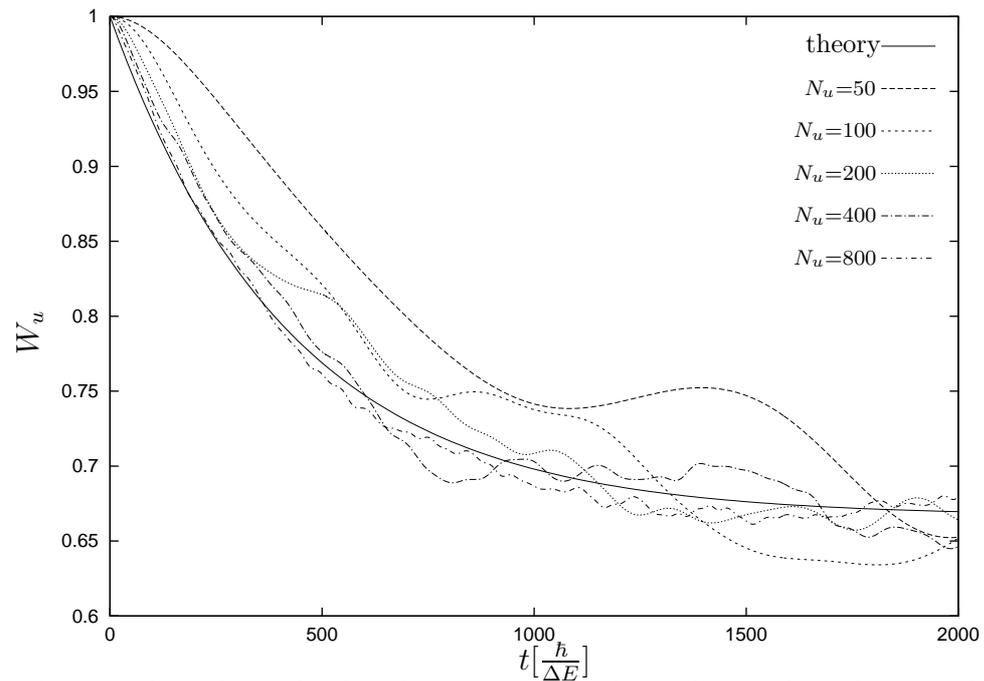


Fig. 11.2: Numerical simulation for the relaxation period. The predictions from theory get better with increasing container systems.

the differential equation (11.15). This essentially shifts the corresponding curves to later times, compared to the theoretical prediction. All those effects vanish if the container system becomes sufficiently big. The simulation for $N_u = 800$ is obviously in good agreement with the theoretical prediction.

12 Summary and Conclusion

This thesis essentially has two intentions: To define what precisely thermodynamical behavior is, and to show that a certain class of quantum mechanical systems show this thermodynamical behavior, even if treated merely quantum mechanically, i.e., if their behavior is exclusively described by the Schrödinger equation. Since it seems plausible that this class of systems may contain almost all systems that one typically thinks of as being thermodynamical (and even some more), this thesis may be viewed, as the title says, as a quantum approach to thermodynamics.

First of all a brief overview over the historical attempts to “derive” thermodynamics from an underlying theory, for the biggest part from classical mechanics, with an emphasis on their problems is given **2**. (It may be noticed here, that the author, despite of his critical attitude in this chapter, does not think that those attempts are all wrong, in fact he feels that some of those belong to the most brilliant work that has ever been done in theoretical physics). The main problem in this field seems to be the precise definition of entropy as an observable (function of the micro state), and the demonstration that this entropy exhibits the behavior, one expects from an axiomatically defined entropy, without exploiting additional assumptions like ergodicity.

After that, some sort of axiomatic structure, a set properties of thermodynamical quantities and their dynamics (entropy, pressure, etc) is given, that is meant to specify thermodynamical behavior. This set essentially corresponds to the different statements that are contained in the first and second law of thermodynamics. It is not necessarily considered unique, irreducible or entirely new, but organized in such a way, that every rule or property can be shown to result from an underlying theory -Schrödinger-type quantum mechanics- in the context of the approach presented here.

The main part of this thesis consists of the definitions of thermodynamical quantities on the basis of the momentary, local, quantum mechanical micro state (density operator) of the considered system, and the demonstration that those quantities indeed exhibit, under certain conditions, the claimed behavior (Those conditions than define the thermodynamical limit). The general set-up always consists of an considered system weakly coupled to another system. This weak coupling may or may not allow for energy exchange. The micro state of the full system, i.e., considered system and environment, is assumed to be initially given by a pure state, i.e., a wavefunction, and, due to the Schrödinger equation, remains to be so. A micro state of the considered system only, is produced by tracing out the environment of the micro state of the full system, thus getting a density matrix. Entropy, e.g, is now defined as the Von Neumann entropy of this reduced density matrix. In this or similar ways, all thermodynamical quantities are

defined as functions of the micro state of the full system, which may be represented by a point in Hilbertspace. The whole Hilbertspace may now be viewed as being divided into precisely defined cells, by demanding that all points that belong to one zone, have to yield the same thermodynamical quantity like entropy, energy etc. Those cells that correspond to quantities that are conserved, like energy under microcanonical conditions, represent restrictions to the possible evolutions of the full system in Hilbertspace. It now turns, out that the biggest part of all different micro states within some “accessible cell” yield the same equilibrium values or the thermodynamical quantities, like maximum entropy or identical temperatures for subsystems etc. That means the biggest part of the accessible Hilbertspace is filled with states that correspond to thermodynamical equilibrium. The assumption now is that a system which starts in a region in Hilbertspace that does not correspond to equilibrium, will typically eventually end up wandering through a region that does correspond to equilibrium, simply because this region is so much bigger. It should be noted here, that this assumption, although it is an assumption, is much weaker than the assumption of ergodicity or quasi-ergodicity. (The concrete systems we analyzed are definitely not ergodic, nevertheless they show thermodynamical behavior as may be seen from the numerical simulations).

In this scenario, entropy is a measure for the entanglement between the considered system and its environment. Thus saying that almost all states of the full system are maximum local entropy states, amounts to saying that almost all full system states are highly entangled states. Thus, in this picture, the increase of entropy is due to the built up of entanglement of the considered system with its environment induced by the interaction, regardless of whether this interaction allows for any exchange of extensive quantities.

It turns out that some features of thermodynamics, e.g., the full independence of the equilibrium state of some system coupled to a large heat bath, from its initial state, require a certain structure of the spectrum of this heat bath. Since pretty much any larger macroscopic system can possibly play the role of a heat bath, this structure should be somehow generic. It is thus shown in **6.2**, that this structure, which is also linked to the extensivity of entropy, namely that the logarithm of the statedensity is a homogeneous function of the first grade, results, whenever very many weakly interacting identical units form the entire system, i.e., whenever the system shows some modular structure (This consideration is not necessarily restricted to quantum mechanics).

The next chapters deal with the intensive quantities and their property as derivatives of energy with respect to a corresponding extensive quantity. It is shown, that there are processes (heating, cooling, adiabatic volume changes, etc.), that fundamentally have to be described by the Schrödinger equation, possibly with time dependent Hamiltonians, in which the quantum mechanically defined intensive quantities, happen to result as those derivatives. In the case of pressure that amounts, first of all, to showing that there are processes in which volume changes but entropy does not. This is essentially done by enlarging the concept of adiabatic following, known from quantum optics.

Trying to roughly summarize the different claims and assumptions that have been exploited to establish thermodynamical behavior, one can say that quantum mechanical systems that are weakly coupled to other quantum mechanical systems with higher state

densities, should already exhibit a lot of the features of thermodynamical behavior, full thermodynamical behavior, however, can only be established this way, if all systems in quest have spectra, featuring this structure which is typical for modular systems.

Since entropy as a thermodynamical potential determines the concrete thermodynamical behavior of a system entirely and statedensity as a function of energy and, say, volume, determines entropy, two contradicting types of behavior might result for the same system, depending on whether the statedensity is computed based on a classical or a quantum mechanical model. (In the classical case the “statedensity” is just the volume of some energy shell in phase space.) Thus the connection between classical and quantum mechanical statedensity has been investigated, with the result that, whenever the energy spread of some minimum position-momentum uncertainty wavepackage is small compared to its mean energy, classical state density divided by \hbar may be seen as a good approximation to quantum mechanical state density. The above condition is fulfilled either by systems consisting of heavy objects, or systems made of many identical weakly interacting subsystems, which is again the condition of “modularity”.

In the last chapter the way to equilibrium has been examined. Usually it is stated that a system in some excited state has a certain statistical transition probability to its ground state, and thus an exponential decay results. If the system is entirely controlled by the Schrödinger equation, this statistical character needs explanation, and there are several attempts to explain it, by additional influences that are not described by the Schrödinger equation. Here we investigated, if, and under what conditions, such an exponential decay may result for some system, if this system is again weakly coupled to some other system, and the whole system is only described by the Schrödinger equation. It is found that this is possible and the conditions for it to happen are pretty much the same as the ones that guarantee the applicability of “Fermi’s golden rule”. Some numerical simulations are presented to back up and illustrate the results from this theory.

As usual a concise conclusion from this thesis is hard to draw. Certainly it cannot be that all approaches to a foundation of thermodynamics done so far are entirely wrong, and this work establishes the only unique access that is beyond any doubt. Nevertheless the idea of thermodynamics being an emergent pattern of behavior of special compound quantum systems might contribute as a valuable alternative point of view to an old discussion. It carries the potential to see the limits of the applicability of thermodynamics somewhat clearer. And those seem not necessarily bound to some minimum number of particles of the system in quest. In the context of this theory, thermodynamical behavior of microscopic single objects seems possible. Furthermore violations of the second law appear as not entirely impossible, but extremely unlikely. Nevertheless the theory might provide hints on how systems could possibly be made, that allow to see such a violation. It is our hope that future experiments might help to clarify the applicability and significance of this theory, rather than mere theoretical reasoning.

13 Zusammenfassung

13.1 Einleitung und historischer Hintergrund

Die Thermodynamik mit ihren Grundbegriffen und Hauptsätzen ist eine Theorie, deren Gültigkeit und Anwendungsrelevanz heute niemand mehr bezweifeln kann. Dessen ungeachtet findet, seit Wissenschaftler wie Boltzmann ihre Grundsteine legten, eine Diskussion statt, über die Möglichkeit bzw. Notwendigkeit, sie aus anderen, fundamentaleren Theorien abzuleiten. Eng damit verknüpft ist die Frage nach den Grenzen ihrer Anwendbarkeit. Heute, da Experimente mit einzelnen Quanten-Objekten im Zentrum des Interesses stehen, gewinnen diese Fragen wieder an Brisanz.

Als Beginn der wissenschaftlichen Thermodynamik könnte man die Versuche ein Thermometer zu bauen nennen, die Santorio und Galileo erst um 1600 unternahmen, als z.B. der Bau des Sonnensystems im wesentlichen schon bekannt war. Der entscheidende Vorstoß zur Eingliederung thermodynamischer Phänomene in das damalige Gebäude der Physik gelang dann Joule 1840, als er seine Versuche zum Wärmeäquivalent durchführte und damit Wärme als Energieform etablierte. Bald danach entstand die Formulierung des ersten Hauptsatzes durch Clausius, der auch die ersten Versuche zur phänomenologischen Definition der Entropie unternahm und mit dem zweiten Hauptsatz die Irreversibilität einführte.

Ein grundlegender Beitrag zur Berechnung thermodynamischen Verhaltens aus der Newtonschen Theorie kam von Boltzmann, der 1866 forderte, dass die phänomenologische Entropie identifiziert werden müsse mit dem Logarithmus der Zahl der Mikrozustände eines Systems, die zu einem gegebenen Makrozustand gehören. Mit diesem "Rezept" gelang zum erstenmal die Bestimmung des thermodynamischen Verhaltens eines Gases aus seiner Hamiltonfunktion. Boltzmann war allerdings mit seiner Arbeit nicht vollständig zufrieden, da die Entropie nach dieser Definition keine Observable (keine Funktion des Mikrozustandes) im Hamiltonschen Sinn ist und ihre Dynamik sich deshalb der Analyse entzieht.

In der Folge bestanden daher eigentlich alle weiteren Versuche, die Thermodynamik auf die klassische Mechanik zu reduzieren darin, eine Observable zu definieren, die mit der Zeit immer bis zu der von Boltzmann definierten Grenze anwachsen würde. Boltzmann selbst führte im H -Theorem eine solche Größe ein, von der mit Hilfe der Hamiltonschen Theorie gezeigt werden konnte, dass sie immer wachse, allerdings nur unter der

Annahme der Unkorreliertheit der Teilchen, zudem konvergiert H nicht gegen die gewünschte Grenze.

Ein weiterer Ansatz von ihm bestand darin, die Entropie mit dem Volumen zu verknüpfen, welches von der Trajektorie eines Systems im Phasenraum, also dem $6N$ -dimensionalen Raum der Orte und Impulse aller N Teilchen, zu einem bestimmten Zeitpunkt bereits durchlaufen worden war. Damit eine so definierte Größe aber die gewünschten Eigenschaften aufweisen kann, muss die Trajektorie jedem Punkt des zugänglichen Phasenraums beliebig nahekommen, und sich in jeden Volumenelement gleich lange aufhalten. Diese Eigenschaft ist die berühmte (Quasi-) Ergodizität. Da es unmöglich war, diese Eigenschaft für die riesige Klasse der thermodynamischen Systeme explizit zu zeigen, entstand die Ergodenhypothese.

Einen anderen Zugang entwickelte Gibbs um 1900. Er ging von der Unmöglichkeit aus, den Zustand eines Makro-Systems genau zu bestimmen; so beschrieb er den Zustand durch ein kleines Volumen statt durch einen Punkt im Phasenraum. Alle Punkte dieses Raums bilden dann das Ensemble. Dieses Volumen sollte durch die Dynamik der Hamiltonschen Gleichungen wachsen und schließlich den gesamten Phasenraum füllen. Da dieses Volumen aber nach dem Liouvillschen Satz invariant ist, nahm Gibbs an, dass es sich zu einer schwammartigen Struktur verformen würde, die letztlich doch effektiv über den ganzen Raum verteilt wäre. Gibbs verglich dieses Verhalten mit dem eines Tropfens Tinte, den man in ein Glas gibt und verrührt. Daher nennt man es noch heute "mischend". Allerdings ließ sich diese Eigenschaft auch nur für spezielle Systeme nachweisen. Ein weiterer Versuch, die Entropie stringent zu formulieren, geht auf Ehrenfest zurück, der den Phasenraum so in Zellen unterteilte, dass eine Zelle nur solche Zustände enthielt, die makroskopisch nicht zu unterscheiden sein sollten. In einer Weiterentwicklung wurde die Entropie eines Zustandes dann verknüpft mit dem Volumen der Zelle, in der er sich gerade aufhält. Die Zelle, die dem makroskopischen Gleichgewicht entspricht, sollte so überwältigend viel größer sein als alle anderen Zellen, dass das System schlussendlich immer in sie hineinlaufen würde. Diese Tendenz ist oft vorhanden, dennoch bereiten die Zellgrenzen Probleme, da die Methode zu ihrer Festlegung nicht eindeutig klar ist.

Aus heutiger Perspektive erscheinen alle diese Versuche, die Thermodynamik aus der klassischen Mechanik abzuleiten, fragwürdig. Seit langem ist klar, dass typische thermodynamische Systeme aus mikroskopischen Teil(ch)en bestehen, und daher quantenmechanisch behandelt werden müssen. Zwar können im Ehrenfest-Limit, also wenn die Ausdehnung der Wellenpakete klein ist gegen die Strukturen des Potentials, die ihre Dynamik bestimmen, quantenmechanische Mittelwerte wie klassische Koordinaten behandelt werden, aber es ist leicht auszurechnen, dass sich ein typisches Gas, wegen der Kleinheit und der Leichtigkeit seiner konstituierenden Teilchen, nicht lange in diesem Limit befinden kann. Aber selbst abgesehen von dieser Problematik sind Aspekte der Thermodynamik und Statistik unübersehbar, welche bereits die Mitnahme von Quanteneffekten notwendig machen: die korrigierte Boltzmann-Statistik, der dritte Hauptsatz, sowie die Einflüsse der Quantenstatistik.

Auch ist heute klar, dass die klassische Zustandsumme, ein zentraler Begriff der Thermostatistik, als ein Grenzfall der quantenmechanischen gesehen werden kann. Das alles legt den Versuch nahe, ganz allgemein die Prinzipien der Thermodynamik (Anwachsen

der Entropie, entstehen eines Gleichgewichts) direkt auf quantenmechanische Evolutionsgleichungen wie die Schrödinger-Gleichung zurückzuführen. So versuchten bereits Schrödinger und Von Neumann, die Thermodynamik konsequent aus quantenmechanischen Überlegungen herzuleiten. Lindblad stellte fest, dass die Entropie eines Vielteilchensystems, definiert als die Summe der Teilentropien, anwachsen sollte durch die Vernachlässigung der "Verschränkung" zwischen den Teilchen. Zurek und Mitarbeiter argumentierten, dass der zweite Hauptsatz letztlich aus der Unmöglichkeit folgen sollte, makroskopische Quantensysteme von ihrer Umgebung zu isolieren. Landau und Lifschitz erhofften dagegen eine Deutung des 2.Hauptsatzes über den quantenmechanischen Messprozess.

13.2 Axiomatischer Aufbau der Thermodynamik

In dieser Arbeit wurde der Versuch unternommen, einen Satz von Regeln und Eigenschaften bzgl. thermodynamischer Größen und ihrer Dynamik zusammenzustellen, der dann in seiner Gesamtheit thermodynamisches Verhalten definiert. Die meisten dieser Eigenschaften ergeben sich aus dem ersten oder zweiten Hauptsatz der Thermodynamik. Wenn alle Eigenschaften aus einer darunterliegenden, fundamentaleren Theorie (hier der Schrödingerschen Quantenmechanik) abgeleitet werden können, wie es in dieser Arbeit versucht wird, ist eine Herleitung der Thermodynamik gelungen.

Definitionen und thermodynamischer Limes

Alle thermodynamischen Größen also Entropie, Temperatur, etc, sollten als echte Observable, also als lokale und momentane Funktionen der Größen definiert sein, deren Dynamik von der darunterliegenden Theorie beschrieben wird. Da die darunterliegende Theorie notwendig mehr Phänomene als die thermodynamischen erfasst, werden thermodynamische Größen auch dann definiert sein wenn sie sich nicht thermodynamisch verhalten. Man benötigt also einen thermodynamischen Limes, ein (hinreichend einfaches) Kriterium, welches Auskunft darüber gibt, wann thermodynamisches Verhalten überhaupt erwartet werden kann.

Zweiter Hauptsatz

Thermodynamische Systeme sollten nach einer Relaxationszeit Gleichgewichtszustände erreichen, die durch ein mit den Randbedingungen verträgliches Maximum an Entropie gekennzeichnet sind. Gleichgewicht bedeutet dass alle thermodynamischen Größen vernachlässigbar zeitlich fluktuieren. Der Gleichgewichtszustand sollte im mikrokanonischen Fall nur von der Energie des Anfangszustandes, und im kanonischen Fall überhaupt nicht vom Anfangszustand abhängen.

Erster Hauptsatz und intensive Variablen

Es sollte gezeigt werden, dass die wie oben beschrieben definierten intensiven Größen, in bestimmten thermodynamischen Prozessen (also hier letztlich durch die Schrödinger-Gleichung mit zeitabhängigen Hamiltonoperatoren kontrollierten Prozessen), momentan

die Werte der Ableitungen der Energie nach den entsprechenden extensiven Variablen annehmen.

Es sollte außerdem gezeigt werden, dass intensive Variablen im Gleichgewicht dieselben Werte annehmen, wenn die entsprechenden extensiven Variablen ausgetauscht werden können.

13.3 Grundprinzipien der Theorie

Ausgangspunkt aller folgenden Überlegungen ist letztlich, dass jedes thermodynamische System eine Umgebung hat mit der es interagiert. Dies gilt insbesondere auch für Systeme unter mikrokanonischen Bedingungen. Ein Gas, z.B. in einem isolierenden Gefäß, wechselwirkt sicherlich mit dem Gefäß, selbst wenn dabei keinerlei Energie ausgetauscht wird. Sogar diese Art "isolierender Kopplung" kann für quantenmechanische Systeme zu einem Verhalten führen, dass von dem, das entstünde, wenn die Kopplung und die Umgebung gar nicht vorhanden wären, wie im Folgenden gezeigt werden soll, stark abweicht.

Geht man weiter davon aus dass alle thermodynamischen System auch quantenmechanisch sind, so hat man die Dynamik schwach gekoppelter, zusammengesetzter Quantensysteme zu untersuchen. Die (reinen) Gesamtzustände solcher zusammengesetzter Systeme, (wir bezeichnen hier aus Gewohnheitsgründen das betrachtete System als "Gas" (g), z.B. mit N^g Energieeigenzuständen, das Umgebungssystem als "Container" (c), z.B. mit N^c Energieeigenzuständen) sind in einem Produkt-Hilbertraum der Dimension $N = N^g N^c$ als Zustandsvektoren (Wellenfunktionen) definiert. (Wir beschränken uns hier zunächst auf endliche Dimensionen.) Der Zustand des "eigentlich interessierenden" Systems, g , ist dann im allgemeinen nicht mehr als Zustandsvektor, sondern nur noch als Dichte-Operator $\hat{\rho}^g$ darstellbar. Dieser Dichte-Operator entsteht durch "Reduktion" (Teilspurbildung) aus dem vollen Zustand des Gesamtsystems. In diesem Dichteoperator sind alle ausschließlich das Gassystem betreffenden Informationen enthalten, auf der Basis dieses reduzierten Dichteoperators werden also auch alle thermodynamischen Größen des Gassystems definiert, z.B. dessen Entropie als die Von Neumann Entropie S^g des zugeordneten Dichteoperators. Auf diese Weise sind alle thermodynamischen Größen letztlich momentan und lokal als Funktionen des Gesamtzustandes im Hilbertraum definiert.

Ähnlich wie sich der klassische Zustand auch von Vielteilchensystemen als ein Punkt im klassischen Phasenraum visualisieren läßt, kann man eine kartesische Darstellung des Hilbertraums finden, in der man sich den Zustand eines Quantensystems als Punkt veranschaulichen kann. Es lässt sich sogar eine Darstellung finden, sodass von der Schrödingergleichung generierte Trajektorien mit konstanter Geschwindigkeit durchlaufen werden. Durch die punktuell auf ihm definierten thermodynamischen Größen, lässt sich der so veranschaulichte Hilbertraum in verschiedene Zellen unterteilen. Dies geschieht dadurch dass alle Punkt die zu ein und dem selben Wert einer thermodynamischen Größe gehören, zu einer Zelle zusammenfasst werden. Zellen die zu Erhaltungsgrößen gehören, also z.B. die Zelle gebildet aus allen Zuständen gleicher Energie, stellen "zugängliche Bereiche" für alle möglichen Entwicklungen dar, Trajektorien die in einer solchen Zelle beginnen

können diese nicht verlassen, ähnlich wie klassische Trajektorien in einer Energieschale im Phasenraum.

Kann nun gezeigt werden, dass innerhalb einer zugänglichen Zelle fast alle Zustände zu den selben Gleichgewichtswerten für alle thermodynamischen Größen gehören, und für die konkrete Trajektorie keinerlei weitere Einschränkungen existieren als die Bindung an eben Zelle, so kann man folgern: Im allgemeinen werden Trajektorien, insbesondere solche die in einem (verschwindend kleinen) Nichtgleichgewichtsgebiet beginnen, irgendwann mit hoher Wahrscheinlichkeit in den (überwältigend großen) Gleichgewichtsbereich gelangen, und diesen dann praktisch nicht mehr verlassen. Die Wahrscheinlichkeit für ein solches Verhalten wird umso größer, je drastischer das Verhältnis der Größe des Gleichgewichtsbereichs zu der des Nichtgleichgewichtsbereichs ist. Diese Folgerung sollte nicht mit der Ergodenhypothese verwechselt werden. Die (Quasi)Ergodenhypothese fordert ja, dass eine Trajektorie in einer gegebenen Zeit durch alle ihr zugänglichen Volumenelemente des Phasenraums bewegt, und sich dabei in jedem gleich lange aufhält. Hier wird nur vermutet, dass sich eine Trajektorie aus einem kleinen Bereich herausbewegt, um dann einen großen zu durchwandern. Die Details dieser Bewegung sind völlig unerheblich.

In dem hier beschriebenen Szenario stellt die Entropie gleichzeitig auch ein Maß der Verschränkung des Gassystems mit dem Containersystem dar. Die Aussage dass die meisten möglichen Zustände, Zustände maximaler Entropie seien, ist also äquivalent zu der Aussage, dass die meisten möglichen Zustände hochverschränkte Zustände sind. Wenn sich also ein System von niederen zu immer höheren Entropien entwickelt, wächst dabei seine Verschränkung mit der Umgebung. Deswegen ist die oben erwähnte Kopplung unverzichtbar. Ohne diese Wechselwirkung könnte keine Verschränkung entstehen. Man könnte vermuten dass, wenn die Kopplung schwach ist, und sie muss in der Tat schwach sein um die Anwendbarkeit der im Folgenden dargestellten Konzepte zu garantieren, die Verschränkung auch nur wenig anwachsen könne, dass ist aber nicht der Fall. Sobald irgendeine Kopplung vorhanden ist, ist die Verschränkung keine Erhaltungsgröße mehr, und kann im Prinzip beliebig anwachsen. Mit abnehmender Kopplungsstärke werden einfach die Zeiten in denen die maximale Entropie erreicht wird immer länger.

Es sollte noch erwähnt werden dass Entropie hier, wie in vielen klassischen Konzepten auch als Informationsmangel interpretiert werden kann. Aber im Gegensatz zu den klassischen Konzepten in denen Informationsmangel ja immer "subjektiven" Informationsmangel bedeutet, den irgendjemand der genauer misst oder exakter rechnet vielleicht irgendwie überwinden könnte, ist die Entropie hier ein fundamentaler Informationsmangel, genau wie der aus der Unschärferelation resultierende Informationsmangel über den Ort eines Teilchens dessen Impuls genau bekannt ist. In diesem Fall betrifft der Informationsmangel aber nicht nur eine Observable wie den Ort, sondern alle lokalen Größen des betreffenden Systems sind maximal unscharf, scharf sind allein irgendwelche Quantenkorrelationen zwischen dem System und seiner Umgebung die technisch praktisch unmöglich zu messen sind.

13.4 Die Zellstruktur des Hilbertraums

Die mathematische Hauptschwierigkeit der vorliegenden Theorie besteht in der Analyse der Zellstruktur des Hilbertraums. Man benötige eine Methode um herauszufinden ob es für eine bestimmte Größe tatsächlich ein größte Zelle im Hilbertraum gibt, die einen zugänglichen Bereich fast vollständig ausfüllt, und in der sich nur Zustände befinden die alle demselben Wert dieser Größe entsprechen. Betrachtet man diese Größe als eine Art Höhe die über dem (hochdimensionalen) Hilbertraum aufgetragen wird, so ergibt sich eine "Landschaft". Ist diese zu einer Größe gehörige Landschaft weitgehend "flach", so existiert eine Gleichgewichtszelle und man kann annehmen, dass die Größe dynamisch eine Tendenz zum Gleichgewicht zeigt.

Eine Methode zur Analyse dieser Landschaft ist numerischer Natur. Man kann einen Satz von Zuständen (letztlich Punkte in einem hochdimensionalen kartesischen Raum) erzeugen, die auf dem zugänglichen Bereich (letztlich eine Menge bestehend aus Hypersphären) gleichverteilt sind. Zu jedem dieser Zustände berechnet man dann die zu untersuchende Größe und sortiert die Zustände dieser Größe entsprechend in ein Histogramm ein. Ergibt sich so eine Verteilung die an einer Stelle scharf gepeakt ist, so ist die entsprechende Landschaft tatsächlich weitgehend flach auf einer Höhe die dem Wert entspricht, an dem sich der Peak befindet. Die Breite der Verteilung ist ein Maß dafür, wie groß die zeitlichen Fluktuationen sind die man für die Größe im Gleichgewicht erwarten kann. Diese Methode hat den Vorteil dass sie ein relativ detailliertes Bild liefert, und mit entsprechen rechenstarken Computern immer durchgeführt werden kann, aber sie bietet natürlich keine Möglichkeit auf die Form solcher Landschaften im allgemeinen zu schließen.

Um ein weniger detailliertes, aber dafür umfassenderes Bild zu erhalten, bietet sich eine andere, analytische Methode an. Diese besteht darin, zunächst einmal den Mittelwert der entsprechende Größe über den zugänglichen Bereich zu berechnen. Ist die Größe als $f = f(|\psi\rangle)$ definiert so ist dieser Mittelwert zu berechnen als:

$$\langle f \rangle = \frac{\int f(|\psi\rangle) \prod_i \psi_i}{\int \prod_i \psi_i}, \quad (13.1)$$

wobei die ψ_i den Dimensionen des Hilbertraums entsprechen, und sich das Integral nur über den zugänglichen Bereich erstreckt. Das ist im allgemeinen nur durch Parametrisierung des Bereichs möglich, so dass zur Bewertung des Integrals die entsprechende Funktionaldeterminante berechnet werden muss. Dieser Mittelwert bestimmt dann den möglichen Gleichgewichtswert der Größe, nicht aber ob diese tatsächlich ein Gleichgewicht ausbilden wird. Der Wert bestimmt ja nur die mittlere Höhe der Landschaft, nicht aber ob diese flach ist. Um das festzustellen, muss man weitere Betrachtungen anstellen. Falls z.B die Größe f auf dem zugänglichen Bereich eine untere Schranke, f_{min} , hat, und der Mittelwert nur wenig über dieser Schranke liegt $\langle f \rangle \approx f_{min}$, so kann man folgern dass die Landschaft weitgehend flach sein muss, denn, um es bildlich auszudrücken, wenn die mittlere Höhe einer Landschaft nur wenig über dem tiefsten Punkt ihres tiefsten Tales liegt, können sich in ihr nicht viele, hohe Berge erheben. Im allgemeinen wird man aber einen solchen Fall nicht vorfinden und muss die "Welligkeit" einer solchen Landschaft Δf

berechnen:

$$\Delta^2 f = \frac{\int f^2(|\psi\rangle) \prod_i \psi_i}{\int \prod_i \psi_i} - \langle f \rangle^2 \quad (13.2)$$

Ist dieses Maß klein, so ist die zugehörige Landschaft flach, und die entsprechende Größe wird eine Tendenz zum Gleichgewicht zeigen.

13.4.1 Mikrokanonische Bedingungen

Unter mikrokanonischen Bedingungen ist der zugängliche Bereich durch die innere Energie des betrachteten Systems selbst, bzw seine quantenmechanische Energiewahrscheinlichkeitsverteilung, die eine Erhaltungsgröße ist, bestimmt. Die zu untersuchende Größe ist hier also die Entropie. Da diese aber analytisch schon für die einzelnen Zustände schwer zu bestimmen ist (das würde die Bestimmung der Eigenwerte hochdimensionaler Matrizen beinhalten) ist ihr Hilbertraummittel kaum zu berechnen. Deshalb wählt man hier eine indirekte Methode. Zur höchsten möglichen Entropie im zugänglichen Bereich gehört nur eine einzige Dichtematrix für das betrachtete System, $\hat{\rho}_{eq}$. Die allgemeine Dichtematrix läßt sich schreiben als

$$\hat{\rho} = \hat{\rho}_{eq} + \hat{\epsilon} \quad (13.3)$$

wobei $\hat{\epsilon}$ die Abweichungen vom Gleichgewicht beschreibt. Diese Abweichungen lassen sich skalar quantifizieren als

$$F = \text{Tr} \{ \hat{\epsilon}^2 \} \quad (13.4)$$

Aufgrund der Hermitizität ist F immer größer Null. F ist mathematisch simpler als die Entropie und der entsprechende Mittelwert läßt sich analytisch berechnen:

$$\langle F \rangle \approx \sum_B \frac{(W_B^c)^2}{N_B^c} \quad (13.5)$$

wobei W_B^c für die Besetzungswahrscheinlichkeit des Energieniveaus E_B^c des Containersystems steht, welche unter mikrokanonischen Bedingungen eine Erhaltungsgröße und daher durch den Anfangszustand bestimmt ist, und N_B^c den Entartungsgrad des Energieniveaus E_B^c bedeutet. Wenn also das Containersystem hochentartete Energieniveaus besetzt und/oder die Besetzungswahrscheinlichkeit über viele Niveaus verteilt ist kann dieser Mittelwert extrem klein werden. Da F auf dem zugänglichen Gebiet nach unten durch Null beschränkt ist, kann man in diesem Fall folgern dass F und damit die Abweichungen vom Gleichgewichtszustand höchster Entropie, überall sehr gering sind. In diesem Fall ist die Zellstruktur die zu einer Größe mit Tendenz zum Gleichgewicht führt etabliert, und man kann erwarten, dass unter solchen mikrokanonischen Bedingungen, obwohl das System energetisch von seiner Umgebung isoliert ist, die Entropie bis zu einem Maximum anwächst. Diese maximale Entropie berechnet sich aus dem Gassystem und seinem Anfangszustand als

$$S_{max} = -k \sum_A W_A^g \ln \frac{W_A^g}{N_A^g} \quad (13.6)$$

und ist damit genau die Größe, die man durch Anwendung des ‘‘a priori’’ Postulats auf Energieeigenzustände erhalten würde.

In 5.8.1 sind diese Prinzipien numerisch illustriert.

13.4.2 Kanonische Bedingungen

Unter kanonischen Bedingungen sind die Besetzungswahrscheinlichkeiten der Energieniveaus der individuellen Subsysteme keine Erhaltungsgrößen sondern nur noch die Besetzungswahrscheinlichkeiten der Energieniveaus des Gesamtsystems. Daher verändert sich der zugängliche Bereich, und um den Gleichgewichtszustand zu bestimmen müssen zunächst die Hilbertraummittel der individuellen Besetzungswahrscheinlichkeiten bestimmt werden. Man findet:

$$\langle W(E_A^g) \rangle = N^g(E_A^g) \sum_E \frac{N^c(E - E_A^g) W(E)}{N(E)} \quad (13.7)$$

für das Gassystem und entsprechendes für das Containersystem, wobei $W(E)$ und $N(E)$ die Besetzungswahrscheinlichkeit für das Gesamtenergieniveau E , sowie dessen Entartungsgrad bedeuten. Auch $\Delta W(E_A^g)$ läßt sich berechnen, es ergibt sich eine komplizierte Funktion, die aber im wesentlichen mit größer werdenden Entartungsgraden des Gesamtsystems immer kleiner wird, also ist, für ausreichend große Umgebungen, auch für die $W(E_A^g)$ ein Gleichgewicht zu erwarten. Ob fast alle Zustände die zu dieser ‘‘Besetzungswahrscheinlichkeitsgleichgewichtszelle’’ gehören tatsächlich auch Zustände der maximalen, mit dieser Wahrscheinlichkeitsverteilung verträglichen Entropie sind, läßt sich feststellen in dem man gedanklich die Gleichgewichtsverteilung als mikrokanonisch festgelegt annimmt, und dann $\langle F \rangle$ prüft. Wiederum wird das für große Umgebungssysteme tatsächlich der Fall sein. In 5.8.2 sind diese Prinzipien numerisch illustriert.

13.5 Spektren modularer Systeme

Die kanonische Gleichgewichtsverteilung (13.7) ist offensichtlich im allgemeinen nicht die bekannte Boltzmannverteilung. Diese ergibt sich erst, wenn die Zustandsdichte des Containersystems, zumindest über einen gewissen Bereich, durch eine Exponentialfunktion beschreibbar ist, d.h. $N^c(E) \approx N(U) \exp(\beta E)$. Eine solche exponentielle Struktur wird auch sonst oft für eine Ableitung der Boltzmannverteilung benutzt, sie ist nicht unplausibel, rigide begründen läßt sie sich aber nicht, was man schon daran sieht, dass es Systeme gibt, die ihr nicht entsprechen. Man kann aber zeigen, dass für Systeme die aus vielen (M) identischen, schwach gekoppelten Subsysteme bestehen, der Logarithmus der Zustandsdichte folgende Form annimmt

$$\ln N(E) \approx M \quad s \left(\frac{E}{M} \right) \quad (13.8)$$

wobei s sich aus dem Spektrum des Subsystems ergibt. In jedem Fall handelt es sich aber um eine homogene Funktion ersten Grades was bedeutet dass die Funktion auf einem

gegebenen Intervall für zunehmende M immer "linearer" wird. Das wiederum bedeutet, dass für zunehmende M die Zustandsdichte tatsächlich über immer weitere Bereiche durch eine Exponentialfunktion beschrieben wird. Letztlich gilt also, dass das Gassystem sich in eine Boltzmannverteilung hinein entwickeln wird, wenn das Containersystem eine solche modulare Struktur zeigt. Genau in diesem Fall ist auch der kanonische Gleichgewichtszustand völlig unabhängig vom Anfangszustand des Gassystems.

13.6 Spektrale Temperatur

Auf der Basis der Energieniveaubesetzungswahrscheinlichkeiten und dem Spektrum eines Systems läßt sich eine Funktion, die spektrale Temperatur, definieren die für den Fall dass zwei Systeme sich im oben beschriebenen kanonischen Gleichgewicht befinden folgende Eigenschaften hat:

- Falls beide Systeme so dichte Spektren haben dass sie gut durch kontinuierliche Zustandsdichten beschrieben sind, sind ihre spektralen Temperaturen gleich
- Falls ein System ein Spektrum hat welches diskret behandelt werden muss, das andere aber im relevanten Bereich die für modulare Systeme exponentielle Zustandsdichte zeigt, sind ihre spektralen Temperaturen gleich
- Falls eine Änderung der inneren Energie in einem beliebigen System welches an ein großes modulares System angekoppelt ist, dadurch herbeigeführt wird dass es an ein anderes großes, modulares System mit infinitesimal benachbarter spektraler Temperatur angekoppelt wird, ist der Quotient aus Energieänderung und Entropieänderung gerade die spektrale Temperatur.
- Verschiebt sich eine relativ scharf gepeakte Energiewahrscheinlichkeitsverteilung innerhalb des Spektrums eines modularen Systems infinitesimal, bleibt aber dabei gepeakt, so ist der Quotient aus der damit einhergehenden Energieänderung und der Entropieänderung gerade die spektrale Temperatur.

Damit weißt die spektrale Temperatur, die immer definiert ist, für viele typisch thermodynamische Situationen gerade die Eigenschaften auf, die man axiomatisch von der Temperatur fordert.

13.7 Parametrischer Druck

Ist für ein System lokal ein Hamiltonoperator definiert, der sich zeitlich verändern kann, und lässt sich diese Veränderung über einen Parameter erfassen, ist also abstrakt

$$\hat{H}^g(a) = \sum_n |n(a)\rangle E(a) \langle n(a)| \quad (13.9)$$

mit $a = a(t)$, so gilt für die zeitliche Entwicklung des Systems

$$\hat{\rho}^g(t) = \sum_n W_n(0) |n(t)\rangle \langle n(t)| \quad (13.10)$$

dann, wenn die zeitliche Veränderung des Hamiltonoperators, langsam ist, gegen das, durch (mikrokanonische) Ankopplung an die Umgebung verursachte, Anwachsen der

Entropie. Dieses Verhalten wird als “adiabatic following” bezeichnet und tritt, z.B. im Bereich der Quantenoptik auch ohne jede Kopplung an eine Umgebung auf. Die genauen Bedingungen für dieses Verhalten sind aber in komplizierteren Systemen schwer abzuschätzen. Es kann aber gezeigt werden, dass eine Ankopplung an die Umgebung ein adiabatisches Folgen stark stabilisieren kann, d.h. es wird z.B. auch dann noch auftreten wenn die Veränderung ohne die Kopplung schon zu schnell wäre. Das adiabatische Folgen ist ein Verhalten das effektiv ausschließlich durch die Parameterveränderung kontrolliert wird, und das die Entropie erhält. Es entspricht einer adiabatischen Zustandsänderung in der Thermodynamik, die ja den Druck definiert. Daher kann man den parametrischen Druck definieren als

$$\frac{\partial \text{Tr} \left\{ \hat{\rho}^g \hat{H}^g \right\}}{\partial a} = -p \quad (13.11)$$

Dieses Prinzip gilt zunächst für jede Parametervariation, die Verbindung zum “normalen” Druck ist aber natürlich nur gegeben wenn $a = V$, also wenn der variierte Parameter dem Volumen entspricht

Für diesen Fall kann auch gezeigt werden dass ein Gleichgewichtsverhalten in dem Sinne, dass der Volumenparameter sich nicht weiter verändern wird, für zwei Subsysteme die miteinander Volumen austauschen können (verschiebbare Wand), nur dann verwirklicht werden kann, wenn der parametrische Druck in beiden Subsystemen gleich ist.

13.8 Klassische und quantenmechanische Zustandsdichte

Definiert man als klassische Zustandsdichte das Volumen einer Energieschale im Phasenraum geteilt durch \hbar^{3N} , so gibt es eine große Klasse von (Hamilton)Modellen, für die auf zwei vollkommen unterschiedliche Arten zwei verschiedene Zustandsdichten berechnet werden können, die klassische und die quantenmechanische. Daher ist es schon an und für sich interessant zu untersuchen ob für ein und dasselbe Modell diese Zustandsdichten in allgemeinen gleich oder unterschiedlich sein werden. Darüberhinaus dienen Zustandsdichten, gleich welcher Natur, als Basis für Entropiedefinitionen und diese wiederum sind thermodynamische Potentiale. Wären also die Zustandsdichten unterschiedlich, so gäbe es für ein und dasselbe System verschiedene Prognosen seines konkreten thermodynamischen Verhaltens, je nachdem ob eine klassische oder eine quantenmechanische Betrachtungsweise zugrunde gelegt wird.

Die Untersuchungsmethode die hier verwendet wurde um die Verbindung der Zustandsdichten im allgemeinen zu klären, besteht im wesentlichen darin das quantenmechanische Spektrum aus den Spektren einzelner Wellenpakete im entsprechenden Modell zusammenzusetzen. Verwendet werden Gauß'sche Wellenpakete von minimaler Orts-Impuls Unschärfe. Diese erlauben eine Zuordnung zu Punkten im Phasenraum so dass hier eine Verbindung zu den klassischen Energieschalen hergestellt werden kann.

Das Ergebnis dieser Betrachtung ist, dass sich näherungsweise in einem Energieintervall,

das groß ist gegen die Energieunschärfe des Gauß'schen Wellenpaketes, so viele quantenmechanische Energieeigenzustände befinden wie Volumenelemente der Größe \hbar^{3N} in dem entsprechenden, im selben Energieintervall liegenden Volumen im Phasenraum. Ist also die Energieunschärfe eines solchen Wellenpaketes klein, insbesondere gegen seine mittlere Energie, so sind quantenmechanische und klassische Zustandsdicht näherungsweise gleich.

Zu den Modellen für die eine solche Gleichheit erwartet werden kann, gehören daher alle Modelle mit großen Massen und "glatten" Potentialen, sowie, für die Thermodynamik wichtiger, modulare Modelle die aus vielen identischen, schwach gekoppelten Untereinheiten bestehen.

13.9 Wege zum Gleichgewicht

Bis hierher bezogen sich alle Aussagen auf das Verhalten im Gleichgewicht. Offen ist daher die Frage was für ein Verhalten auf dem Weg in das Gleichgewicht zu erwarten ist. Für quantenmechanische Modelle führt dort meistens die Annahme statistischer Übergangsprozesse zu einem exponentiellen Zerfall angeregter Zustände. Diese statistische Verhalten ist aber im allgemeinen nicht durch die Schrödingergleichung erklärt.

Es wurde gezeigt dass mit Hilfe zeitabhängiger Störungsrechnung und der Theorie der Hilbertraummittel für bestimmte "Systemszenarien" dieser exponentielle Zerfall letztlich aus der Schrödingergleichung gefolgert werden kann. Diese Szenarien beinhalten wiederum dass das System mit dem zerfallenden Gleichgewichtszustand an ein anderes, mit möglichst hoher Zustandsdichte schwach gekoppelt ist. Es gibt noch einige weitere Kriterien die im wesentlichen mit denen zur Anwendbarkeit von "Fermis goldener Regel" identisch sind. Wiederum zeigen Simulationen **11.1.5** eine gute Übereinstimmung mit der Theorie für Modelle die im Rahmen ihrer Anwendbarkeit liegen.

13.10 Schlussbetrachtung und Ausblick

Die in dieser Arbeit angestellten Überlegungen haben gezeigt dass der Einfluss einer schwach gekoppelten Umgebung ein Quantensystem dazu bringen kann, thermodynamisches Verhalten zu zeigen. Die Umgebungseinflüsse sind dabei keineswegs statistisch oder unkontrolliert, und das gesamte, zusammengesetzte System entwickelt sich streng deterministisch nach der Schrödingergleichung. In diesem Zugang sind alle thermodynamischen Größen streng als Funktionen des Mikrozustandes definiert und die Emergenz thermodynamischer Eigenschaften folgt nicht aus einer Ergodenhypothese oder aus dem subjektiven Informationsmangel eines Beobachters. Das Auftreten thermodynamischen Verhaltens ist hier nicht von einer Mindestgröße des Systems abhängig, bei einer geeigneten Kopplung an eine geeignete Umgebung können einzelne, beliebig kleine Quantenobjekte thermodynamisches Verhalten zeigen. Dieser Aspekt gibt Anlass zu der Hoffnung, dass die Theorie helfen könnte, mögliche Experimente an den Grenzen zwischen Quantenmechanik und Thermodynamik besser zu verstehen. Solche Experimente könnten sicherlich auch zur besseren Beurteilung der Relevanz dieser Theorie insgesamt

beitragen.

Appendix

A Minimization of the effective Interaction, eq.5.24

The functional to be minimized is:

$$\begin{aligned} \langle \psi | (\hat{I} - \hat{V}^g)^2 | \psi \rangle = & \quad (A.1) \\ \int \left(\hat{I}^2(\{\vec{x}_i^g\}, \{\vec{x}_j^c\}) - 2\hat{I}(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})\hat{V}^g(\{\vec{x}_i^g\}) + (\hat{V}^g)^2(\{\vec{x}_i^g\}) \right) |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \\ \prod_j d\vec{x}_j^c \prod_i d\vec{x}_i^g \end{aligned}$$

Since the minimalization has to be done with respect to \hat{V}^g which only depends on the gas variables $\{\vec{x}_i^g\}$, the Euler-Lagrange equation reads:

$$-2 \int \hat{I}(\{\vec{x}_i^g\}, \{\vec{x}_j^c\}) |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \prod_j d\vec{x}_j^c + 2\hat{V}^g(\{\vec{x}_i^g\}) \int |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \prod_j d\vec{x}_j^c = 0 \quad (A.2)$$

from which we find, solving for \hat{V}^g

$$\hat{V}^g(\{\vec{x}_i^g\}) = \frac{\int \hat{I}(\{\vec{x}_i^g\}, \{\vec{x}_j^c\}) |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \prod_j d\vec{x}_j^c}{\int |\psi(\{\vec{x}_i^g\}, \{\vec{x}_j^c\})|^2 \prod_j d\vec{x}_j^c}. \quad (A.3)$$

B Conserved Probabilities, eq.5.28

From

$$\left[\hat{H}^g, \hat{H} \right] = 0 \quad \text{follows} \quad \left[(\hat{H}^g)^n, \hat{H} \right] = 0 \quad (B.1)$$

If we expand \hat{H}^g in terms of projectors $\hat{\Pi}_A^g$, projecting out the different energy eigenspaces of the gas system, such that

$$\sum_A E_A^g \hat{\Pi}_A^g = \hat{H}^g, \quad (B.2)$$

inserting (B.2) into (B.1) yields

$$\sum_A (E_A^g)^n \left[\hat{\Pi}_A^g, \hat{H} \right] = 0. \quad (B.3)$$

Since with different n infinitely many linearly independent equations can be generated we can conclude that the projectors have to commute separately:

$$\left[\hat{\Pi}_A^g, \hat{H} \right] = 0. \quad (B.4)$$

The corresponding considerations are of course true for the container system as well, such that

$$\left[\hat{\Pi}_B^c, \hat{H} \right] = 0. \quad (B.5)$$

Since $\hat{\Pi}_A^g$ and $\hat{\Pi}_B^c$ commute anyway (they act on different subspaces) we can conclude that

$$\left[\hat{\Pi}_A^g \hat{\Pi}_B^c, \hat{H} \right] = 0. \quad (\text{B.6})$$

which means that $\langle \hat{\Pi}_A^g \hat{\Pi}_B^c \rangle$ is a conserved quantity. Writing out $\hat{\Pi}_A^g$ and $\hat{\Pi}_B^c$ as

$$\hat{\Pi}_A^g = \sum_a |A, a\rangle \langle A, a|, \quad \hat{\Pi}_B^c = \sum_b |B, b\rangle \langle B, b| \quad (\text{B.7})$$

we get

$$\langle \psi(t) | \hat{\Pi}_A^g \hat{\Pi}_B^c | \psi(t) \rangle = \sum_{a,b} |\psi_{ab}^{AB}(t)|^2 = \sum_{a,b} |\psi_{ab}^{AB}(0)|^2 \quad (\text{B.8})$$

C Averages over Hyperspheres, sect. 5.4.3:

The integrals that are to be evaluated 5.4.3 are essentially of the form:

$$Z = \int z_l^{u_l}(\{\phi_n\}) z_m^{u_m}(\{\phi_n\}) \det \mathcal{F}(\{\phi_n\}) \prod_n^{d-1} d\phi_n \quad (\text{C.1})$$

where the u are integers and the z are parametrized by the ϕ to lie on a hypersphere of radius R and dimension d . (The z_l correspond to the real and imaginary parts of the amplitudes, respectively, d is the degree of degeneracy of a given subspace and R^2 the probability of this subspace to be occupied.) The z_s are related to cartesian coordinates x_s by:

$$\frac{z_s(\{\phi_n\})r}{R} = x_s \quad (\text{C.2})$$

where r is a radial variable. The technique to solve this integral, Z , is basically the same as used to calculate the surface area of a hypersphere of arbitrary dimension. This surface area is also the special case of Z for $u_l = u_m = 0$.

Defining

$$Z_1 := \int_0^\infty e^{-r^2} \left(\frac{r}{R}\right)^{u_l u_m + d - 1} dr \quad (\text{C.3})$$

the product $Z_2 = Z Z_1$ may be written as:

$$Z_2 := \int_0^\infty e^{-r^2} \left(\frac{z_l r}{R}\right)^{u_l} \left(\frac{z_m r}{R}\right)^{u_m} \left(\frac{r}{R}\right)^{d-1} \det \mathcal{F} \prod_n^{d-1} d\phi_n dr \quad (\text{C.4})$$

Now, this is an integral over all space, written in angular and radial coordinates that can be converted to cartesian coordinates yielding:

$$Z_2 = \int e^{-\sum_s x_s^2} x_l^{u_l} x_m^{u_m} \prod_s^d dx_s \quad (\text{C.5})$$

Since this integral factorizes completely, both Z_2 and Z_1 can be evaluated using standard tables of integrals, and depend only on R, d, u_l, u_m . (Note that Z_2 vanishes if any of the u is odd.) Z is then found to be

$$Z(R, d, u_l, u_m) = \frac{Z_2(R, d, u_l, u_m)}{Z_1(R, d, u_l, u_m)} \quad (\text{C.6})$$

The averages that are to be computed in **5.4.3** are of the form:

$$A(R, d, u_l, u_m) := \frac{Z(R, d, u_l, u_m)}{Z(R, d, 0, 0)} \quad (\text{C.7})$$

They are all invariant with respect to exchange of the u 's. Here we need only:

$$A(R, d, 0, 1) = A(R, d, 1, 1) = 0, \quad A(R, d, 0, 2) = \frac{R^2}{d} \quad (\text{C.8})$$

which could also have been found from simple symmetry considerations, and

$$A(R, d, 2, 2) = \frac{R^4 \Gamma(\frac{d}{2})}{4\Gamma(\frac{d}{2} + 2)} \quad A(R, d, 0, 4) = \frac{3R^4 \Gamma(\frac{d}{2})}{4\Gamma(\frac{d}{2} + 2)} \quad (\text{C.9})$$

All A 's are invariant with respect to exchange of the u 's.

D From Multiplication to Gaussian, eq. 6.16

To show that any function with a global maximum that is multiplied N times with itself will essentially be a Gaussian, we consider a function $f(x)$ with a global maximum at $x = 0$. Such a function can be written as:

$$f(x) = Ae^{-Bx^2} e^{Cx^3} \dots \quad (\text{D.1})$$

Since multiplying the function with itself will amplify the maximum in the center extremely we can truncate the decomposition in this way. We then get:

$$f(x)^N = A^N e^{-NBx^2} e^{NCx^3} \dots \quad (\text{D.2})$$

The value of x for which the quadratic part will have reduced the function to half maximum, i.e., for which

$$e^{-NBx_H^2} = \frac{1}{2} \quad (\text{D.3})$$

is

$$x_H = \sqrt{\frac{-\ln \frac{1}{2}}{NB}} \quad (\text{D.4})$$

Evaluating the third order part of $f(x)^N$ at x_H yields

$$e^{NCx_H^3} = e^{\frac{c\left(\frac{-\ln \frac{1}{2}}{B}\right)^{\frac{3}{2}}}{\sqrt{N}}} \quad (\text{D.5})$$

which tends to 1 if N approaches infinity. We thus conclude that for the relevant region in which the function is peaked

$$f(x)^N \approx A^N e^{-NBx^2} \quad (\text{D.6})$$

that is, essentially Gaussian.

E Entropy of an ideal Gas, eq. 6.23

The total energy of a classical gas depends $3N$ degrees of freedom, corresponding to the components of the momenta of all N particles. Thus its spectrum can be calculated as the $3N$ -fold convolution of the spectrum of an individual degree of freedom.

From the dispersion relation of a classical free particle confined to one dimension

$$E = \frac{1}{2M}p^2 \quad (\text{E.1})$$

we find

$$\frac{dp}{dE} = \frac{M}{p} = \sqrt{\frac{M}{2E}} \quad (\text{E.2})$$

Since there are two momenta corresponding to one energy, we get, taking h as the the volume of an elementary cell, for a particle restricted to the length L , the state density:

$$g(E) = \frac{L}{h} \sqrt{\frac{2M}{E}} \quad (\text{E.3})$$

With this state density we find, using a standard table of integrals, for the quantities defined in 6.2, A and m

$$A = \frac{L}{h} \sqrt{2M} \frac{\Gamma\left(\frac{1}{2}\right)}{\alpha^{\frac{1}{2}}}, \quad m = \frac{1}{2\alpha} \quad (\text{E.4})$$

Setting $m = \left(\frac{U}{N'}\right)$ and writing α and A as functions of this argument we get:

$$\alpha = \frac{1}{2\left(\frac{U}{N'}\right)}, \quad A = \frac{L}{h} \left(4M\pi \left(\frac{U}{N'}\right)\right)^{\frac{1}{2}} \quad (\text{E.5})$$

Inserting this into (6.23) yields:

$$\ln G(U) = N' \left(\ln \frac{L}{h} + \frac{1}{2} \ln \left(4M\pi \left(\frac{U}{N'}\right)\right) + \frac{1}{2} \right) \quad (\text{E.6})$$

Connecting the number of degrees of freedom N' with the number of particles N as $N' = 3N$, we eventually find:

$$\ln G(U) = N \left(3 \ln \frac{L}{h} + \frac{3}{2} \ln \left(\frac{4}{3} M \pi \left(\frac{U}{N} \right) \right) + \frac{3}{2} \right) \quad (\text{E.7})$$

which is exactly the same result which is in standard textbooks usually obtained by using completely different methods.

F Stabilized adiabatic Approximation, eq. 8.18

Taking the derivative of $\rho^g(t)$ according to (8.17) reads:

$$\frac{\partial}{\partial t} \hat{\rho}^g = \sum_n W_n \left(\left| \frac{\partial}{\partial t} n \right\rangle \langle n| + |n\rangle \left\langle \frac{\partial}{\partial t} n \right| \right) + \frac{\partial r}{\partial t} \hat{\epsilon} + r \frac{\partial}{\partial t} \hat{\epsilon} \quad (\text{F.1})$$

Inserting this into (8.15) we find:

$$i\hbar \left(\sum_n W_n \left(\left| \frac{\partial}{\partial t} n \right\rangle \langle n| + |n\rangle \left\langle \frac{\partial}{\partial t} n \right| \right) + \frac{\partial r}{\partial t} \hat{\epsilon} + r \frac{\partial}{\partial t} \hat{\epsilon} \right) = \left[\hat{H}^g, r\hat{\epsilon} \right] - i\hbar \sum_{n,n'} |n\rangle \langle n'| \langle n|r\hat{\epsilon}|n'\rangle C_{n,n'} \quad (\text{F.2})$$

Realizing that

$$\text{Tr}\{\hat{\epsilon}^2\} = 1, \quad \text{Tr}\left\{\hat{\epsilon} \frac{\partial r}{\partial t} \hat{\epsilon}\right\} = \frac{1}{2} \frac{\partial}{\partial t} \text{Tr}\{\hat{\epsilon}^2\} = 0 \quad \text{Tr}\left\{\hat{\epsilon} \left[\hat{H}^g, r\hat{\epsilon} \right]\right\} = 0 \quad (\text{F.3})$$

we get, multiplying (F.2) from the left by $\hat{\epsilon}$, taking the trace and solving for $\frac{\partial r}{\partial t}$

$$\frac{\partial r}{\partial t} = -r \sum_{n,n'} |\langle n|\hat{\epsilon}|n'\rangle|^2 C_{n,n'} - \sum_n W_n \left(\langle n|\hat{\epsilon} \left| \frac{\partial}{\partial t} n \right\rangle + \left\langle \frac{\partial}{\partial t} n \right| \hat{\epsilon} |n\rangle \right) \quad (\text{F.4})$$

which is exactly (8.18).

G Sizes of Hypersphereregions, sect. 5.6

Here we want to explain how to find the size of a zone on a hypersphere of dimension $N_1 + N_2$ and radius 1 that consists only of points for which the sum of the squares of (the first) N_1 coordinates add up to equal W .

We start by parametrizing the full space that the hypersphere lives in, in the following way:

$$\vec{Z}(r_1, r_2, \vec{\phi}, \vec{\theta}) = \left(r_1 \vec{X}(\vec{\phi}), r_2 \vec{Y}(\vec{\theta}) \right) \quad (\text{G.1})$$

where the $\vec{X}(\vec{\phi}), \vec{Y}(\vec{\theta})$ are vectors with N_1, N_2 coordinates, parametrized by the $N_1 - 1, N_2 - 1$ parameters denoted as $\vec{\phi}, \vec{\theta}$ to lay on hyperspheres of radii 1 themselves.

If we now set up a volume integral, such as to integrate the volume of the full space, we have to compute the functional matrix. Due to the special parameterization chosen here, this functional matrix will have diagonal block form. The upper left block essentially consists of the functional submatrix $\frac{\partial \vec{X}}{\partial \vec{\phi}}$, the lower right block of the submatrix $\frac{\partial \vec{Y}}{\partial \vec{\theta}}$. Those matrices result from the parametrization of the first N_1 coordinates of \vec{Z} to lay on a hypersphere of radius r_1 and the last N_2 coordinates to lay on a hypersphere of radius r_2 . Thus an integration of the determinants of those matrices over $\vec{\phi}, \vec{\theta}$ results in quantities proportional to $r_1^{N_1-1}, r_2^{N_2-1}$ since those are simply the surface areas of the N_1, N_2 dimensional hyperspheres of radii r_1, r_2 . The determinant of the full functional matrix will, due to its having block form, be just the product of the determinants of the submatrices. Thus integrating the determinant of the full matrix over $\vec{\phi}$ and $\vec{\theta}$ we get:

$$V \propto \int r_1^{N_1-1} r_2^{N_2-1} dr_1 dr_2 \quad (\text{G.2})$$

(From here we only write proportionalities rather than equalities since factors that do not depend on the crucial variables are of no further interest here.) in order to be able to change from an integral over full space to an integral over the surface of the $N_1 + N_2$ dimensional hypersphere we transform the remaining variables according to:

$$r_1 = r \cos(\phi) \quad r_2 = r \sin(\phi) \quad (\text{G.3})$$

thus getting

$$V \propto \int \cos^{N_1-1}(\phi) \sin^{N_2-1}(\phi) r^{N_1+N_2-1} dr d\phi \quad (\text{G.4})$$

Since in those variables the radius of the full sphere $|\vec{Z}|$ simply equals r we may, to get the surface integral, simply introduce a delta-function, $\delta(r - 1)$, into the integrand and integrate over r , finding

$$A \propto \int \cos^{N_1-1}(\phi) \sin^{N_2-1}(\phi) d\phi \quad (\text{G.5})$$

Since the crucial quantity, the sum of the squares of the first N_1 coordinates, W now equals $\cos^2(\phi)$, and it is this W in terms of which we want to write the integral, we have to do yet another variable transformation:

$$W = \cos^2(\phi) \quad dW = -2 \cos(\phi) \sin(\phi) d\phi = -2\sqrt{W(1-W)} d\phi \quad (\text{G.6})$$

yielding

$$A \propto \int \sqrt{W}^{N_1-2} \sqrt{(1-W)}^{N_2-2} dW \quad (\text{G.7})$$

Now the surface area of the hypersphere is eventually given by an integration over only one variable which is exactly the quantity we want to connect with the sizes of the corresponding zones. That means the integrand gives those relative sizes. Since we are

eventually interested in the relative frequency of points featuring some W among all points on the surface of the hypersphere, $f(W)$, we have to normalize A to equal 1. This can be done using either a standard table of integrals and a lot of self discipline, or some standard software for symbolic computation finding:

$$f(W) = \frac{1}{C} \sqrt{W}^{N_1-2} \sqrt{(1-W)}^{N_2-2} \quad C = \frac{\Gamma(\frac{N_1}{2})\Gamma(\frac{N_2}{2})}{\Gamma(\frac{N_1+N_2}{2})} \quad (\text{G.8})$$

This guaranteeing that

$$\int_0^1 f(W)dW = 1 \quad (\text{G.9})$$

Obviously the function is 0 at $W = 0$ and $W = 1$ and peaked somewhere in between. The mean value is

$$\langle W \rangle := \int_0^1 W f(W)dW = \frac{N_1}{N_1 + N_2} \quad (\text{G.10})$$

as is in perfect agreement with (C.8). The maximum of this function is ant

$$W_{max} = \frac{N_1 - 2}{N_1 + N_2 - 4} \quad (\text{G.11})$$

thus the maximum converges against the mean value for large N_1, N_2 . The variance of the function $f(w)$ is

$$\Delta W := \sqrt{\int_0^1 W^2 f(W)dW - \langle W \rangle^2} = \sqrt{\frac{2N_1 N_2}{(N_1 + N_2)^2(N_1 + N_2 + 2)}} \quad (\text{G.12})$$

which means that the peak will be the sharper, the bigger N_1, N_2 are. Especially if both, N_1, N_2 , scale with the same factor A , the variance will scale as:

$$\Delta W = \sqrt{\frac{1}{A}} \quad (\text{G.13})$$

H Hilbertspace Averages, sect. 11.1.4

To calculate $\langle \alpha \rangle$ we write the quantity out in detailed tensor notation:

$$\langle \alpha \rangle = \langle \langle \psi_u | \hat{U} | \psi_o \rangle \rangle = \langle \sum_{i,j} \psi_i^{u*} U_{ij} \psi_j^o \rangle = \sum_{i,j} \langle \psi_i^{u*} \psi_j^o \rangle U_{ij} \quad (\text{H.1})$$

Since $|\psi_o\rangle, |\psi_u\rangle$ refer to entirely different hyperspheres we get, using the results from 2.7

$$\langle \psi_i^{u*} \psi_j^o \rangle = \langle \psi_i^{u*} \rangle \langle \psi_j^o \rangle = 0 \quad (\text{H.2})$$

and thus

$$\langle \alpha \rangle = 0 \quad (\text{H.3})$$

For $\langle \beta \rangle$ we proceed in a very similar way.

$$\langle \beta \rangle = \langle \langle \psi_u | \hat{U}^2 | \psi_u \rangle \rangle = \langle \sum_{i,j} \psi_i^{u*} U_{ij}^2 \psi_j^u \rangle = \sum_{i,j} \langle \psi_i^{u*} \psi_j^u \rangle U_{ij}^2 \quad (\text{H.4})$$

From **C** we find

$$\langle \psi_i^{u*} \psi_j^u \rangle = \frac{\langle \psi_u | \psi_u \rangle}{N_u} \delta_{ij} \quad (\text{H.5})$$

and thus

$$\langle \beta \rangle = \frac{\langle \psi_u | \psi_u \rangle}{N_u} \sum_i U_{ii}^2 = \frac{\langle \psi_u | \psi_u \rangle}{N_u} \text{Tr}^u \{ \hat{U}^2 \} \quad (\text{H.6})$$

In the same way $\langle \gamma \rangle$ is evaluated.

$\langle \alpha^2 \rangle$ reads in tensor notation

$$\begin{aligned} \langle \alpha^2 \rangle &= \langle \langle \psi_u | \hat{U} | \psi_o \rangle^2 \rangle = \langle \sum_{i,j,k,l} \psi_i^{u*} U_{ij} \psi_j^o \psi_k^{u*} U_{kl} \psi_l^o \rangle \\ &= \sum_{i,j,k,l} \langle \psi_i^{u*} \psi_k^{u*} \rangle \langle \psi_l^o \psi_j^o \rangle U_{ij} U_{kl} \end{aligned} \quad (\text{H.7})$$

To evaluate those averages we write them out in even more detail,

$$\langle \psi_l^o \psi_j^o \rangle = \langle X_l^o X_j^o + i(X_l^o Y_j^o + Y_l^o X_j^o) - Y_l^o Y_j^o \rangle = \frac{\delta_{lj}}{2N_u} - 0 - \frac{\delta_{lj}}{2N_u} = 0 \quad (\text{H.8})$$

where X, Y are the real and imaginary parts of the ψ . (The same applies to the average over the lower hypersphere). Thus

$$\langle \alpha^2 \rangle = 0 \quad \langle \alpha^{*2} \rangle = \langle \alpha^2 \rangle^* = 0 \quad (\text{H.9})$$

To evaluate $\langle |\psi_o\rangle \langle \psi_o| \rangle$ we again use tensor notation

$$\langle |\psi_o\rangle \langle \psi_o| \rangle = \langle \sum_{ij} \psi_i^o \psi_j^{o*} |i\rangle \langle j| \rangle = \langle \sum_{ij} \frac{W_o}{N_o} \delta_{ij} |i\rangle \langle j| \rangle = \frac{W_o}{N_o} \hat{1}^o \quad (\text{H.10})$$

where $\hat{1}^o$ stands for a projector, projecting out the upper subspace.

To compute the Hilbertspace variance of the expectation value of some hermitian operator \hat{A} , $\Delta \langle \psi | \hat{A} | \psi \rangle$,

$$\Delta \langle \psi | \hat{A} | \psi \rangle = \sqrt{\langle \langle \psi | \hat{A} | \psi \rangle^2 \rangle - \langle \langle \psi | \hat{A} | \psi \rangle \rangle^2} \quad (\text{H.11})$$

we use the eigenbasis of \hat{A} , A_i being the eigenvalues of \hat{A}

$$\Delta^2 \langle \psi | \hat{A} | \psi \rangle = \langle \sum_{i,j} |\psi_i|^2 A_i |\psi_j|^2 A_j \rangle - \left(\frac{1}{N} \sum_i A_i \right)^2 \quad (\text{H.12})$$

where we used (H.6) to infer the second term. To evaluate the first term we find either from **C** or from comparing with (5.40) and (5.44)

$$\langle |\psi_i|^2 |\psi_j|^2 \rangle = \frac{1 + \delta_{ij}}{N(N+1)} \quad (\text{H.13})$$

thus getting

$$\langle \sum_{i,j} |\psi_i|^2 A_i |\psi_j|^2 A_j \rangle = \frac{1}{N(N+1)} \left(\left(\sum_i A_i \right)^2 + \sum_i A_i^2 \right) \quad (\text{H.14})$$

putting it all together yields

$$\Delta^2 \langle \psi | \hat{A} | \psi \rangle = \frac{1}{N(N+1)} \sum_i A_i^2 + \left(\frac{1}{N(N+1)} - \frac{1}{N} \right) \left(\sum_i A_i \right)^2 \quad (\text{H.15})$$

$$= \frac{1}{N+1} \left(\frac{\text{Tr} \{ \hat{A}^2 \}}{N} - \left(\frac{\text{Tr} \{ \hat{A} \}}{N} \right)^2 \right) \quad (\text{H.16})$$

thus, the Hilbertspace variance squared is simply the spectral variance squared, divided by the dimension plus one.

Bibliography

- [1] Boltzmann, L. *Lectures on Gas theory 1896-1889* (University of California Press, Los Angeles, 1964)
- [2] Gibbs, W. *Elementary Principles in Statistical Mechanics* (Yale University Press, New Haven, 1974)
- [3] Birkhoff, G., Proc. Nat. Acad., **17**, 650, (1931).
- [4] Ehrenfest, P., Ehrenfest, T. *The Conceptual Foundations of the Statistical Approach in Mechanics* (Cornell University Press, Ithaca, 1959)
- [5] Von Neumann, J., Proc. Nat. Acad., **18**, 70, (1932).
- [6] Ben-Menahem, Y., Pitowsky, I., J. Stud. Hist. Phil. Mod. Phys. **32**, 503, (2001).
- [7] Von Neumann, J., Z. Phys., **57**, 30, (1930).
- [8] Pauli, W., Fierz, Z., Z. Phys., **106**, 572, (1937).
- [9] Landau, L., Lifshitz, I., *Statistical Mechanics* (Pergamon Press, Oxford, 1978); *Quantum Mechanics* (Pergamon Press, Oxford, 1977)
- [10] Schrödinger, E., *Statistical Thermodynamics* (Dover, New York, 1989)
- [11] Lindblad, G. *Non-equilibrium entropy and Irreversibility* (D. Reidel Publishing Company, Dordrecht, 1983)
- [12] Zurek, W., Phys. Today, **44**, 36, (1991)
- [13] Zurek, W., Paz, J., Phys. Rev. Lett., **72**, 2508, (1994)
- [14] Theis, W. *Grundzüge der Quantentheorie* (Teubner, Stuttgart, 1985)
- [15] Davies, P. *The Physics of time Asymmetry* (University of California Press, Los Angeles, 1974)
- [16] Penrose, R. *The Emperor's New Mind* (Oxford University Press, New York, 1990)
- [17] Münster, A. *Statistische Thermodynamik* (Springer Verlag, Berlin, 1956)
- [18] Jaynes, E., Phys. Rev. **106**, 620, (1957)

-
- [19] von Neumann, J., *Die mathematischen Grundlagen der Quantenmechanik* (Springer Verlag, Berlin, 1932)
- [20] Fick, E., Sauermann, G., *Quantenstatistik dynamischer Prozesse* (Verlag Harri Deutsch Thun, Frankfurt a. Main, 1983)
- [21] Davies, E., *Quantum Theory of Open Systems* (Academic Press, London, 1976)
- [22] Weiss, U., *Quantum Dissipative Systems* (World Scientific, Singapore, 1999)
- [23] Mahler, G. *Quantum Networks* (Springer Verlag, Berlin, 1995)
- [24] Gemmer, J., Mahler, G., Eur. Phys. J. D, **17**, 385, (2001)
- [25] Einstein, A., Podolsky, B., Rosen, N., Phys. Rev. **47**, 777, (1935)
- [26] Ballantine, L., *Quantum Mechanics* (World Scientific, Singapore, 1999)
- [27] Blum, K., *Density matrix. Theory and applications* (World Scientific, Singapore, 1999)
- [28] Tsallis, C., J. Stat. Phys. **52**, 479, (1988)
- [29] Diu, B., Guthmann, C., Lederer, D., Roulet, B., *Elements de Physique Statistique* (Hermann Editeurs des Sciences et des Arts, Paris, 1989)
- [30] Allahverdyan, A., Nieuwenhuizen, T., Phys. Rev. Lett., **85**, 1799, (2000)
- [31] Gemmer, J., Mahler, G., Europhys. Lett. **59**, 159, (2002)
- [32] Bergmann, P., Lebowitz, J., **99**, 578, (1955)
- [33] Blatt, J., Progr. Theor. Phys., **22**, 745, (1959)
- [34] Lubkin, E., Math. Phys., **19**, 1028, (1978)
- [35] Tasaki, H., Phys. Rev. Lett., **80**, 1373, (1998)
- [36] Jensen, R., Shankar, R., Phys. Rev. Lett., **54**, 1879, (1985)
- [37] Bocchieri, P., Loigner, A., Phys. Rev., **114**, 948, (1959)
- [38] Scarani, V., Phys. Rev. Lett., **88**, 097905, (2002)
- [39] Brenig, W., *Statistische Theorie der Wärme* (Springer, Berlin, 1996)
- [40] Toda, M., Kobo, R., Saito, N., *Statistical Physics I* (Springer, Berlin, 1978)
- [41] Allahverdyan, A., Nieuwenhuizen, T., Phys. Rev. Lett., **85**, 232, (2000)
- [42] Scully, M., Phys. Rev. Lett., **85**, 1799, (2000)

- [43] Rugh, H., Phys. Rev. Lett., **78**, 722, (1997)
- [44] Yckayzen, G., J. Chem. Phys., **114**, 4333 (2001)
- [45] Pick, Spaeth, Paus, *Kursvorlesung Experimentalphysik I* (Universität Stuttgart, 1969)
- [46] Ehrenfest, P., J., Ann. d. Phys., **51** 327 (1916)
- [47] Schiff, L., *Quantum Mechanics* (McGraw-Hill, Tokyo, 1955)
- [48] Römer, H., Filk, T., *Statistische Mechanik* (VCH, Weinheim, 1994)
- [49] Scully, M., Zubairy, M., *Quantum Optics* (Cambridge University Press, 1997)