Lecture Notes in Physics

The Editorial Policy for Monographs

The series Lecture Notes in Physics reports new developments in physical research and teaching - quickly, informally, and at a high level. The type of material considered for publication includes monographs presenting original research or new angles in a classical field. The timeliness of a manuscript is more important than its form, which may be preliminary or tentative. Manuscripts should be reasonably self-contained. They will often present not only results of the author(s) but also related work by other people and will provide sufficient motivation, examples, and applications.

Acceptance

The manuscripts or a detailed description thereof should be submitted either to one of the series editors or to the managing editor. The proposal is then carefully refereed. A final decision concerning publication can often only be made on the basis of the complete manuscript, but otherwise the editors will try to make a preliminary decision as definite as they can on the basis of the available information.

Contractual Aspects

Authors receive jointly 30 complimentary copies of their book. No royalty is paid on Lecture Notes in Physics volumes. But authors are entitled to purchase directly from Springer other books from Springer (excluding Hager and Landolt-Börnstein) at a $33\frac{1}{3}$% discount off the list price. Resale of such copies or of free copies is not permitted. Commitment to publish is made by a letter of interest rather than by signing a formal contract. Springer secures the copyright for each volume.

Manuscript Submission

Manuscripts should be no less than 100 and preferably no more than 400 pages in length. Final manuscripts should be in English. They should include a table of contents and an informative introduction accessible also to readers not particularly familiar with the topic treated. Authors are free to use the material in other publications. However, if extensive use is made elsewhere, the publisher should be informed. As a special service, we offer free of charge LaTeX macro packages to format the text according to Springer's quality requirements. We strongly recommend authors to make use of this offer, as the result will be a book of considerably improved technical quality. The books are hardbound, and quality paper appropriate to the needs of the author(s) is used. Publication time is about ten weeks. More than twenty years of experience guarantee authors the best possible service.

LNP Homepage (springerlink.com)

On the LNP homepage you will find:
—The LNP online archive. It contains the full texts (PDF) of all volumes published since 2000. Abstracts, table of contents and prefaces are accessible free of charge to everyone. Information about the availability of printed volumes can be obtained.
—The subscription information. The online archive is free of charge to all subscribers of the printed volumes.
—The editorial contacts, with respect to both scientific and technical matters.
—The author's / editor's instructions.

J. Gemmer M. Michel G. Mahler

Quantum Thermodynamics

Emergence of Thermodynamic Behavior Within Composite Quantum Systems

Springer

Authors

J. Gemmer
Universität Osnabrück
FB Physik
Barbarastr. 7
49069 Osnabrück, Germany

M. Michel
G. Mahler
Universität Stuttgart
Pfaffenwaldring 57
70550 Stuttgart, Germany

J. Gemmer M. Michel G. Mahler , *Quantum Thermodynamics*, Lect. Notes Phys. **657** (Springer, Berlin Heidelberg 2005), DOI 10.1007/b98082

Library of Congress Control Number: 2004110894

ISSN 0075-8450
ISBN 3-540-22911-6 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by the authors/editor
Data conversion: PTP-Berlin Protago-TeX-Production GmbH
Cover design: *design & production*, Heidelberg

Printed on acid-free paper
54/3141/ts - 5 4 3 2 1 0

Lecture Notes in Physics

For information about Vols. 1–609
please contact your bookseller or Springer
LNP Online archive: springerlink.com

Vol.611: A. Buchleitner, K. Hornberger (Eds.), Coherent Evolution in Noisy Environments.

Vol.612: L. Klein, (Ed.), Energy Conversion and Particle Acceleration in the Solar Corona.

Vol.613: K. Porsezian, V.C. Kuriakose (Eds.), Optical Solitons. Theoretical and Experimental Challenges.

Vol.614: E. Falgarone, T. Passot (Eds.), Turbulence and Magnetic Fields in Astrophysics.

Vol.615: J. Büchner, C.T. Dum, M. Scholer (Eds.), Space Plasma Simulation.

Vol.616: J. Trampetic, J. Wess (Eds.), Particle Physics in the New Millenium.

Vol.617: L. Fernández-Jambrina, L. M. González-Romero (Eds.), Current Trends in Relativistic Astrophysics, Theoretical, Numerical, Observational

Vol.618: M.D. Esposti, S. Graffi (Eds.), The Mathematical Aspects of Quantum Maps

Vol.619: H.M. Antia, A. Bhatnagar, P. Ulmschneider (Eds.), Lectures on Solar Physics

Vol.620: C. Fiolhais, F. Nogueira, M. Marques (Eds.), A Primer in Density Functional Theory

Vol.621: G. Rangarajan, M. Ding (Eds.), Processes with Long-Range Correlations

Vol.622: F. Benatti, R. Floreanini (Eds.), Irreversible Quantum Dynamics

Vol.623: M. Falcke, D. Malchow (Eds.), Understanding Calcium Dynamics, Experiments and Theory

Vol.624: T. Pöschel (Ed.), Granular Gas Dynamics

Vol.625: R. Pastor-Satorras, M. Rubi, A. Diaz-Guilera (Eds.), Statistical Mechanics of Complex Networks

Vol.626: G. Contopoulos, N. Voglis (Eds.), Galaxies and Chaos

Vol.627: S.G. Karshenboim, V.B. Smirnov (Eds.), Precision Physics of Simple Atomic Systems

Vol.628: R. Narayanan, D. Schwabe (Eds.), Interfacial Fluid Dynamics and Transport Processes

Vol.629: U.-G. Meißner, W. Plessas (Eds.), Lectures on Flavor Physics

Vol.630: T. Brandes, S. Kettemann (Eds.), Anderson Localization and Its Ramifications

Vol.631: D. J. W. Giulini, C. Kiefer, C. Lämmerzahl (Eds.), Quantum Gravity, From Theory to Experimental Search

Vol.632: A. M. Greco (Ed.), Direct and Inverse Methods in Nonlinear Evolution Equations

Vol.633: H.-T. Elze (Ed.), Decoherence and Entropy in Complex Systems, Based on Selected Lectures from DICE 2002

Vol.634: R. Haberlandt, D. Michel, A. Pöppl, R. Stannarius (Eds.), Molecules in Interaction with Surfaces and Interfaces

Vol.635: D. Alloin, W. Gieren (Eds.), Stellar Candles for the Extragalactic Distance Scale

Vol.636: R. Livi, A. Vulpiani (Eds.), The Kolmogorov Legacy in Physics, A Century of Turbulence and Complexity

Vol.637: I. Müller, P. Strehlow, Rubber and Rubber Balloons, Paradigms of Thermodynamics

Vol.638: Y. Kosmann-Schwarzbach, B. Grammaticos, K.M. Tamizhmani (Eds.), Integrability of Nonlinear Systems

Vol.639: G. Ripka, Dual Superconductor Models of Color Confinement

Vol.640: M. Karttunen, I. Vattulainen, A. Lukkarinen (Eds.), Novel Methods in Soft Matter Simulations

Vol.641: A. Lalazissis, P. Ring, D. Vretenar (Eds.), Extended Density Functionals in Nuclear Structure Physics

Vol.642: W. Hergert, A. Ernst, M. Däne (Eds.), Computational Materials Science

Vol.643: F. Strocchi, Symmetry Breaking

Vol.644: B. Grammaticos, Y. Kosmann-Schwarzbach, T. Tamizhmani (Eds.) Discrete Integrable Systems

Vol.645: U. Schollwöck, J. Richter, D.J.J. Farnell, R.F. Bishop (Eds.), Quantum Magnetism

Vol.646: N. Bretón, J. L. Cervantes-Cota, M. Salgado (Eds.), The Early Universe and Observational Cosmology

Vol.647: D. Blaschke, M. A. Ivanov, T. Mannel (Eds.), Heavy Quark Physics

Vol.648: S. G. Karshenboim, E. Peik (Eds.), Astrophysics, Clocks and Fundamental Constants

Vol.649: M. Paris, J. Rehacek (Eds.), Quantum State Estimation

Vol.650: E. Ben-Naim, H. Frauenfelder, Z. Toroczkai (Eds.), Complex Networks

Vol.651: J.S. Al-Khalili, E. Roeckl (Eds.), The Euroschool Lectures of Physics with Exotic Beams, Vol.I

Vol.652: J. Arias, M. Lozano (Eds.), Exotic Nuclear Physics

Vol.653: E. Papantonoupoulos (Ed.), The Physics of the Early Universe

Vol.654: G. Cassinelli, A. Levrero, E. de Vito, P. J. Lahti (Eds.), Theory and Appplication to the Galileo Group

Vol.655: M. Shillor, M. Sofonea, J.J. Telega, Models and Analysis of Quasistatic Contact

Vol.656: K. Scherer, H. Fichtner, B. Heber, U. Mall (Eds.), Space Weather

Vol.657: J. Gemmer, M. Michel, G. Mahler (Eds.), Quantum Thermodynamics

Preface

This monograph views thermodynamics as an *incomplete description* of many freedom quantum systems. Left unaccounted for may be an environment with which the system of interest interacts; closed systems can be described incompletely by focussing on any subsystem with fewer particles and declaring the remainder as the environment. Any interaction with the environment brings the open system to a mixed quantum state, even if the closed compound state is pure. Moreover, observables (and sometimes even the density operator) of an open system may relax to equilibrium values, while the closed compound state keeps evolving unitarily à la Schrödinger forever.

The view thus taken can hardly be controversial for our generation of physicists. And yet, the authors offer surprises. Approach to equilibrium, with equilibrium characterized by maximum ignorance about the open system of interest, does not require excessively many particles: some dozens suffice! Moreover, the precise way of partitioning which might reflect subjective choices is immaterial for the salient features of equilibrium and equilibration. And what is nicest, quantum effects are at work in bringing about universal thermodynamic behavior of modest size open systems. Von Neumann's concept of entropy thus appears as being much more widely useful than sometimes feared, way beyond truely macroscopic systems in equilibrium.

The authors have written numerous papers on their quantum view of thermodynamics, and the present monograph is a most welcome coherent review.

Essen, *Fritz Haake*
June 2004

Acknowledgements

The authors thank Dipl. Phys. Peter Borowski (MPI Dresden) for the first numerical simulations to test our theoretical considerations and for contributing several figures as well as some text material, and Dipl. Phys. Michael Hartmann (DLR Stuttgart) for contributing some sections. Furthermore, we thank Cand. Phys. Markus Henrich for helping us to design some diagrams and, both M. Henrich and Cand. Phys. Christos Kostoglou (Institut für theoretische Physik, Universität Stuttgart) for supplying us with numerical data. We have profited a lot from fruitful discussions with Dipl. Phys. Harry Schmidt, Dipl. Phys. Marcus Stollsteimer and Dipl. Phys. Friedemann Tonner (Institut für theoretische Physik, Universität Stuttgart) and Prof. Dr. Klaus Bärwinkel, Prof. Dr. Heinz-Jürgen Schmidt and Prof. Dr. Jürgen Schnack (Fachbereich Physik, Universität Osnabrück). We benefitted much from conversations with Prof. Dr. Wolfram Brenig and Dipl. Phys. Fabian Heidrich-Meisner (Technische Universität Braunschweig) as well as Dr. Alexander Otte and Dr. Heinrich Michel (Stuttgart). It is a pleasure to thank Springer-Verlag, especially Dr. Christian Caron, for continuous encouragement and excellent cooperation. This cooperation has garanteed a rapid and smooth progress of the project. Financial support by the “Deutsche Forschungsgesellschaft” and the “Landesstiftung Baden-Württemberg” is gratefully acknowledged. Last but not least, we would like to thank Björn Butscher, Kirsi Weber and Hendrik Weimer for helping us with typesetting the manuscript, proof-reading and preparing some of the figures.

Contents

Part I Background

Part II Quantum Approach to Thermodynamics

Part III Applications and Models

Part IV Appendices

List of Symbols

$\hat{1}$	Unit operator
$\hat{1}^{(\mu)}$	Unit operator in Hilbert space of subsystem μ
∇	Nabla operator
δ_{ij}	Kronecker delta
$\delta(\dots)$	Dirac δ-function
$\otimes$	Dyadic product
$\mathrm{d}A$	Infinitesimal change of an integrable quantity, complete differential
$\varDelta A$	Finite change of A
$\updelta A$	Infinitesimal change of a non-integrable quantity
$\mathcal{C}_N\{f\}$	Convolution of N identical functions f
$[\![f]\!]$	Hilbert space average of a quantity f
$\mathcal{O}(R,n)$	Surface of an n-dimensional hypersphere with radius R
$\big[\hat{A},\hat{B}\big]$	Commutator of $\hat{A}$ with $\hat{B}$
$\mathcal{F}\{f\}$	Fourier transformation of f
$\overline{f}$	Mean value of f
$\langle i\vert j\rangle$	Scalar product in Hilbert space
$\{\eta_i,\xi_i\}$	Set of all coordinates $i = 1, \dots n$
$\{\eta_{ab}^{AB},\xi_{ab}^{AB}\}$	Set of coordinates in subspace AB
$\{\eta_i^J,\xi_i^J\}$	Subset of coordinates in subspace J
$\{r,\phi_i\}$	Generalized spherical coordinates
$\{r^J,\phi_i^J\}$	Spherical coordinates of subspace J
$\{r^{AB},\phi_i^{AB}\}$	Spherical coordinates of subspace AB
$\langle\hat{A}\rangle$	Expectation value of operator $\hat{A}$
$\hat{A}$	Operator
$\hat{A}^{(\mu)}$	Operator in subspace μ
$\hat{A}^{\dagger}$	Adjoint operator
$\hat{A}_{\mathrm{H}}(t)$	Time-dependent operator in Heisenberg picture
$\hat{A}_{\mathrm{I}}$	Operator in the interaction picture
$\vert A,a\rangle$	Basis state of the gas system g
A	Index of the energy subspace of the gas system with energy E_A^{g}
a	Index of the degenerate eigenstates belonging to one energy subspace A of the gas system

$A, B/E$	Index operation under the constraint $E_A^{\mathrm{g}} + E_B^{\mathrm{c}} = E$
A_{ij}	Matrix elements of an operator $\hat{A}$
A_{ij}^*	Complex conjugate matrix elements of an operator $\hat{A}$
AR	Accessible region
$\|B, b\rangle$	Basis state of the container system c
B	Index of the energy subspace of the container system with energy E_B^{c}
b	Index of the degenerate eigenstates belonging to one energy subspace B of the container system
c	Label for container
$D_{\hat{\rho}\hat{\rho}'}^2$	Distance measure (Bures metric)
D	Diagonal deviation matrix
d	Dimension of the Liouville space
d_A	Dimension of special subspace A
E_i	Energy of the state $\|i\rangle$
E	Energy
E_0	Zero-point energy
E^{c}	Total energy of container system
E_B^{c}	Energy eigenvalues of the container system
E^{g}	Total Energy of gas system g
E_A^{g}	Energy eigenvalues of the gas system g
E	Off-diagonal part of a matrix, deviation matrix
$F_{\hat{\rho}\hat{\rho}'}$	Fidelity between $\hat{\rho}$ and $\hat{\rho}'$
$F(T, V)$	Free energy
F	Jacobian matrix (functional matrix)
F^J	Jacobian matrix in subspace J
F^{AB}	See F^J
F	Force
$G(T, p)$	Gibbs free energy
$g(E)$	State density of subsystems
$g(\boldsymbol{\gamma}, E)$	Energy spectrum of wave package $\|\gamma\rangle$
$G^{\mathrm{c}}(E_B^{\mathrm{c}})$	State density of the container system
$G^{\mathrm{g}}(E_A^{\mathrm{g}})$	State density of the gas system
$G(E)$	State density at energy E
g	Label for gas subsystem
$H(S, p)$	Enthalpy
$H(\boldsymbol{q}, \boldsymbol{p})$	Hamilton function
$\hat{H}$	Hamiltonian
$\hat{H}_{\mathrm{loc}}^{(\mu)}$	Local Hamiltonian of subsystem μ
$\hat{H}_{\mathrm{F}}^{(\mu,\mu+1)}$	Next neighbor Förster coupling
$\hat{H}_{\mathrm{R}}^{(\mu,\mu+1)}$	Next neighbor random coupling
$\hat{H}_{\mathrm{NR}}^{(\mu,\mu+1)}$	Next neighbor non-resonant coupling
$\hat{H}_0$	Unperturbed Hamiltonian
$\hbar = h/2\pi$	Planck's constant

$\hat{H}^{\mathrm{c}}$	Hamiltonian of the container system c
$\hat{H}^{\mathrm{g}}$	Hamiltonian of the gas system g
$\mathcal{H}$	Hilbert space
$\mathcal{H}^{(\mu)}$	Hilbert space of subsystem μ
i	Imaginary unit
$\hat{I}$	Interaction operator
$\hat{I}^{\mathrm{gc}}$	Interaction between gas system and container system
$\|i\rangle$	Basis state
$\langle i\|$	Adjoint basis state
$\|i,j,\dots\rangle$	Product state of several subsystems
$\|i,t\rangle$	Time-dependent state
$\|i\rangle \otimes \|j\rangle = \|ij\rangle$	Product state of two subsystems
J	Label for a side condition
$\boldsymbol{j}$	Current
$\boldsymbol{j}_u$	Energy current
$\boldsymbol{j}_s$	Entropy current
$\hat{J}$	Current operator
$\hat{J}^{(\mu,\mu+1)}$	Local current operator between subsystem μ and $\mu+1$
k_{B}	Boltzmann constant
$\hat{\mathcal{L}}$	Super-operator acting on operators of the Liouville space (Lindblad operator)
$\hat{\mathcal{L}}_{\mathrm{coh}}$	Coherent part of Lindblad super-operator
$\hat{\mathcal{L}}_{\mathrm{inc}}^{1,2}$	Incoherent part of Lindblad super-operator
L	Transport coefficient (in general a matrix)
m	Number of micro states accessible for a system; mass
N_0^{c}	Number of levels in lower band
N_1^{c}	Number of levels in upper band
$N^{\mathrm{c}}(E_B^{\mathrm{c}})$	Total number of states in a subspace B, degeneracy
$N(E)$	Number of levels under the constraint $A, B/E$
$N^{\mathrm{g}}(E_A^{\mathrm{g}})$	Total number of states in a subspace A, degeneracy
n_{var}	Number of macro variables
$n^{(\mu)}$	Number of levels of subsystem μ
n_{tot}	Dimension of the total Hilbert space
N	Number of subsystems, particle number
n	Number of levels of a subsystem
$N_{AB} = N_A N_B$	Number of states in subspace AB
N_A	See $N^{\mathrm{g}}(E_A^{\mathrm{g}})$
N_B	See $N^{\mathrm{c}}(E_B^{\mathrm{c}})$
X_i	Extensive macro variable
$\boldsymbol{p}$	Vector of all momentum coordinates
$\boldsymbol{p}_\mu$	Momentum of the μ^{th} particle
$\hat{\boldsymbol{p}}_\mu$	Momentum operator of the μ^{th} particle
$\hat{P}^{\mathrm{ex}}$	Projector projecting out some part ex of the total state space

p	Pressure
$\hat{P}_{ii}^{(\mu)}$	Projector within subspace μ
$\hat{P}_{ij}$	Transition operator, projector for $i = j$
P	Purity
Q	Heat
$\hat{\boldsymbol{q}}_\mu$	Position operator of the μ^{th} particle
q	Position
$\boldsymbol{q}$	Vector of position coordinates of N particles
$\boldsymbol{q}_\mu$	Position of the μ^{th} particle
$\{\boldsymbol{q}_\nu^{\text{c}}\}$	Set of position coordinates of all container particles
$\{\boldsymbol{q}_\mu^{\text{g}}\}$	Set of position coordinates of all gas particles
r	Radial coordinate of a hypersphere
R^J	Radius of hypersphere in subspace J
R^{AB}	See R^J
S	Entropy
S^{tot}	Entropy of system and environment
S_{lin}	Linearized von Neumann entropy
$s(\boldsymbol{q}, t)$	Entropy density
$s(U, V)$	Specific entropy
$\text{sph}(n)$	n-dimensional hypersphere
$T(\boldsymbol{q}, t)$	Temperature field
T	Temperature
$T^{(\mu)}$	Local temperature of subsystem μ
t	Time
$\text{Tr}\{\dots\}$	Trace operation
$\text{Tr}_\mu\{\dots\}$	Partial trace over subsystem μ
$u(\boldsymbol{q}, t)$	Energy density
U	Internal energy
$\hat{U}$	Unitary transformation
$\hat{U}_0(t, t_0)$	Unitary transformation into the interaction picture
U_{ij}	Unitary matrix
$\hat{U}(t)$	Time evolution operator
$\hat{U}^{(\mu)}(t)$	Time evolution operator of subsystem μ
$\boldsymbol{v}$	Velocity vector
$\hat{V}$	Potential
$\hat{V}_{\text{I}}(t)$	Potential in the interaction picture
V	Volume
v	Hilbert space velocity
$W(E_A^{\text{g}}, E_B^{\text{c}})$	See W_{AB}
$W(E_B^{\text{c}})$	See W_B
$W^{\text{d}}(E_B^{\text{c}})$	Dominant probability of finding the container in E_B^{c}
$W^{\text{d}}(E_A^{\text{g}})$	Dominant probability of finding the gas in E_A^{g}
$W(E)$	Probability of finding the complete system at energy E
$W(E_A^{\text{g}})$	See W_A

$W(\boldsymbol{q},\boldsymbol{p},t)$	Probability of finding a point in phase space at position $(\boldsymbol{q},\boldsymbol{p})$ at time t
$W_{ij}^{(12)}$	Probability of finding subsystem 1 in state i and subsystem 2 in state j
$W_{1\to 0}$	Rate for a decay from $\|1\rangle$ to $\|0\rangle$
$W_{ij}(t)$	Transition probability from state j into state i
$W_i = \rho_{ii}$	Statistical weight of, or probability of finding the system in the pure state $\hat{P}_{ii}$
W_{AB}	Joint probability of finding the gas system at energy E_A^{g} and the container system at energy E_B^{c}
W_B	Probability of finding the container system at energy E_B^{c}
W_{AB}^{d}	Dominant probability distribution for the whole system
W_A	Probability of finding the gas system at energy E_A^{g}
$\{W_{AB}\}$	Set of all probabilities W_{AB}, probability distribution
$\{W_{AB}^{\mathrm{d}}\}$	Set of dominant probabilities
W_A^{d}	Dominant probability distribution for the gas system, equivalent to $W^{\mathrm{d}}(E_A^{\mathrm{g}})$
W_B^{d}	Dominant probability distribution for the container system, equivalent to $W^{\mathrm{d}}(E_B^{\mathrm{c}})$
$\|x\rangle$	State vector indexed by position x
Z_i	Macro variables
Z	Canonical partition function
α	Lagrange parameter or exponent for spectra
β	Lagrange parameter; inverse temperature
$\gamma_i = (p_i, q_i)$	Wavepackage coordinates
$\|\boldsymbol{\gamma}\rangle$	Complete but not necessarily orthogonal basis
$\delta\epsilon$	Band width
ΔE_0^{c}	Energy level spacing in lower band
ΔE_1^{c}	Energy level spacing in upper band
$\Delta_{\mathrm{H}}^2(f)$	Hilbert space variance of a quantity f
ΔW_A	Variance of W_A
ε_i	Deviation, small quantity
ε_{AB}	Deviation from W_{AB}^{d}
η	Carnot efficiency
η_i	Real part of ψ_i
κ	Heat conductivity
λ	Coupling constant
λ_0	Mean absolute value of interaction matrix element
μ	Subsystem index; chemical potential
ξ_i	Imaginary part of ψ_i; intensive macro variable
$\hat{\rho}_{\mathrm{eq}}^{\mathrm{g}}$	Equilibrium state of subsystem g
$\hat{\rho}$	Density operator
$\hat{\rho}^{(\mu)}$	Reduced density operator subsystem μ
$\hat{\rho}_{\mathrm{irrel}}$	Irrelevant part of density operator

$\hat{\rho}_{\mathrm{rel}}$	Relevant part of density operator
$\rho_{\mathrm{irrel}}(\boldsymbol{q},\boldsymbol{p})$	Irrelevant part of the prob. dist. func. $\rho(\boldsymbol{q},\boldsymbol{p})$
$\rho_{\mathrm{rel}}(\boldsymbol{q},\boldsymbol{p})$	Relevant part of the prob. dist. func. $\rho(\boldsymbol{q},\boldsymbol{p})$
$\rho(\boldsymbol{q},\boldsymbol{p})$	Probability distribution function
ρ_{ij}	Density matrix
$\hat{\rho}$	Density operator
$\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$	Minimum purity state of the gas g
$\hat{\sigma}_{\pm}$	Raising/lowering operator
$\hat{\sigma}_i$	Pauli operator
σ	Standard deviation
$\|\psi\rangle$	Wave function, arbitrary pure state
$\|\psi_{\mathrm{I}}(t)\rangle$	Arbitrary wave function in the interaction picture
$\|\psi^{(\mu)}\rangle$	Pure state of subsystem μ
ψ^{AB}_{ab}	Amplitute of a product state $\|Aa\rangle \otimes \|Bb\rangle$
ψ^{A}_{a}	Product state amplitude of gas system
ψ^{B}_{b}	Product state amplitude of container system
ψ_{ij}	Amplitude of a basis product state $\|ij\rangle$
ψ_i	Amplitude of basis state $\|i\rangle$
ω_{ji}	Transition frequency from state j into state i
$\Omega(E)$	Total volume of phase space below the energy surface $H(\boldsymbol{q},\boldsymbol{p}) = E$
$\mathcal{V}\ (\{W_{AB}\})$	Size of region in Hilbert space with probability distribution $\{W_{AB}\}$
$\mathcal{V}^{\,d}$	Size of dominant region in Hilbert space

1 Introduction

Over the years enormous effort was invested in proving ergodicity, but for a number of reasons, confidence in the fruitfulness of this approach has waned.
— Y. Ben-Menahem and I. Pitowsky [11]

Originally, thermodynamics was a purely phenomenological science. Early scientists (Galileo, Santorio, Celsius, Fahrenheit) tried to give definitions for quantities which were intuitively obvious to the observer, like pressure or temperature, and studied their interconnections. The idea that these phenomena might be linked to other fields of physics, like classical mechanics, e.g., was not common in those days. Such a 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.

At the end of the 19th century, when the atomic theory became popular, researchers began to think of a gas as a huge amount of bouncing balls inside a box. 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 [18], when he connected entropy, a quantity which so far had only been defined phenomenologically, to the volume of a certain region in phase space, an object defined within classical mechanics. This was an enormous step forward, especially from a practical point of view. Taking this connection for granted one could now calculate all sorts of thermodynamic properties of a system from its Hamilton function. This gave rise to modern thermodynamics, a theory the validity of which is beyond any doubt today. Its results and predictions are a basic ingredient for the development of all kinds of technical apparatuses ranging from refrigerators to superconductors.

Boltzmann himself, however, tried to prove the conjectured connection between the phenomenlogical and the theoretical entropies, but did not succeed without invoking other assumptions like the famous ergodicity postulate or the hypothesis of equal "a priori probabilities". Later on, other physicists (Gibbs [41], Birkhoff [14], Ehrenfest [32], von Neumann [88], etc.) tried to prove those assumptions, but none of them seems to have solved the problem satisfactorily. It has been pointed out, though, that there are more properties of the entropy to be explained than its mere equivalence with the region in phase space, before thermodynamics can be reduced to classical mechanics, thus the discussion is still ongoing [11]. The vast majority of the work done in this field is based on classical mechanics.

Meanwhile, quantum theory, also initially triggered by the atomic hypothesis, has made huge progress during the last century and is today believed

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 3–5 (2004)
http://www.springerlink.com/

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 caricature of what a gas really is. Furthermore, thermodynamic 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 has led to the desired result (e.g., demonstration of ergodicity) or not. The fact that a basically classical approach apparently did so well may even be considered rather surprising.

Of course, there have been suggestions of how to approach the problem on the basis of quantum theory [69, 70, 73, 87, 96, 116, 134, 136], but again, none of them seems to have established the emergence of thermodynamics from quantum mechanics as an underlying theory in a conclusive way.

The text at hand can be viewed as a contribution to this ongoing discussion. Thus, on the one hand, one might consider this work as a somewhat new explanation for the emergence of thermodynamic behavior. This point of view definitely leaves one question open: whether or not all macroscopic thermodynamic systems belong to the class of systems that will be examined in the following. The answer to this question is beyond the scope of this text.

Furthermore, this quantum approach to thermodynamics may turn out not to be a one-way road. In fact, this delicate interplay between quantum mechanics and thermodynamics could possibly shed new light on some interpretational problems within quantum mechanics: with the “exorcism” of subjective ignorance as a guiding principle underlying thermodynamic states, the general idea of quantum states representing subjective knowledge might lose much of its credibility.

However, this book might be looked at also from another, less speculative angle. Rather than asking how thermodynamic behavior of typical systems might be explained, one can ask whether the principles of thermodynamics are powerful tools for predictions and whether their descriptions might be applicable to systems other than the pertinent large, many-particle systems. It is a main conclusion of this work that the answer has to be positive. For it turns out that a large class of small quantum systems without any restriction concerning size or particle number show thermodynamic behavior with respect to an adequately defined set of thermodynamic variables. This behavior (“nano-thermodynamics”) requires some embedding but is nevertheless established entirely on the basis of the Schrödinger equation.

So it is left to the reader to decide whether there is room and need for a foundation of thermodynamics on the basis of quantum theory or whether he simply wants to gain insight into how the applicability of thermodynamics can be extended down to the microscopical scale; in both cases we hope the reading will be interesting and clarifying.

This book is not intended to be a review, not even of the most important contributions, which more or less point in a similar direction. Related work includes, in particular, the so-called decoherence theory. We cannot do justice to the numerous investigations; we merely give a few references [21, 42, 45, 135]. It might be worth mentioning here that decoherence has, during the last years, mainly been discussed as one of the main obstacles in the implementation of large-scale quantum computers [90], possibly neglecting other aspects of the phenomenon.

Last but not least a short “manual” for reading this book shall be given here.

Chapters 2 and 3 are not meant to serve as a full-fledged introduction, more as a reminder of the central topics in quantum mechanics and thermodynamics. They may very well be skipped by a reader familiar with these subjects. Chapter 4 is a collection of historical approaches to thermodynamics with a focus on their insufficiencies hereby neglecting their undoubtable brilliance. Again this chapter is not imperative for the understanding of Part II.

Chapter 5 lists the properties of thermodynamic quantities that need to be derived from an underlying theory (quantum mechanics). This derivation is then given in the remainder of Part II. In Chap. 6 the central ideas of this quantum approach to thermodynamics are explained in plain text (no formulas). For a quick survey it might be read without referring to anything else. Starting with Chap. 7 and throughout Part II, these ideas are derived in detail, and will probably only be enlightening if read from the beginning. Exceptions are Chap. 11 and Chap. 14, which are important for the general picture, but have their own “selfcontained” messages.

Chapter 18 mainly consists of numerical illustrations of the analytically derived principles in Part II. In order to get an idea of the benefits of the theories derived in Part II, it might be read as a “stand alone” chapter. In Chap. 19 recent results on quantum heat conduction are presented, it may also be read individually, for it is only loosely connected to the rest of the book as far as mathematical techniques are concerned.

2 Basics of Quantum Mechanics

> *Indeed my favourite key to understanding quantum mechanics is that a subsystem cannot be isolated by tracing from an enveloping pure state without generating impurity: the probability associated with measurement develops because the observer must implicitly trace himself away from the observed system.*
>
> — E. Lubkin [76]

Before we can start with the quantum mechanical approach to thermodynamics we have to introduce some fundamental terms and definitions of standard quantum mechanics for later reference. This chapter should introduce the reader only to some indispensible concepts of quantum mechanics necessary for the text at hand, but is far from being a complete overview of this subject. For a complete introduction we refer to standard textbooks [9, 70, 79, 89, 113, 126].

2.1 Introductory Remarks

The shortcomings of classical theories had become apparent by the end of the 19th century. Interestingly enough, one of the first applications of quantum ideas has been within thermodynamics: Planck's famous formula for black body radiation was based on the hypothesis that the exchange of energy between the container walls and the radiation field should occur in terms of fixed energy quanta only. Later on, this idea has been put on firmer ground by Einstein postulating his now well known rate equations [75].

Meanwhile quantum mechanics has become a theory of unprecedented success. So far, its predictions have always been confirmed by experiment.

Quantum mechanics is usually defined in terms of some loosely connected axioms and rules. Such a foundation is far from the beauty of, e.g., the "principles" underlying classical mechanics. Motivated, in addition, by notorious interpretation problems, there have been numerous attempts to modify or "complete" quantum mechanics.

A first attempt was based on so-called "hidden variables" [10]. Its proponents essentially tried to expel the non-classical nature of quantum mechanics. More recent proposals intend to "complete" quantum mechanics not within mechanics, but on a higher level: by means of a combination with gravitation theory (Penrose [102]), with psychology (Stapp [122]) or with (quantum-) information theory [26, 38].

While the emergence of classicality from an underlying quantum substrate has enjoyed much attention recently, it has so far not been appreciated that

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 7–20 (2004)
http://www.springerlink.com/

the understanding of quantum mechanics may benefit also from subjects like quantum thermodynamics.

2.2 Operator Representations

In quantum mechanics we deal with systems (Hamilton models), observables, and states. They all are represented by Hermitian operators. Their respective specification requires data (parameters), which have to be defined with respect to an appropriate reference frame. These frames are operator representations. Let us in the following consider some aspects of these operator representations in detail. First of all we will concentrate on simple systems and their representation.

2.2.1 Transition Operators

If we restrict ourselves to systems living in a finite and discrete Hilbert space $\mathcal{H}$ (a complex vector space of dimension n_{tot}), we may introduce a set of orthonormal state vectors $|i\rangle \in \mathcal{H}$. From this orthonormal and complete set of state vectors with

$$\langle i|j\rangle = \delta_{ij} \ , \quad i,j = 1,2,\dots,n_{\mathrm{tot}} \ , \tag{2.1}$$

we can define n_{tot}^2 transition operators (in general non-Hermitian)

$$\hat{P}_{ij} = |i\rangle\langle j| \ , \quad \hat{P}_{ij}^{\dagger} = \hat{P}_{ji} \ . \tag{2.2}$$

These operators are, again, orthonormal in the sense that

$$\mathrm{Tr}\left\{\hat{P}_{ij}\hat{P}_{i'j'}^{\dagger}\right\} = \delta_{ii'}\delta_{jj'} \ , \tag{2.3}$$

where $\mathrm{Tr}\{\dots\}$ denotes the trace operation. Furthermore, they form a complete set in so-called Liouville space, into which any other operator $\hat{A}$ can be expanded,

$$\hat{A} = \sum_{i,j} A_{ij}\hat{P}_{ij} \ , \tag{2.4}$$

$$A_{ij} = \mathrm{Tr}\left\{\hat{A}\,\hat{P}_{ij}^{\dagger}\right\} = \langle i|\hat{A}|j\rangle \ . \tag{2.5}$$

The n_{tot}^2 parameters are, in general, complex ($2n_{\mathrm{tot}}^2$ real numbers). For Hermitian operators we have, with

$$\hat{A}^{\dagger} = \sum_{i,j} A_{ij}^{*}\hat{P}_{ij}^{\dagger} = \sum_{i,j} A_{ji}\hat{P}_{ji} = \hat{A} \ , \tag{2.6}$$

$$A_{ij}^{*} = A_{ji} \ , \tag{2.7}$$

i.e., we are left with n_{tot}^2 independent real numbers. All these numbers must be given to uniquely specify any Hermitian operator $\hat{A}$.

2.2.2 Pauli Operators

There are many other possibilities to define basis operators, besides the transition operators. For $n_{\text{tot}} = 2$ a convenient set is given by the so-called Pauli operators $\hat{\sigma}_i$ $(i = 0, \dots, 3)$. The new basis operators can be expressed in terms of transition operators

$$\hat{\sigma}_1 = \hat{P}_{12} - \hat{P}_{21} \ , \tag{2.8}$$

$$\hat{\sigma}_2 = \mathrm{i}(\hat{P}_{21} - \hat{P}_{12}) \ , \tag{2.9}$$

$$\hat{\sigma}_3 = \hat{P}_{11} - \hat{P}_{22} \ , \tag{2.10}$$

$$\hat{\sigma}_0 = \hat{1} \ . \tag{2.11}$$

These operators are Hermitian and – except for $\hat{\sigma}_0$ – traceless. The Pauli operators satisfy several important relations: $(\hat{\sigma}_i)^2 = \hat{1}$ and $[\hat{\sigma}_1, \hat{\sigma}_2] = 2\mathrm{i}\hat{\sigma}_3$ and their cyclic extensions. Since the Pauli operators form a complete orthonormal operator basis, it is possible to expand any operator in terms of these basis operators. Furthermore we introduce raising and lowering operators, in accordance with

$$\hat{\sigma}_+ = \hat{\sigma}_1 + \mathrm{i}\hat{\sigma}_2 \ , \quad \hat{\sigma}_- = \hat{\sigma}_1 - \mathrm{i}\hat{\sigma}_2 \ . \tag{2.12}$$

Also for higher dimensional cases, $n_{\text{tot}} \geq 2$, one could use as a basis the Hermitian generators of the $\mathrm{SU}(n_{\text{tot}})$ group.

2.2.3 State Representation

The most general way to note the information about a state of a quantum mechanical system is by its density matrix, ρ_{ij}, which specifies the representation of the density operator,

$$\hat{\rho} = \sum_{i,j} \rho_{ij} \hat{P}_{ij} \tag{2.13}$$

subject to the condition

$$\mathrm{Tr}\,\{\hat{\rho}\} = \sum_i \rho_{ii} = 1 \ . \tag{2.14}$$

The expectation value for some observable $\hat{A}$ in state $\hat{\rho}$ is now given by

$$\langle A \rangle = \mathrm{Tr}\{\hat{A}\hat{\rho}\} = \sum_{i,j} A_{ij} \rho_{ij} \ . \tag{2.15}$$

The density matrix $\rho_{ij} = \langle i|\hat{\rho}|j\rangle$ is a positive definite and Hermitian matrix. The number of independent real numbers needed to specify $\hat{\rho}$ is thus $d =$

$n_{\mathrm{tot}}^2 - 1$. For the density operator of an arbitrary pure state $|\psi\rangle$ we have $\hat{\rho} = |\psi\rangle\langle\psi|$. In the eigenrepresentation one finds, with $W_i = \rho_{ii}$,

$$\hat{\rho} = \sum_i W_i \hat{P}_{ii} \ , \tag{2.16}$$

which can be seen as a "mixture" of pure states $\hat{P}_{ii} = |i\rangle\langle i|$ with the statistical weight W_i. From this object the probability $W(|\psi\rangle)$ to find the system in an arbitrary pure state, expanded in the basis $|i\rangle$

$$|\psi\rangle = \sum_i \psi_i |i\rangle \ , \tag{2.17}$$

can be calculated as

$$W(|\psi\rangle) = \langle\psi|\hat{\rho}|\psi\rangle = \sum_i |\psi_i|^2 \, W_i \ . \tag{2.18}$$

To measure the distance of two arbitrary, not necessarily pure states, given by $\hat{\rho}$ and $\hat{\rho}'$ we define a "distance measure"

$$D^2_{\hat{\rho}\hat{\rho}'} = \mathrm{Tr}\left\{(\hat{\rho} - \hat{\rho}')^2\right\} \ . \tag{2.19}$$

This commutative measure (sometimes called Bures metric) has a number of convenient properties: $D^2_{\hat{\rho}\hat{\rho}'} \geq 0$ with the equal sign holding if and only if $\hat{\rho} = \hat{\rho}'$; the triangle inequality holds as expected for a conventional distance measure; for pure states

$$D^2_{|\psi\rangle|\psi'\rangle} = 2\Big(1 - |\langle\psi|\psi'\rangle|^2\Big) \leq 2 \tag{2.20}$$

and D^2 is invariant under unitary transformations. A second measure of distance is the fidelity defined by [90]

$$F_{\hat{\rho}\hat{\rho}'} = \mathrm{Tr}\left\{\left(\sqrt{\hat{\rho}}\,\hat{\rho}'\sqrt{\hat{\rho}}\right)^{1/2}\right\} \ . \tag{2.21}$$

For pure states F is just the modulus of the overlap: $F_{|\psi\rangle|\psi'\rangle} = |\langle\psi|\psi'\rangle|$.

2.2.4 Purity and von Neumann Entropy

For a pure state all matrix elements in (2.13) of the density matrix are zero except $\rho_{ii} = 1$, say, i.e., the density operator $\hat{\rho} = \hat{P}_{ii}$ is a projection operator. Obviously in this case $\hat{\rho}^2 = \hat{\rho}$, due to the properties of the projection operator, so that the so-called purity becomes

$$P = \mathrm{Tr}\left\{\hat{\rho}^2\right\} = 1 \ . \tag{2.22}$$

In general, we have

$$P(\hat{\rho}) = \sum_{i,j} \sum_{i',j'} \rho_{ij}\, \rho_{i'j'}\, \mathrm{Tr}\left\{ \hat{P}_{ij}\, \hat{P}^{\dagger}_{i'j'} \right\} = \sum_{i,j} |\rho_{ij}|^2 > \frac{1}{n_{\mathrm{tot}}} \ . \tag{2.23}$$

Because of the Cauchy–Schwarz relation

$$|\rho_{ij}|^2 \leq \rho_{ii}\, \rho_{jj} \ , \tag{2.24}$$

we conclude that $P \leq 1$. The equality sign holds for pure states only. P can be calculated for any density matrix without prior diagonalization. In the diagonal representation (cf. (2.16)) the purity is simply the sum of the squares of the probabilities W to find the system in a respective eigenstate,

$$P(\{W_i\}) = \mathrm{Tr}\left\{\hat{\rho}^2\right\} = \sum_i (W_i)^2 \ . \tag{2.25}$$

Note that the purity itself is invariant with respect to unitary transformations. Its value does not depend on the representation chosen.

Furthermore, a very important quantity is another measure called the von Neumann entropy [90]. Also this measure is defined for any state $\hat{\rho}$ as

$$S(\hat{\rho}) = -k_{\mathrm{B}} \mathrm{Tr}\left\{\hat{\rho} \ln \hat{\rho}\right\} \geq 0 \ , \tag{2.26}$$

wherein k_{B} denotes a proportional constant, the Boltzmann constant. (At this point the inclusion of k_{B} is arbitrary and not yet meant to anticipate any connection to thermodynamics.) For a pure state the minimum entropy $S = 0$ is reached. The maximum entropy obtains for

$$\rho_{ij} = \frac{1}{n_{\mathrm{tot}}}\, \delta_{ij} \ , \quad i, j = 1, 2, \ldots, n_{\mathrm{tot}} \ , \tag{2.27}$$

i.e., for a density matrix proportional to the normalized unit matrix, with the entropy

$$S_{\mathrm{max}} = k_{\mathrm{B}} \ln n_{\mathrm{tot}} \ . \tag{2.28}$$

In the same limit the purity P is minimal,

$$P_{\mathrm{min}} = \frac{1}{n_{\mathrm{tot}}} \ . \tag{2.29}$$

The maximum entropy (or minimum purity) is thus found for the broadest possible probability distribution, the equipartition over all pure states (remember (2.27)). Therefore S and P are both measures for the “broadness” of the distribution.

The purity can be expressed as a rather simple function of the full state, the evaluation of which does not require the diagonalization of a matrix, as opposed to the calculation of the von Neumann entropy. We will thus mainly consider P rather than S.

In general, though, these two measures do not uniquely map onto each other. Nevertheless in the limits of maximum S (minimum P) and maximum P (minimum S) they do. The formal approximation $\ln \hat{\rho} \approx \hat{\rho} - \hat{1}$ leads to the "linearized" entropy

$$S_{\text{lin}} = k_{\text{B}}(1 - P) \geq 0 \,. \tag{2.30}$$

Since, as will be shown in Sect. 2.4, S is a constant of motion, the question for the possible origin of $S > 0$ arises. One interpretation is essentially classical and traces a finite S back to subjective ignorance. In the eigenrepresentation of the density operator (see (2.16)) the density operator can be seen as a "mixture" of pure states $\hat{P}_{ii} = |i\rangle\langle i|$ and the entropy then reads

$$S = -k_{\text{B}} \sum_i W_i \, \ln W_i \,. \tag{2.31}$$

Alternatively, a nonpure state may result from the system under consideration being entangled with another system, while the total state is pure. In this case S indicates a principal uncertainty. It is always possible to find such an embedding, as will be discussed in the next section.

2.2.5 Bipartite Systems

Systems typically consist of subsystems. In the case of a bipartite system, the total Hilbert space can be decomposed into a product space

$$\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)} \,, \tag{2.32}$$

with dimension $n_{\text{tot}} = n^{(1)} \cdot n^{(2)}$. A complete set of orthonormal vectors is then given by the product states ($\otimes$ means tensor product of the vectors involved)

$$|ij\rangle = |i\rangle \otimes |j\rangle \,, \tag{2.33}$$

with $i = 1, 2, \ldots, n^{(1)}$ numbering the states in $\mathcal{H}^{(1)}$ and $j = 1, 2, \ldots, n^{(2)}$ in $\mathcal{H}^{(2)}$. The states fulfill the orthonormality relation

$$\langle ij|i'j'\rangle = \delta_{ii'} \, \delta_{jj'} \,. \tag{2.34}$$

Based on this we can define the transition operators

$$\hat{P}_{ij|i'j'} = |ij\rangle\langle i'j'| = \hat{P}^{(1)}_{ii'} \otimes \hat{P}^{(2)}_{jj'} \,, \tag{2.35}$$

where $\hat{P}^{(\mu)}_{ii'}$ is a transition operator in the subspace of the subsystem $\mu = 1, 2$. These, again, form a complete orthogonal set such that any operator $\hat{A}$ can be expanded in the form

$$\hat{A} = \sum_{i,j} \sum_{i',j'} A_{ij|i'j'} \, \hat{P}_{ij|i'j'} \,. \tag{2.36}$$

For a pure state

$$|\psi\rangle = \sum_{i,j} \psi_{ij}\, |i\rangle \otimes |j\rangle \tag{2.37}$$

the density operator $\hat{\rho} = |\psi\rangle\langle\psi|$ has the matrix representation

$$\rho_{ij|i'j'} = \psi_{ij}\, \psi^*_{i'j'} \; . \tag{2.38}$$

If we are interested in the state of one of the subsystems alone we have to trace over the other subsystem. The reduced density operator of the system of interest is now given by

$$\begin{aligned} \hat{\rho}^{(1)} = \mathrm{Tr}_2\left\{\hat{\rho}\right\} &= \sum_{i,i'} \sum_{j} \langle ij|\hat{\rho}|i'j\rangle\, |i\rangle\langle i'| \\ &= \sum_{i,i'} \rho_{ii'}\, \hat{P}^{(1)}_{ii'} \; , \end{aligned} \tag{2.39}$$

with $\rho_{ii'} = \sum_j \rho_{ij|i'j}$. Here $\mathrm{Tr}_2\{\ldots\}$ means trace operation within Hilbert space $\mathcal{H}^{(2)}$. The result for subsystem 2 is obtained by exchanging the indices of the two subsystems.

The expectation value for any local operator $\hat{A}^{(1)} \otimes \hat{1}^{(2)}$ can be calculated from

$$\langle \hat{A}^{(1)} \rangle = \mathrm{Tr}_1\left\{\hat{A}^{(1)} \hat{\rho}^{(1)}\right\} \; . \tag{2.40}$$

The corresponding purity, say, for the reduced state of the first subsystem, is

$$P\big(\hat{\rho}^{(1)}\big) = \sum_{i,i'} |\rho_{ii'}|^2 = \sum_{i,i'} \sum_{j,j'} \rho_{ij|i'j}\; \rho_{i'j'|ij'} \; . \tag{2.41}$$

Furthermore, the reduced von Neumann entropies are given by

$$S\big(\hat{\rho}^{(\mu)}\big) = -k_{\mathrm{B}} \mathrm{Tr}_\mu \left\{\hat{\rho}^{(\mu)} \ln \hat{\rho}^{(\mu)}\right\} \; , \quad \mu = 1, 2 \; . \tag{2.42}$$

One easily convinces oneself that for

$$\hat{\rho} = \hat{\rho}^{(1)} \otimes \hat{\rho}^{(2)} \tag{2.43}$$

the total entropy is additive,

$$S = S\big(\hat{\rho}^{(1)}\big) + S\big(\hat{\rho}^{(2)}\big) \; . \tag{2.44}$$

In general, the theorem by Araki and Lieb [7] tells us that

$$\left| S\big(\hat{\rho}^{(1)}\big) - S\big(\hat{\rho}^{(2)}\big) \right| \leq S \leq S\big(\hat{\rho}^{(1)}\big) + S\big(\hat{\rho}^{(2)}\big) \; . \tag{2.45}$$

This theorem implies that if the total system is in a pure state ($S = 0$) then $S(\hat{\rho}^{(1)}) = S(\hat{\rho}^{(2)})$, no matter how the system is partitioned. Under the same condition $P(\hat{\rho}^{(1)}) = P(\hat{\rho}^{(2)})$. Then if $S(\hat{\rho}^{(1)}) = S(\hat{\rho}^{(2)}) > 0$, it follows that (2.44) does not apply and the total (pure) state cannot be written in a product form. This is interpreted to result from "entanglement", for which the local entropies $S(\hat{\rho}^{(1)}) = S(\hat{\rho}^{(2)})$ thus constitute an appropriate measure.

Such pure entangled states have been of central 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 has raised the question of whether quantum mechanics could be valid at all [33]. Nevertheless, these states can theoretically be shown to result from product states, if the subsystems are allowed to interact for a while. On a small scale such a build-up has been demonstrated experimentally; it is a widespread belief that entanglement as a fundamental quantum mechanical property should show up mainly between very small objects.

2.2.6 Multi-Partite Systems

Alternatively, one may consider a network of N subsystems of dimension n each. Then $n_{\text{tot}} = n^N$. As a consequence of the direct product structure, the number of parameters required to specify a density operator then grows exponentially with N

$$d = n^{2N} - 1 \,. \tag{2.46}$$

For the classical system of N point particles we would need $6N$ real parameters, i.e., we would just have to specify position and momentum of each individual particle. This so-called phase space is the direct sum of the individual particle spaces. The analog in the quantum case would be to specify the local states of the N subsystems, for which we would need $(n^2 - 1)N$ parameters. (This was the dimension of the direct sum of subsystem Liouville spaces.) Defining

$$\gamma = \frac{d}{(n^2 - 1)N} \,, \tag{2.47}$$

we see that for $n = 2$, $N = 3$, $\gamma = 7$, but for $N = 10$, $\gamma \approx 30000$. The tremendous information needed over the local parameters is due to the fact that correlations (entanglement) dominate, in general. For product states $\gamma = 1$.

The blow-up of γ is a typical quantum property, closer to the heart of quantum mechanics than the famous Heisenberg uncertainty relation. Both are due to the non-commutativity of the underlying operators, though.

The number of parameters needed to specify a Hamilton model typically grows only polynomially with N. This is because direct interactions are usually restricted to finite clusters, e.g., up to pairs.

2.3 Dynamics

So far, we have considered some properties of Hilbert spaces, the basis operators and appropriate states. We turn now to some dynamical aspects of quantum systems.

The unitary dynamics of a closed system generated by a Hamilton operator $\hat{H}$ is given by the Schrödinger equation

$$\mathrm{i}\hbar \frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle \ , \tag{2.48}$$

for the time-dependent pure state $|\psi(t)\rangle$. This is the fundamental equation specifying the so-called Schrödinger picture: here the state vectors $|\psi(t)\rangle$ carry all dynamics, while the basic operators are time-independent. But note that the Hamiltonian could include explicitly time-dependent potentials.

From the Schrödinger equation one can easily derive the evolution equation directly for the density operator. This is the Liouville–von–Neumann equation

$$\mathrm{i}\hbar \frac{\partial \hat{\rho}}{\partial t} = \left[\hat{H}, \hat{\rho}\right] \ , \tag{2.49}$$

with $\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A}$ defining the commutator. This equation can be written in the form

$$\frac{\partial \hat{\rho}}{\partial t} = \hat{\mathcal{L}}\hat{\rho} \ , \tag{2.50}$$

where $\hat{\mathcal{L}}$ is a so-called super-operator acting (here) on the operator $\hat{\rho}$ to produce the new operator

$$\hat{\mathcal{L}}\hat{\rho} = -\frac{\mathrm{i}}{\hbar}\left[\hat{H}, \hat{\rho}\right] \ . \tag{2.51}$$

Modified super-operators control the dynamics of open quantum systems, which we will consider in detail in Sect. 4.8.

The Liouville–von–Neumann equation can formally be solved by

$$\hat{\rho}(t) = \hat{U}(t)\, \hat{\rho}(0)\, \hat{U}^{\dagger}(t) \ , \tag{2.52}$$

where the unitary time evolution operator, $\hat{U}^{\dagger}\hat{U} = \hat{U}\hat{U}^{\dagger} = \hat{1}$, also obeys the Schrödinger equation,

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\hat{U}(t) = \hat{H}(t)\, \hat{U}(t) \ . \tag{2.53}$$

For $\partial\hat{H}/\partial t = 0$, i.e., no explicit time-dependent Hamiltonian, it has the formal solution

$$\hat{U}(t) = \mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar} \; . \tag{2.54}$$

When represented with respect to a specific set of basis operators, the Liouville–von–Neumann equation is equivalent to

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\rho_{ij}(t) = \mathrm{Tr}\left\{ \left[\hat{H}, \hat{\rho}(t)\right] \hat{P}_{ij}^{\dagger} \right\} \; . \tag{2.55}$$

This equation determines the evolution of the matrix elements of the density operator. The solution $\rho_{ij}(t)$, subject to the condition $\sum_i \rho_{ii} = 1$, can thus be visualized as a deterministic quantum trajectory in Liouville space, controlled by the Hamiltonian and by the initial state $\hat{\rho}(0)$.

In the Heisenberg picture, the dynamics is carried by time-dependent observables

$$\hat{A}_{\mathrm{H}}(t) = \hat{U}^{\dagger}(t)\,\hat{A}\,\hat{U}(t) \; , \tag{2.56}$$

while the states are constant, $\hat{\rho}_{\mathrm{H}}(t) = \hat{\rho}(0)$. If $\partial\hat{A}/\partial t = 0$ in the Schrödinger picture, the corresponding evolution equation for the now time-dependent operators reads

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\hat{A}_{\mathrm{H}} = -\left[\hat{H}, \hat{A}_{\mathrm{H}}\right] \; . \tag{2.57}$$

In either picture the time-dependence of the expectation value of an operator $\langle\hat{A}\rangle = \mathrm{Tr}\{\hat{A}\,\hat{\rho}\} = \mathrm{Tr}\{\hat{A}_{\mathrm{H}}\,\hat{\rho}(0)\}$ is given by

$$\begin{aligned} \mathrm{i}\hbar \frac{\partial}{\partial t}\langle\hat{A}\rangle &= \mathrm{Tr}\left\{ \left[\hat{H}, \hat{\rho}\right] \hat{A} \right\} = -\mathrm{Tr}\left\{ \left[\hat{H}, \hat{A}_{\mathrm{H}}\,\hat{\rho}(0)\right] \right\} \\ &= -\langle\left[\hat{H}, \hat{A}_{\mathrm{H}}\right]\rangle \; , \end{aligned} \tag{2.58}$$

which is known as the “Ehrenfest theorem”. Since this evolution equation is similar to the classical equation of motion based on the Poisson bracket, this theorem can be interpreted to state that “the classical equations of motion are valid for expectation values in quantum mechanics”.

2.4 Invariants

According to the Heisenberg equation of motion (2.57), conserved quantities are those which commute with the system Hamiltonian $\hat{H}$. In eigenrepresentation $\hat{H}$ can be written as

$$\hat{H} = \sum_{i=1}^{n_{\mathrm{tot}}} E_i\,\hat{P}_{ii} \; . \tag{2.59}$$

As a consequence, the projectors commute with the Hamiltonian itself,

$$\left[\hat{P}_{jj}, \hat{H}\right] = 0 \;, \quad j = 1, \dots, n_{\text{tot}} \;. \tag{2.60}$$

Since commutators are invariant under unitary transformations, the above relation thus holds in the Schrödinger as well as in the Heisenberg pictures. For the change of the energy distribution we find

$$\mathrm{i}\hbar \frac{\partial}{\partial t} W_j = \mathrm{i} \frac{\partial}{\partial t} \operatorname{Tr}\left\{\hat{P}_{jj}\,\hat{\rho}\right\} = -\operatorname{Tr}\left\{\left[\hat{H}, \hat{P}_{jj}^{(\mathrm{H})}\right] \hat{\rho}(0)\right\} = 0 \;, \tag{2.61}$$

i.e., the energy distribution, the probability of finding the system in state j, is a constant of motion.

Furthermore, defining the expectation value of an arbitrary function of the density operator $\hat{\rho}$

$$\langle f(\hat{\rho})\rangle = \operatorname{Tr}\left\{\hat{\rho}\, f(\hat{\rho})\right\} \;, \tag{2.62}$$

one infers that

$$\begin{aligned} \mathrm{i}\hbar \frac{\partial}{\partial t} \langle f(\hat{\rho})\rangle &= \mathrm{i}\hbar \operatorname{Tr}\left\{\frac{\partial}{\partial t} \hat{\rho} f(\hat{\rho})\right\} + \mathrm{i}\hbar \operatorname{Tr}\left\{\hat{\rho} \frac{\partial}{\partial t} f(\hat{\rho})\right\} \\ &= \operatorname{Tr}\left\{\left[\hat{H}, \hat{\rho}\right] f(\hat{\rho})\right\} + \operatorname{Tr}\left\{\hat{\rho}\left[\hat{H}, f(\hat{\rho})\right]\right\} \;. \end{aligned} \tag{2.63}$$

Here we have made use of the Liouville equation (2.49) and its variant

$$\mathrm{i}\hbar \frac{\partial}{\partial t} f(\hat{\rho}) = \left[\hat{H}, f(\hat{\rho})\right] \;. \tag{2.64}$$

Observing the invariance of the first trace term in (2.63) under cyclic permutations, we see that the right hand side cancels,

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f(\hat{\rho})\rangle = 0 \;. \tag{2.65}$$

Taking now $f(\hat{\rho}) = \hat{\rho}$, the term $\langle f(\hat{\rho})\rangle = \operatorname{Tr}\{\hat{\rho}^2\}$ is just the purity, so that

$$\frac{\mathrm{d}}{\mathrm{d}t} P = 0 \;. \tag{2.66}$$

For $f(\hat{\rho}) = \ln \hat{\rho}$ one concludes that the von Neumann entropy is invariant, too. In fact, any moment $\operatorname{Tr}\left\{(\hat{\rho})^k\right\}$ is a constant of motion in closed quantum systems. But note that the local reduced von Neumann entropy of a part of the system defined in (2.42) is not necessarily conserved under a unitary time evolution of the full system (see Sect. 6.1).

For later reference we finally investigate a bipartite system with the total Hamiltonian $\hat{H}$. Here we may encounter a situation for which

$$\left[\hat{A}^{(1)}, \hat{H}\right] = 0 \;, \quad \left[\hat{B}^{(2)}, \hat{H}\right] = 0 \;, \tag{2.67}$$

where the operator

$$\hat{A}^{(1)} = \sum_i A_i \, \hat{P}_{ii}^{(1)} \tag{2.68}$$

acts only on subsystem 1, and

$$\hat{B}^{(2)} = \sum_j B_j \, \hat{P}_{jj}^{(2)} \tag{2.69}$$

acts only on subsystem 2. As a consequence,

$$\left[\left(\hat{A}^{(1)} \right)^k , \hat{H} \right] = \sum_i A_i^k \left[\hat{P}_{ii}^{(1)} , \hat{H} \right] = 0 \; . \tag{2.70}$$

As this has to hold for any k, we conclude that

$$\left[\hat{P}_{ii}^{(1)} , \hat{H} \right] = 0 \; , \tag{2.71}$$

and correspondingly,

$$\left[\hat{P}_{jj}^{(2)} , \hat{H} \right] = 0 \; . \tag{2.72}$$

According to these considerations the expectation value

$$W_{ij}^{(12)} = \mathrm{Tr} \left\{ \hat{P}_{ii}(1) \, \hat{P}_{jj}(2) \, \hat{\rho} \right\} \tag{2.73}$$

is thus a conserved quantity, too. This expectation value is the joint probability for finding subsystem 1 in state i and subsystem 2 in state j.

2.5 Time-Dependent Perturbation Theory

Just as in classical mechanics or any other theoretical framework, there are very few examples that allow us to achieve an exact analytical solution. In quantum mechanics even numerical solutions are seriously constrained by exponential explosion of state parameters (cf. Sect. 2.2.6). Many, quite powerful, approximation schemes have been developed. For later reference we summarize here the basics of standard perturbation theory.

To consider time-dependent phenomena it is often very helpful – if not unavoidable – to use a perturbation theory instead of a full solution of the time-dependent problem. To outline such a theory, we use in addition to the Schrödinger and the Heisenberg pictures the interaction or Dirac picture.

2.5.1 Interaction Picture

In the interaction picture, both observables as well as states, are time-dependent. We consider the Hamilton operator

$$\hat{H} = \hat{H}_0 + \hat{V}(t) \; , \tag{2.74}$$

where $\hat{H}_0$ represents the unperturbed Hamiltonian and $\hat{V}(t)$ the time-dependent perturbation. According to the unitary transformation

$$\hat{U}_0(t, t_0) = \exp\Bigl(-\frac{\mathrm{i}}{\hbar}\hat{H}_0(t - t_0)\Bigr) \; , \tag{2.75}$$

where t_0 is the time at which the perturbation is switched on, one can transform the states as well as the operators of the Schrödinger picture into the interaction picture (index I)

$$|\psi(t)\rangle = \hat{U}_0(t, t_0)\,|\psi_\mathrm{I}(t)\rangle \; , \tag{2.76}$$

$$\hat{A}_\mathrm{I} = \hat{U}_0^\dagger(t, t_0)\;\hat{A}\;\hat{U}_0(t, t_0) \; . \tag{2.77}$$

Based on these transformations, the Schrödinger equation reads

$$\mathrm{i}\hbar\Bigl(\frac{\partial}{\partial t}\hat{U}_0\Bigr)|\psi_\mathrm{I}(t)\rangle + \mathrm{i}\hbar\hat{U}_0\frac{\partial}{\partial t}|\psi_\mathrm{I}(t)\rangle = (\hat{H}_0 + \hat{V})\hat{U}_0|\psi_\mathrm{I}(t)\rangle \; . \tag{2.78}$$

Observing that

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\hat{U}_0 = \hat{H}_0\,\hat{U}_0 \tag{2.79}$$

and

$$\hat{U}_0\hat{U}_0^\dagger = \hat{1} \; , \tag{2.80}$$

the above equation reduces to an effective Schrödinger equation for $|\psi_\mathrm{I}(t)\rangle$

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}|\psi_\mathrm{I}(t)\rangle = \hat{V}_\mathrm{I}(t)\,|\psi_\mathrm{I}(t)\rangle \; , \tag{2.81}$$

identifying $\hat{V}_\mathrm{I}(t) = \hat{U}_0^\dagger\hat{V}(t)\hat{U}_0$. This equation has the formal solution

$$|\psi_\mathrm{I}(t)\rangle = \hat{U}_\mathrm{I}(t, t_0)\,|\psi_\mathrm{I}(t_0)\rangle \tag{2.82}$$

with the evolution equation

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\hat{U}_\mathrm{I}(t, t_0) = \hat{V}_\mathrm{I}(t)\;\hat{U}_\mathrm{I}(t, t_0) \; . \tag{2.83}$$

The corresponding dynamics for observables in the interaction picture (remember (2.77)) is then controlled by

$$\frac{\mathrm{d}\hat{A}_\mathrm{I}}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\bigl[\hat{A}_\mathrm{I}(t), \hat{H}_0\bigr] + \hat{U}_0^\dagger\,\frac{\partial\hat{A}}{\partial t}\,\hat{U}_0 \; . \tag{2.84}$$

2.5.2 Series Expansion

The formal solution (2.82) of the effective Schrödinger equation (2.81) may be written as

$$\hat{U}_I(t,t_0) = \hat{1} - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}t_1 \hat{V}_{\mathrm{I}}(t_1)\,\hat{U}_{\mathrm{I}}(t_1,t_0)\;. \tag{2.85}$$

This integral equation can be solved for $\hat{U}_{\mathrm{I}}(t,t_0)$ by iteration,

$$\begin{aligned}\hat{U}_{\mathrm{I}}(t,t_0) &= \hat{1} + \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}t_1\,\hat{V}_{\mathrm{I}}(t_1) + \left(\frac{\mathrm{i}}{\hbar}\right)^2 \int_{t_0}^{t}\mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2\,\hat{V}_{\mathrm{I}}(t_1)\,\hat{V}_{\mathrm{I}}(t_2) + \cdots \\ &= \sum_{n=0}^{\infty}\left(\frac{\mathrm{i}}{\hbar}\right)^n \int_{t_0}^{t}\mathrm{d}t_1 \int_{t_0}^{t_1}\mathrm{d}t_2 \cdots \int_{t_0}^{t_{n-1}} \mathrm{d}t_n\,\hat{V}_{\mathrm{I}}(t_1)\cdots\hat{V}_{\mathrm{I}}(t_n)\;,\end{aligned} \tag{2.86}$$

which is called the Dyson series expansion. In first order the transition probability due to $\hat{V}_{\mathrm{I}}(t)$ is given by

$$W_{ij}(t) = \left|\delta_{ij} + \frac{1}{\mathrm{i}\hbar}\int_{t_0}^{t} \mathrm{d}t_1\;\langle j|\,\hat{V}_{\mathrm{I}}(t_1)\,|i\rangle\right|^2 \;. \tag{2.87}$$

For $i \neq j$ and going back to the Schrödinger picture, we find

$$W_{ij}(t) = \frac{1}{\hbar^2}\left|\int_{t_0}^{t}\mathrm{d}t_1\;\exp\!\left(\frac{\mathrm{i}(E_j - E_i)t_1}{\hbar}\right)\langle j|\,\hat{V}(t_1)\,|i\rangle\;\right|^2 \;. \tag{2.88}$$

Let the time-dependent perturbation be

$$\hat{V}(t) = \begin{cases} 0 \\ \hat{V} \end{cases} \quad \text{for} \quad \begin{array}{l} t \leq 0 \\ t > 0 \end{array} \tag{2.89}$$

and

$$\frac{E_j - E_i}{\hbar} = \omega_{ji}\;. \tag{2.90}$$

Then we find for the transition probability

$$W_{ij}(t) = \frac{1}{\hbar^2}\left|\frac{\mathrm{e}^{\mathrm{i}\omega_{ji}t} - 1}{\omega_{ji}}\;\langle j|\hat{V}|i\rangle\right|^2 \tag{2.91}$$

$$= \frac{1}{\hbar^2}\left|\frac{\sin(\omega_{ji}\,t/2)}{\omega_{ji}/2}\right|^2\;\left|\langle j|\hat{V}|i\rangle\right|^2\;, \tag{2.92}$$

which gives Fermi's Golden Rule for large times

$$W_{ij}(t) = t\,\frac{2\pi}{\hbar}\,\delta(E_j - E_i)\left|\langle j|\hat{V}|i\rangle\right|^2\;, \tag{2.93}$$

i.e., a constant transition rate.

3 Basics of Thermodynamics and Statistics

> *Not knowing the 2nd law of thermodynamics is like never having read a work of Shakespeare.*
>
> — C. P. Snow [121]

After having introduced some central concepts, results, and equations from quantum mechanics, we will now present the main definitions and laws of phenomenological thermodynamics and of thermostatistics. The aim is not at all to give a complete overview of the concepts of classical thermodynamics, but a brief introduction and summary of this old and useful theory. For a complete exposition of the subject we refer to some standard textbooks [24, 65, 108, 131]. (A timeline of notable events ranging from 1575 to 1980 can be found in [13].)

3.1 Phenomenological Thermodynamics

In spite of the fact that all physical systems are finally constructed out of basic subunits, the time evolution of which follows complicated coupled microscopical equations, the macro state of the whole system is typically defined only by very few macroscopic observables like temperature, pressure, volume, etc. From a phenomenological point of view one finds simple relations between these macroscopic observables, essentially condensed into four main statements – the fundamental laws of thermodynamics. These central statements are based on experience and cannot be founded within the phenomenological theory.

3.1.1 Basic Definitions

A physical system is understood to be an operationally separable part of the physical world. Microscopically such a system can be defined by a Hamilton function (classically) or a Hamilton operator (quantum mechanically), whereas in the macroscopic domain of thermodynamics systems are specified by state functions. Such a state function, like the internal energy U, is defined on the space of so-called macro states.

Such a macro state of the system is defined by a complete and independent set of state variables (macro variables) Z_i, where $i = 1, 2, \dots, n_{\text{var}}$. The dimension n_{var} is small compared to the number of microscopic variables for the system under consideration. The macroscopic state variables Z_i come

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 21–36 (2004)
`http://www.springerlink.com/`

in two variants: *extensive variables*, X_i (e.g., volume V, entropy S), which double if the system is doubled, and *intensive variables*, ξ_i (e.g., pressure p, temperature T), which remain constant under change of system size. For each extensive variable X_i there is a conjugate intensive variable ξ_i, with $i = 1, 2, \ldots, n_{\text{var}}$. Starting from an all-extensive macro state, $Z_i = X_i$, one can get different representations by replacing X_j by ξ_j (for some given j). For the all-extensive macro state one usually choses the internal energy $U(X_i)$ as the appropriate state function.(Another choice would be the entropy, see below.) The coordinate transformation to other state variables, or more precisely, Legendre transformation, leads to new state functions.

There are no isolated macro systems: system and environment constitute the most fundamental partition of the physical world underlying any physical description. In the thermodynamic regime the environment can be used to fix certain state variables like volume V, temperature T, pressure p etc. The system proper is usually classified according to the allowed exchange processes with the environment. “Completely closed” means no matter-, no energy-exchange; “closed” means no exchange of matter; otherwise the system is termed “open”.

The existence of equilibrium states is taken as a fundamental fact of experience. After a certain relaxation time any completely closed macro system approaches an equilibrium state (stationary state), which the system will then not leave anymore spontaneously. The number of independent state variables becomes a minimum in equilibrium given by n_{var}. There are n_{var} state equations, relations between extensive and intensive macro variables in equilibrium, which help to specify the experimentally verifiable properties of the system.

As a thermodynamic process we consider a sequence of state changes defined in the state space of macro variables of the system and its environment. A reversible process must consist of equilibrium states only: relaxation from a non-equilibrium state to an equilibrium state is always irreversible by definition.

Moderate deviations from global equilibrium are based on the concept of local equilibrium. In this case the macro system can further be partitioned into macroscopic subsystems, which, by themselves, are still approximately in equilibrium. The local state would thus be time-independent, if isolated from the other neighboring parts.

Conventional thermodynamics is sometimes also called thermostatics, as the state changes are studied here without explicit reference to time. As a phenomenological theory thermodynamics cannot define its own range of validity. In particular, it does not give any criteria, according to which a given system should be expected to behave thermodynamically or not.

3.1.2 Fundamental Laws

To consider thermodynamic phenomena in detail, we often need, besides the macro variables and the state function, some additional quantities A, which are functions of the independent macro variables Z_i. In thermodynamic processes the total change of A over a closed cycle may not be independent of the path, i.e.,

$$\oint \delta A \neq 0 . \tag{3.1}$$

Such a quantity is non-integrable and is said to have no complete differential. Nevertheless, it is possible to define an infinitesimal change,

$$\delta A = \sum_{i=1}^{n_{\text{var}}} \frac{\partial A}{\partial Z_i} \, \mathrm{d}Z_i , \tag{3.2}$$

for which, however,

$$\frac{\partial}{\partial Z_j} \frac{\partial A}{\partial Z_i} \neq \frac{\partial}{\partial Z_i} \frac{\partial A}{\partial Z_j} . \tag{3.3}$$

Sometimes one can introduce an integrating factor for the quantity A such that the last relation is fulfilled and A becomes integrable. Furthermore two non-integrable quantities may add up to form an integrable one. State functions are always integrable. In the following $\mathrm{d}A$ will denote a complete differential, δA an infinitesimal change (not necessarily a complete differential) and ΔA a finite change of A.

Zeroth Law:

For a thermodynamic system there exists an empirical temperature T such that two systems are in thermal equilibrium, if $T^{(1)} = T^{(2)}$. Any monotonic function $f(T)$ of T can also be used as an empirical temperature.

First Law:

For any thermodynamic system the total internal energy U is an extensive state function. In a completely closed system U is constant in time,

$$\delta U = 0 . \tag{3.4}$$

U may change only due to external energy transfer: $\delta U = \delta U^{\text{ext}}$. Examples are:

- Change of volume V: $\delta U^{\text{ext}} = -p \, \mathrm{d}V$ (p: pressure).
- Change of magnetization $\boldsymbol{M}$: $\delta U^{\text{ext}} = \boldsymbol{B} \, \mathrm{d}\boldsymbol{M}$ ($\boldsymbol{B}$: magnetic field).
- Change of particle number N: $\delta U^{\text{ext}} = \mu \, \mathrm{d}N$ (μ: chemical potential).

The total contribution has the general form

$$\delta A = \sum_{i=1}^{n_{\mathrm{var}}-1} \xi_i \, \mathrm{d}X_i \; , \tag{3.5}$$

where δA is called the total applied work, X_i an extensive work variable (excluding entropy) and ξ_i the conjugate intensive variable to X_i (excluding temperature). Why just these variables X_i, no others? The answer is that there exist environments (i.e., some appropriate apparatus) such that these energy changes can actually be carried out in a controlled fashion.

For thermodynamic systems we need to have, in addition, a “heat contribution” δQ. The *first law of thermodynamics* thus reads explicitly

$$\mathrm{d}U = \delta Q + \delta A = \delta Q + \sum_{i=1}^{n_{\mathrm{var}}-1} \xi_i \, \mathrm{d}X_i \; , \tag{3.6}$$

δQ and δA do not constitute complete differentials by themselves, but their sum does. For any closed path in macro state space we thus have

$$\oint \mathrm{d}U = 0 \; , \tag{3.7}$$

which constitutes a form of energy conservation; there is no perpetual motion machine (perpetuum mobile) of the first kind, i.e., there is no periodic process in which work is extracted without supplying energy or heat. Periodic means that the machine is exactly in the same state after each cycle (ready for the next one), which is not necessarily true for the environment.

Second Law:

The first law guarantees that each process conserves the energy of the whole system (system and environment together). However, there are processes that we can never find in physics even though they would not violate the first law. According to Clausius:

> Heat never flows spontaneously from a cold body to a hotter one.

An important alternative formulation of the second law makes use of the concept of the perpetuum mobile of second kind (Thomson’s formulation):

> It is impossible to construct a periodically operating machine, which does nothing else but transforms heat of a single bath into work.

Experience tells us that the above two formulations of the second law of thermodynamics are fulfilled, in general. However, it is not possible to prove this law within the phenomenological theory of thermodynamics. In statistical mechanics there have been numerous attempts to do just this. We will introduce some of them later in Chap. 4.

For reversible processes one finds that

$$\mathrm{d}S = \frac{\delta Q}{T} \ . \tag{3.8}$$

In this case $1/T$ is an integrating factor for δQ. S is called entropy. If irreversible processes participate, the quantity is not integrable any more. In general, for arbitrary (reversible and irreversible) processes we have

$$\delta S \geq \frac{\delta Q}{T} \ . \tag{3.9}$$

As long as irreversible processes take place, entropy is increased until the system reaches equilibrium. In equilibrium S takes on a maximum value, usually constrained by some conservation laws.

The entropy of a system can thus change due to internal entropy production and external entropy transfer:

$$\delta S = \delta S^{\mathrm{int}} + \delta S^{\mathrm{ext}} \ . \tag{3.10}$$

The second law states that

$$\delta S^{\mathrm{int}} \geq 0 \ . \tag{3.11}$$

A system is called adiabatically closed, if

$$\delta S^{\mathrm{ext}} = \frac{\delta Q}{T} = 0 \quad \Longrightarrow \quad \delta S \geq 0 \ . \tag{3.12}$$

In the case of a reversible process we need to have $\mathrm{d}S^{\mathrm{tot}} = \mathrm{d}S^{\mathrm{g}} + \mathrm{d}S^{\mathrm{c}} = 0$, where $\mathrm{d}S^{\mathrm{g}}$ is the entropy change of the system, $\mathrm{d}S^{\mathrm{c}}$ the entropy change of the environment. Only under this condition can a process run backwards without violating the second law, i.e., after the reverse process everything, the system as well as the environment, is in exactly the same state as before.

It is important to note that entropy changes are measurable. For this purpose we couple the system to an external system (bath at temperature T^{c}) and perform a reversible process with $\Delta S^{\mathrm{tot}} = \Delta S^{\mathrm{g}} + \Delta S^{\mathrm{c}} = 0$. For fixed $T = T^{\mathrm{c}}$ we can identify

$$\Delta S^{\mathrm{g}} = \int \frac{\delta Q}{T^{\mathrm{c}}} \ . \tag{3.13}$$

ΔS^{g} can thus be measured via the reversible heat exchange.

Remark: the measurability of entropy changes for any individual thermodynamic system has far-reaching consequences. It excludes the possibility to consistently interpret S in terms of subjective ignorance (though one may still use this metaphor in a pragmatic way). Furthermore, in so far as quantum physical uncertainties can be shown to give rise to thermodynamic entropy, this may shed new light on the question whether quantum mechanical states could be interpreted as representing our subjective knowledge or ignorance, indicating that this is not the case. Quantum thermodynamics, as treated in this book, should thus help to clarify ongoing controversial disputes.

Third Law (Nernst):

For $T \to 0$ we have for systems without “frozen-in disorder”,

$$S \to 0 \tag{3.14}$$

independent of X_i, or ξ_i, respectively. As a consequence, specific heats go to zero for $T \to 0$. This is interpreted to imply that the zero point of the absolute temperature T cannot be reached. In support of the above remark, $S = 0$ cannot mean that we have “complete knowledge” of the respective ground state; this is hardly ever the case.

3.1.3 Gibbsian Fundamental Form

The so-called Gibbsian fundamental form now follows as a combination of the first and the second law (for reversible processes)

$$\mathrm{d}U = T\,\mathrm{d}S - \sum_{i=1}^{n_{\mathrm{var}}-1} \xi_i\,\mathrm{d}X_i \; . \tag{3.15}$$

The “natural” independent macro variables for U are thus S and X_i, which are all extensive. Euler’s homogeneity relation (definition of a complete differential)

$$\mathrm{d}U = \frac{\partial U}{\partial S}\,\mathrm{d}S + \sum_{i=1}^{n_{\mathrm{var}}-1} \frac{\partial U}{\partial X_i}\,\mathrm{d}X_i \tag{3.16}$$

allows us to identify

$$T(S, X_i) = \left(\frac{\partial U}{\partial S}\right)_{X_i} \; , \tag{3.17}$$

$$\xi_j(S, X_i) = \left(\frac{\partial U}{\partial X_j}\right)_{S, X_i \neq X_j} \; . \tag{3.18}$$

The absolute temperature T thus has the property of an empirical temperature as defined in the Zeroth Law, it is the conjugate variable to S.

3.1.4 Thermodynamic Potentials

So far we have restricted ourselves to the internal energy $U(S, X_i)$ of the system as a thermodynamic state function or potential (see (3.15)). Instead of U we may alternatively consider the entropy function, $S(U, X_i)$. Both of these basic state functions are functions of extensive variables only. Rewriting the Gibbsian fundamental form we get

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}U - \sum_{i=1}^{n_{\mathrm{var}}-1} \frac{\xi_i}{T}\,\mathrm{d}X_i \tag{3.19}$$

from which, comparing with the Euler equation (cf. (3.16)), we read

$$\left(\frac{\partial S}{\partial U}\right)_{X_i} = \frac{1}{T}\,, \tag{3.20}$$

$$\left(\frac{\partial S}{\partial X_j}\right)_{U, X_i \neq X_j} = -\,\frac{\xi_j}{T}\,. \tag{3.21}$$

However, for concrete physical situations, e.g., special contact conditions of system and environment, it is more appropriate to use a different set of independent variables. This set should be better adapted to the considered situation. The method to perform this coordinate transformation is called the *Legendre transformation*.

We start from the energy function $U(S, X_i)$, and restrict ourselves to simple systems ($n_{\mathrm{var}} = 2$) with the single work variable $X = V$. For this volume V the conjugate intensive variable is the pressure $p = -\partial U/\partial V$. The free energy F (or Helmholtz free energy) results from a Legendre transformation of the function $U(S, V)$ replacing S by its conjugate T:

$$F(T, V) = U(S(T, V), V) - \frac{\partial U}{\partial S}\,S = U - TS\,, \tag{3.22}$$

$$\mathrm{d}F = \mathrm{d}U - T\,\mathrm{d}S - S\,\mathrm{d}T = -S\,\mathrm{d}T - p\,\mathrm{d}V\,. \tag{3.23}$$

For the enthalpy H we replace V by the conjugate variable p,

$$H(S, p) = U - \frac{\partial U}{\partial V}\,V = U + pV\,, \tag{3.24}$$

$$\mathrm{d}H = \mathrm{d}U + V\,\mathrm{d}p + p\,\mathrm{d}V = T\,\mathrm{d}S + V\,\mathrm{d}p\,. \tag{3.25}$$

Finally, the Gibbs free energy (or free enthalpy) results, if we replace S and V by their respective conjugate variables,

$$G(T, p) = U - \frac{\partial U}{\partial S}\,S - \frac{\partial U}{\partial V}\,V = U - TS + pV\,, \tag{3.26}$$

$$\mathrm{d}G = -S\,\mathrm{d}T + V\,\mathrm{d}p\,. \tag{3.27}$$

All these thermodynamic potentials are equivalent. Of course there are additional potentials, if there are more work variables, e.g., magnetic variables, exchange of particle numbers and so on.

What is the use of these thermodynamic potentials? According to the second law the entropy reaches its maximum value in equilibrium. As a consequence these thermodynamic potentials will reach a minimum value for the equilibrium state of the system under specific conditions. This allows us to use these potentials to compute the properties of the system, if a calculation based on the entropy is impossible. Additionally we need these potentials in the statistical theory, as will be seen below.

3.2 Linear Irreversible Thermodynamics

Up to this point all considerations and discussions referred to equilibrium situations, i.e., situations that are reached after sufficient time if systems are left alone. These equilibrium states are time-independent, thus the theory developed so far excludes, technically speaking, all phenomena that feature a time evolution. Of course, relations like the Gibbs fundamental form (see Sect. 3.1.3) or the formulation of thermodynamic potentials are essentially meant to describe processes like adiabatic expansion, isothermal compression, etc., that surely do have a time-dependence, but all these processes are driven by the change of external parameters like volume, temperature, etc. These processes are called "quasi static" since it is assumed that the system will be immediately at rest the very moment in which the external parameter stops changing. Thus, this theory does not include any process in which a system develops without an external influence.

Such processes happen on the way to equilibrium. They are thus irreversible and much harder to deal with because the internal energy is no longer a state function of the extensive quantities. Therefore Gibbs' fundamental form of (3.15) is not necessarily valid and one needs more and more parameters to specify the state of a system.

There is, however, a class of irreversible processes that happen close to equilibrium which are, with some additional assumptions, accessible from a slightly enlarged theory called "linear irreversible thermodynamics".

The first assumption is that in such processes equilibrium thermodynamics remain locally valid, i.e., it is assumed that it is possible to divide the system into spatial cells to each of which equilibrium thermodynamics applies, only the thermodynamic quantities may now vary from cell to cell. Regarding the level of description one does not need entirely new quantities, one only needs the standard quantities for each cell and these are assumed to be small enough so that the quantities may be given in the form of smooth space (and time) dependent functions. So, from extensive quantities one goes to densities (e.g., $U \to u(\boldsymbol{q},t)$, where $\boldsymbol{q}$ is the vector of position coordinates) and from intensive quantities to fields (e.g., $T \to T(\boldsymbol{q},t)$). For the entropy density one gets (see Sect. 3.1.4)

$$s(\boldsymbol{q},t) = s\big(u(\boldsymbol{q},t)\,, X_i(\boldsymbol{q},t)\,\big) \,, \tag{3.28}$$

or, specializing in situations in which no extensive quantities other than energy and entropy vary,

$$s(\boldsymbol{q},t) = s\big(u(\boldsymbol{q},t)\,\big) \,. \tag{3.29}$$

To describe the evolution of the system one has to introduce the "motion" of the energy – the energy current $\boldsymbol{j}_u$. Since the overall energy is conserved, the current is connected with the energy density by a continuity equation,

$$\frac{\partial u}{\partial t} + \nabla \boldsymbol{j}_u = 0 \; . \tag{3.30}$$

If one had an equation connecting $\boldsymbol{j}_u$ to the functions describing the thermal state of the system like $u(\boldsymbol{q}, t)$ or $T(\boldsymbol{q}, t)$, one could insert this equation into (3.30) getting an autonomous equation describing the behavior of the system. The problem is that such an equation depends on the material and could, in principle, take on the most complicated forms. The aim of the considerations at hand is to show that, under some assumptions, this equation can just assume a form into which the properties of the specific material enter via a few constants.

Since equilibrium thermodynamics is supposed to be locally valid, one finds for the differential of the entropy density, with (3.20),

$$\mathrm{d}s = \frac{\partial s}{\partial u} \, \mathrm{d}u = \frac{1}{T} \, \mathrm{d}u \tag{3.31}$$

and thus for the entropy current $\boldsymbol{j}_s$ connected with the energy current,

$$\boldsymbol{j}_s = \frac{1}{T} \boldsymbol{j}_u \; . \tag{3.32}$$

Entropy is no longer a conserved quantity, so the local entropy production rate $\dot{s}$ is determined by

$$\dot{s} = \frac{\partial s}{\partial t} + \nabla \boldsymbol{j}_s \; . \tag{3.33}$$

Now, plugging (3.31) and (3.32) into (3.33), one finds, exploiting (3.30)

$$\dot{s} = -\frac{1}{T^2} \, \nabla T \boldsymbol{j}_u \; . \tag{3.34}$$

Demanding that $\dot{s}$ has to be positive at any point and any time, (3.34) sets restrictions on the above mentioned equation for $\boldsymbol{j}_u$.

A basic restriction deriving from another concept is the Markov assumption, i.e., the assumption that the $\boldsymbol{j}_u(\boldsymbol{q}, t)$ should only depend on the state of the system at the actual time t and not at the configurations at former times $t' < t$. This dependence thus has to be local in time. It could nevertheless, in principle, be non-local in space. $\boldsymbol{j}_u(\boldsymbol{q}, t)$ could depend on the values of, say $T(\boldsymbol{q}', t)$ at all $\boldsymbol{q}' \neq \boldsymbol{q}$ or spatial derivatives of T of arbitrarily high order. This, however, is forbidden by (3.33). $\boldsymbol{j}_u$ can only depend on first order derivatives (no higher, no lower order) of T, otherwise the positivity of $\dot{s}$ could not be guaranteed. Specializing now in cases with very small temperature gradients, one can neglect all terms in which the first order derivatives enter other than linearily, eventually finding

$$\boldsymbol{j}_u(\boldsymbol{q}, t) = -\kappa \, \nabla T(\boldsymbol{q}, t) \; . \tag{3.35}$$

The explicit form of κ depends on the material. This is the above mentioned equation that allows, together with (3.30), for a closed description of linear irreversible processes (equation of heat conduction). Equation (3.35) is also known as Fourier's law, and has turned out to be appropriate for describing a huge class of experiments that proceed close to equilibrium.

This concept can be generalized: the external forces F_i, like gradients of electric potentials, chemical potentials or temperature, are taken to be responsible for the respective currents $\boldsymbol{j}_i$: heat currents, diffusion currents as well as energy currents, which are not independent of each other. In general, we expect:

$$\boldsymbol{j}_i = \sum_j \mathsf{L}_{ij}\, F_j \ , \tag{3.36}$$

where L_{ij} is the matrix of transport coefficients. Due to the Onsager theorem, the matrix L_{ij} should be symmetric ($\mathsf{L}_{ij} = \mathsf{L}_{ji}$). Since the current $\boldsymbol{j}_i$ also induces an entropy flow through the system, we have to be very careful in choosing currents and forces. However, if we choose these quantities ensuring that the entropy density increases, $\dot{s} \geq 0$, while the currents $\boldsymbol{j}_i$ flow, it follows from the continuity equation for the entropy that

$$\dot{s} = \sum_i \boldsymbol{j}_i\, F_i \tag{3.37}$$

and, as a further consequence, we find Onsager's theorem fulfilled (see [78]).

3.3 Statistics

All considerations so far were aimed at a macroscopic description of large systems in contact with different types of environments. To investigate the behavior of such systems, we have introduced the central laws of thermodynamics – phenomenological restrictions for the development of a system, without any fundamental justification. These laws, and especially the second one, are exclusively founded on experience obtained in a large variety of different thermodynamic experiments. This phenomenological theory of thermodynamics does not allow for the calculation of a thermodynamic potential from a microscopic picture; instead, the potential must be found from experiment.

The connection of the macroscopic behavior of a given system to the microscopic time evolution, according to a classical Hamilton function or to a quantum mechanical Hamiltonian, did not enter our considerations yet, and will be the main subject of this book. Nevertheless, even if it might not provide any explanation, the theory of statistical physics, provides a "recipe" for how to calculate an equilibrium entropy from a microscopic picture. This recipe together with the above mentioned phenomenological theory enables us to calculate all sorts of equilibrium behavior. This way of dealing with thermodynamic phenomena has sometimes been compared to "driving a car without knowing how an engine works", nevertheless it is an important technique and therefore briefly described in the following.

3.3.1 Boltzmann's Principle, A Priori Postulate

Boltzmann postulated the following connection between the thermodynamic entropy and the micro state of an isolated system (all extensive state variables, internal energy U, volume V, and particle number N are fixed from the outside), as

$$S = k_{\mathrm{B}} \ln m(U, V, N) \ , \tag{3.38}$$

where k_{B} is the so-called Boltzmann constant and m is the number of micro states accessible for the system under the given restrictions. The number of accessible micro states m is often also called statistical weight or sometimes thermodynamic weight, and we will evaluate this quantity below.

However, let us first consider a macro state of a given system. From phenomenology we have learned that the equilibrium state must have maximum entropy (the second law, see Sect. 3.1.2) and thus, according to Boltzmann, this state should also belong to a maximum number of micro states. For illustration, think of a gas in a container: the states of maximum entropy are states where the gas particles are equally distributed over the whole volume and, of course, the number of such states is very large in comparison to the number of states, where all gas particles are in one corner of the container, say.

The entropy defined above is an extensive quantity in the sense that two systems with statistical weights $m^{(1)}$ and $m^{(2)}$ have the joint weight $m^{(1)} \cdot m^{(2)}$ and the total entropy of both systems $S = k_{\mathrm{B}}(\ln m^{(1)} + \ln m^{(2)})$.

Within statistical mechanics of isolated systems we have another very important postulate – the assumption of equal *a priori* probabilities of finding the system in any one of the m possible micro states belonging to the respective macro state. As a postulate, this statement is not provable either, but, as e.g., the energy of a gas in a volume V does not depend on the position of the gas particles within the container, each of these "micro states" might, indeed, be expected to be equally likely.

This idea of assuming certain probabilities for micro states rather than calculating them led to yet another way of describing a macro state of a system, which is the so-called *statistical ensemble*. This ensemble consists of m identical virtual systems for each accessible micro state and each is represented by a point in phase space. This concept has been supported by the claim that the thermodynamic system should be *quasi-ergodic*, i.e., its trajectory would come infinitesimally close to every possible micro state within its time evolution, thus one would be allowed to replace the time average by the ensemble average. Later we will discuss in more detail the ideas behind this quasi-ergodic theorem and the problems we have to face after its introduction (see Sect. 4.2).

We can now describe the state of a system by the density of points in phase space belonging to the statistical ensemble. This density $W(\boldsymbol{q}, \boldsymbol{p}, t)$ contains the probability of finding a point in phase space at position $(\boldsymbol{q}, \boldsymbol{p})$ at time t. According to the a priori postulate this probability should be constant within the respective energy shell (see below, (3.40)) and elsewhere zero

$$W(\boldsymbol{q}, \boldsymbol{p}, t) = \begin{cases} \frac{1}{m} = \text{const.} & E < H(\boldsymbol{q}, \boldsymbol{p}) < E + \Delta E \\ 0 & \text{else} \end{cases} . \tag{3.39}$$

The statistical ensemble defined by this special density is called the *microcanonical ensemble* (see Fig. 3.1(b)).

3.3.2 Microcanonical Ensemble

The microscopic behavior of any N particle system is described by the respective Hamilton function $H(\boldsymbol{q}, \boldsymbol{p})$, dependent on all generalized coordinates of the system. A micro state of the system is then represented by a point in the systems phase space, the $6N$ dimensional space spanned by all position and momentum coordinates of the N particles. For an isolated system, a system which does not exchange any extensive variable like energy, volume etc. with the environment, the Hamilton function defines an energy surface $H(\boldsymbol{q}, \boldsymbol{p}) = U$ in the phase space. The state evolution is therefore constrained to this hypersurface in phase space. Since the total isolation of a system is a very idealized restriction, let us consider in the following not completely isolated systems, for which the internal energy is fixed only within a small interval

$$E < H(\boldsymbol{q}, \boldsymbol{p}) = U < E + \Delta E . \tag{3.40}$$

The representing trajectory of the system is then restricted to an energy shell of the thickness ΔE in phase space, contrary to the restriction to an energy surface in the case of total isolation.

To exploit Boltzmann's postulate we need to know the number of micro states m in such an energy shell of the respective phase space. Usually we divide the phase space into cells of the size h^{3N}, arguing that in each cell there is exactly one micro state of the system. This assertion is reminiscent of a quantum state in phase space, due to the uncertainty relation. However, we could also have introduced an abstract division into cells. In any case, the number of states should be the phase space volume of the respective energy shell in phase space divided by the cell size.

The total volume of the phase space below the energy surface $H(\boldsymbol{q}, \boldsymbol{p}) = E$ is given by the volume integral

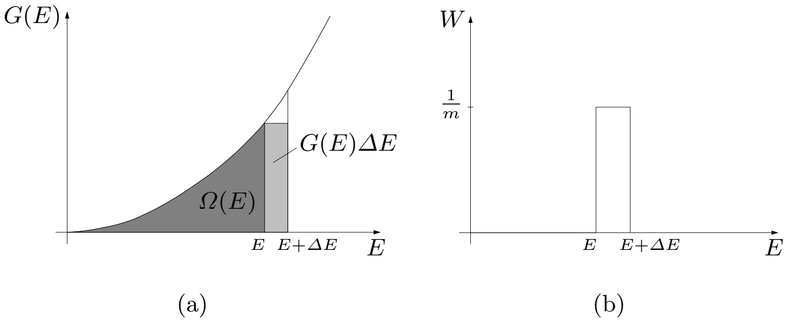


Fig. 3.1. (**a**) state density $G(E)$ (**b**) probability $W(p,q)$ for a microcanonical ensemble

$$\Omega(E) = \iint\limits_{H(\boldsymbol{q},\boldsymbol{p})\leq E} \prod_{\mu=1}^{N} \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p} \,, \tag{3.41}$$

and the volume of the energy shell by $\Omega(E+\Delta E) - \Omega(E)$. The latter can directly be evaluated by the volume integral

$$\Omega(E+\Delta E) - \Omega(E) = \iint\limits_{E<H(\boldsymbol{q},\boldsymbol{p})<E+\Delta E} \prod_{\mu=1}^{N} \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p} \,. \tag{3.42}$$

For further reference we also define here an infinitesimal quantity, the state density (cf. Fig. 3.1(a))

$$G(E) := \lim_{\Delta E\to 0} \frac{\Omega(E+\Delta E) - \Omega(E)}{\Delta E} \tag{3.43}$$

$$= \frac{\mathrm{d}\Omega(E)}{\mathrm{d}E} \,. \tag{3.44}$$

Finally the number of micro states in the respective energy shell, according to the above argumentation, is

$$m = \frac{\Omega(E+\Delta E) - \Omega(E)}{h^{3N}} \,. \tag{3.45}$$

We thus find in linear approximation for small ΔE

$$m \approx \frac{\Delta E\, G(E)}{h^{3N}} \,, \tag{3.46}$$

where $G(E)$ is the state density (3.43) at the energy surface $H(\boldsymbol{q},\boldsymbol{p}) = E$, which we have assumed does not change much in the small interval ΔE. The Boltzmann entropy now reads

$$S = k_{\mathrm{B}} \ln m = k_{\mathrm{B}} \ln \frac{\Delta E \, G(E)}{h^{3N}} \; . \tag{3.47}$$

In most cases ΔE can be considered a constant independent of E. As explained later this is not true in all cases (see Sect. 4.5).

From (3.20) and the entropy definition we are then able to define a temperature of an isolated system in equilibrium by the state density at the energy E,

$$\frac{1}{T} = \frac{\partial S}{\partial E} = k_{\mathrm{B}} \frac{\partial}{\partial E} \ln \frac{\Delta E}{h^{3N}} G(E) = \frac{k_{\mathrm{B}}}{G(E)} \frac{\partial G(E)}{\partial E} \; . \tag{3.48}$$

Due to this result the statistical temperature corresponds to the relative change of the state density with the energy.

So far we have restricted ourselves to a microcanonical ensemble where all possible micro states are equally likely. Because of the assumed isolation of the system this ensemble is not very well adapted for a variety of experimental situations. Therefore we extend our considerations to a more general exchange concept, with more detailed information about the micro state of the system.

3.3.3 Statistical Entropy, Maximum Principle

Firstly, we define a new quantity

$$S'(W_i) = -k_{\mathrm{B}} \sum_i W_i \ln W_i \; , \tag{3.49}$$

where distinguishable states of the system, whatever they might be, are labeled by i and W_i is the probability of finding the system in state i. Originally, this definition was proposed by the information theoretician Shannon, who intended to measure lack of knowledge by this function. In thermostatistics the probabilities W_i are the probabilities for finding the system in a micro state with energy E.

To find the best guess about the probability distribution, provided one knows some property of the system for sure, one has to compute the maximum of S' with respect to the W_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 S. This scheme is often referred to as Jaynes' principle introduced in [56, 57].

Again, if all extensive quantities, like volume and internal energy of a system are fixed (microcanonical situation), one has to maximize S' over all states featuring this energy and volume, which are all states from the accessible region. The only macro condition to meet is the normalization of the distribution

$$\sum_i W_i - 1 = 0 \; . \tag{3.50}$$

As expected from our former considerations, in this case (isolated situation) a uniform distribution (see (3.39)) over all those states results, as claimed in the a priori postulate. Therefore definition (3.49) meets the definition of Boltzmann in the case of a microcanonical situation, if we introduce (3.39) as the respective probability distribution of the microcanonical ensemble.

Much more interesting are other contact conditions, e.g., canonical ones. In a canonical situation energy can be exchanged between system and environment – the system is in contact with a heat bath. As an additional macro condition we require that the mean value of the energy is equivalent to the internal energy of the system U, given by

$$\sum_i W_i E_i = U \; . \tag{3.51}$$

Of course the normalization condition (3.50) should also be obeyed. Now, we maximize S' with respect to the W_i's under observance of both these conditions required for the variation

$$\delta\left(\sum_i W_i \ln W_i + \alpha\Big(\sum_i W_i - 1\Big) + \beta\Big(\sum_i W_i E_i - U\Big)\right) = 0 \; , \tag{3.52}$$

with the Lagrange multipliers α and β. From this variation we find the probability distribution

$$W_i = \frac{1}{Z}\, \mathrm{e}^{-\beta E_i} \quad \text{with} \quad Z = \mathrm{e}^{1+\alpha} \; , \tag{3.53}$$

called the Boltzmann distribution, where we have introduced the partition function Z instead of the Lagrange multiplier α.

It remains to evaluate the two Lagrange multipliers. By introducing the result of the variation in the condition (3.50), we find

$$Z = \sum_i \mathrm{e}^{-\beta E_i} \; . \tag{3.54}$$

For the second Lagrange multiplier we start from the entropy definition introducing the distribution

$$S = -k_{\mathrm{B}} \sum_i W_i \ln W_i = k_{\mathrm{B}} \ln Z \underbrace{\sum_i W_i}_{=1} + k_{\mathrm{B}}\beta \underbrace{\sum_i W_i E_i}_{=U} \; . \tag{3.55}$$

We thus get

$$U - \frac{1}{k_{\mathrm{B}}\beta} S = -\frac{1}{\beta} \ln Z \; . \tag{3.56}$$

The left hand side is the free energy F (see (3.22)), if we identify β and Z, respectively, as

$$\beta = \frac{1}{k_\mathrm{B}T} \quad \text{and} \quad F = -k_\mathrm{B}T \ln Z \; . \tag{3.57}$$

Note that there is no way to avoid this ad hoc identification if one wants to get a connection to phenomenological thermodynamics. In the same way other extensive quantities allowed for exchange can be handled, again yielding results which are in agreement with experiments.

We have thus found a recipe to evaluate the thermodynamic potentials and therefore the entropy of the system only by microscopic properties. These properties are specified by the Hamilton function of the system entering the partition function. If one is able to evaluate the logarithm of the partition function, all other thermodynamic properties, state equations, intensive parameters, etc., follow from phenomenological considerations.

However, a complete derivation of thermodynamics from microscopical theories is still missing. As already mentioned the above statistical considerations are only recipes for concrete evaluation of thermodynamic behavior. A detailed introduction to some approaches to thermodynamics from a microscopical theory will be given in the following chapter.

4 Brief Review of Pertinent Concepts

> *Given the success of Ludwig Boltzmann's statistical approach in explaining the observed irreversible behavior of macroscopic systems . . . , it is quite surprising that there is still so much confusion about the problem of irreversibility.*
>
> — J. L. Lebowitz [71]

> *Boltzmann's ideas are as controversial today, as they were more than hundred years ago, yet they are still defended (Lebowitz 1993). Boltzmann's H-Theorem is based on the unjustifiable assumption that the motions of particles are uncorrelated before collision.*
>
> — H. Primas [105]

In spite of the fact that phenomenological thermodynamics works very well, as outlined in the previous chapter, there have been many attempts to "derive" the laws of thermodynamics from an underlying theory.

Almost all approaches of this type focus on the irreversibility that seems to be present in thermodynamic phenomena, but is most likely absent in any underlying theory. So to a large degree these approaches intend to prove the second law of thermodynamics in terms of this irreversibility. They try to formulate entropy as a function of quantities, the dynamics of which can be calculated within a microscopic picture in such a way that the entropy would eventually increase during any evolution, until a maximum is reached. This maximum value should be proportional to the logarithm of the volume of the accessible phase space (energy shell); see (3.47). Only if this limit is reached will the identification of the "microscopical entropy" with the phenomenological entropy eventually yield state equations that are in agreement with experiment. It has not been appreciated very much that there are further properties of the entropy that remain to be shown, even after the above behavior has been established (see Sect. 4.5 and Sect. 5.1).

One problem of all approaches based on Hamiltonian mechanics is the applicability of classical mechanics itself. To illustrate this, let us 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 were to remain in the Ehrenfest limit (see Sect. 2.3), i.e., if the spread of the wave packages were small compared to the structure of the potentials which the particles encounter. Those potentials are generated 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 wave packages should have a width

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 37–58 (2004)
`http://www.springerlink.com/`

smaller than $10^{-10}\,$m in the beginning. Assuming particle masses between some single and some hundred proton masses and plugging those numbers into the corresponding formulas [126], we find that the spread of such wave packages will be on the order of some meters to 100 m after one second, which means the system leaves the Ehrenfest limit on a timescale much shorter than the one typical for thermodynamic phenomena. If we demand the packages to be smaller in the beginning, their spreading gets even worse. Considering this, it is questionable whether any explanation based on Hamiltonian dynamics in phase 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.

4.1 Boltzmann's Equation and $\boldsymbol{H}$-Theorem

Boltzmann's work was probably one of the first scientific approaches to irreversibility (1866) [18]. 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 is never suddenly found to be concentrated in one corner. This seems to contradict Hamiltonian dynamics according to which any process that is possible forward in time, should also be possible backward in time.

Instead of describing a system in real space (configuration space), Boltzmann tried to describe systems in μ-space, the 6-dimensional space spanned by the positions $\boldsymbol{q}$ and the velocities $\boldsymbol{v}$ of one particle being a point-like object in a 3-dimensional configuration space. Now, to describe the state of the whole system consisting of very many, N, particles, Boltzmann did not introduce N points in this μ-space, he rather used a continuous function $f(\boldsymbol{q}, \boldsymbol{v}, t)$ meant as a sort of particle density in μ-space. His basic idea was to divide μ-space into cells on an intermediate length scale, one cell of size $\mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z$ being big enough to contain very many particles, but small compared to a length scale on which the number of particles within one cell would substantially change from one cell to the next. If such an intermediate scale could be introduced, $f(\boldsymbol{q}, \boldsymbol{v}, t)$ would simply be the number of particles in the cell around $(\boldsymbol{q}, \boldsymbol{v})$. Thus, if f was large at some point, this would simply mean that there are many particles at the corresponding point in configuration space, moving in the same direction with the same velocity.

This description is definitely coarser than the full microscopic description, since information about the exact position of particles within one cell is discarded, which will turn out to be an important point, and it excludes a class of states, namely all those for which the number of particles per cell cannot be given by a smooth continuous function.

Having introduced such a description, Boltzmann tried to give an evolution equation for this function f, which is today known as the Boltzmann equation. Since a full derivation of the Boltzmann equation is beyond the scope of this text (the interested reader will find it in [2]) we only give a qualitative account of the basic ideas and describe in some detail the assumption on which this theory relies.

For a change of f at some point in μ-space and at some time t two different mechanisms have to be taken into account: a change of f due to particles that do not interact and a change due to particles that do interact (scatter). The part corresponding to particles that do not collide with each other does not cause any major problems and results only in some sort of sheer of the function f that changes positions but leaves velocities invariant.

More problematic is the part due to particles that do interact with each other. First of all, only the case of interactions that are short ranged compared to the mean free path are considered. Due to this restriction the dynamics can be treated on the level of scattering processes, just relating incoming to outgoing angles, rather than computing full trajectories. Furthermore, and this is the most important assumption in this context, it is assumed that the particles in cells that collide with each other are uncorrelated within those cells before they scatter. This is called the "assumption of molecular chaos". To understand this assumption in more detail, we consider an infinitesimal time step of the evolution of some special function f depicted in Fig. 4.1. (We restrict ourselves here to a 2-dimensional "gas", nevertheless μ-space is already 4-dimensional, thus we have to visualize f by projections). This function corresponds to a situation with all particles concentrated in one cell in configuration space but moving in opposite directions with the same velocity. (By some "center of mass" coordinate transformation any collision process may by mapped on this one.) After a period $\mathrm{d}t$ f will look as shown in Fig. 4.2. Due to momentum and energy conservation f will only be non-zero on a circle. However, where exactly on this circle the particles end up, depends on their exact positions within the cells before the collisions. If, e.g., the particle configuration at t_0 had been such that all particles had collided head-on, there could only be particles in the marked cells in Fig. 4.2(b). This corresponds to a strong correlation of the particles before the collision. If the particles had been uniformly distributed and completely uncorrelated, the distribution of particles onto the circle at $t_0 + \mathrm{d}t$ would simply be given by the differential scattering cross section $\sigma(\Omega)$ corresponding to scattering into the respective angle Ω. (For scattering from Coulomb interaction, which is not short ranged, f would be given by the famous $\sin^4(\Omega)$ law.) This

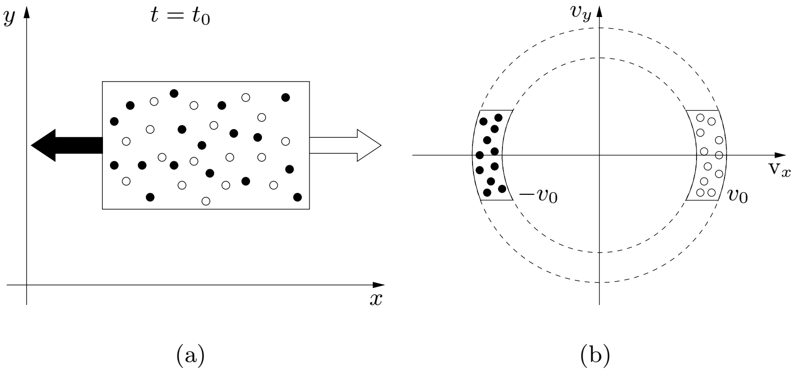


Fig. 4.1. Boltzmann equation: two-dimensional gas in μ-space at $t = t_0$ before any collision. (**a**) projection onto position space, (**b**) projection onto momentum space. White particles are flying with v_0 to the right hand side, black ones with $-v_0$ to the left.

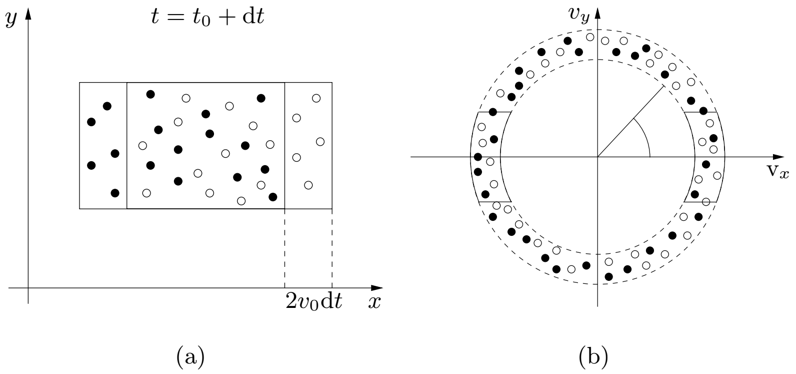
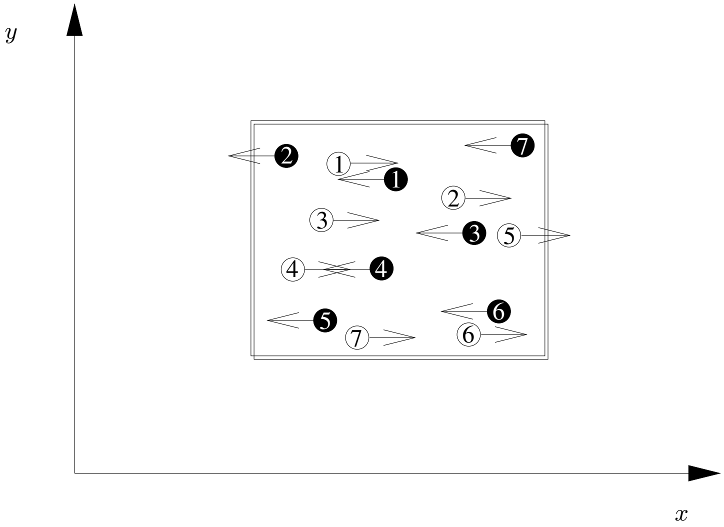


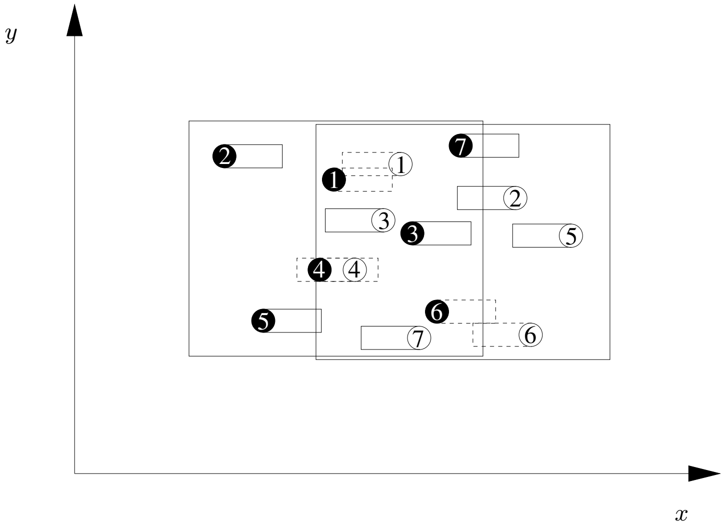
Fig. 4.2. Same as Fig. 4.1, but for $t = t_0 + \mathrm{d}t$. Because of momentum conservation all particles are now concentrated on a ring in momentum space. In the boxes at v_0 and $-v_0$ there are only particles which did not collide or which collided head-on.

is exactly what Boltzmann assumed. By simply plugging in the differential cross section for f after the collision process he could derive an autonomous evolution equation for f which no longer contains the exact positions of the particles.

This assumption, which seems intuitively appealing, has been criticized by other scientists for the following reason: even if the positions of particles within their cells were uncorrelated before the collision, they will no longer be uncorrelated after the collision. To understand this we look at the configuration of the particles before the collision more closely, see Fig. 4.3. Some



(a) $t = t_0$



(b) $t = t_0 + \mathrm{d}t$

Fig. 4.3. Boltzmann equation: (hypothetical) configuration in two cells before (**a**) and after the collision (**b**). Configurations as those depicted for particles "1" and "4" cannot result. Thus, after the collision, particle configurations in corresponding cells are correlated.

particles are going to collide head-on (pair of particles with number 4), some are going to collide on a tilted axis (e.g., pair of particles with number 1), and some are not going to collide at all during dt (e.g., pair of particles with number 3). If we ask which of those particles will still move in the initial direction after the collision, i.e., occupy the marked cells in Fig. 4.2(b), it is the particles that collided head-on (particles 4) and the ones that did not collide at all (particles 2,3,5,7). That means that within a cylinder of length $2v_0\mathrm{d}t$ on the left side of the particles moving to the right (white ones), there is either no particle (e.g., particle 5) to the left or there is exactly one in the cylinder (e.g., particle 4). Thus there are, definitely, correlations between the positions of particles in the cells marked in Fig. 4.2(b) after the collision process. The same is true for all cells on opposite positions on the ring.

To meet this criticism Boltzmann argued that there might be correlations of particles after collisions but not before collisions, since particles would collide so many times with other particles before they collided with themselves again, that in these intermediate collisions all correlations were erased. This is why the assumption of molecular chaos is also called the "Stoßzahlansatz" ("large number of collisions approach").

Exploiting the above described Boltzmann equation, Boltzmann was able to prove that the function

$$H(t) = \iint f(\boldsymbol{q}, \boldsymbol{v}, t) \log f(\boldsymbol{q}, \boldsymbol{v}, t) \, \mathrm{d}\boldsymbol{v} \, \mathrm{d}\boldsymbol{q} \tag{4.1}$$

(already very reminiscent of the entropy proper), which he formulated as a functional of f, can only decrease in time, regardless of the concrete scattering potential. Thus, in a way, irreversibility was introduced. This was the beginning of statistical mechanics.

Based on the Boltzmann equation it is possible to derive the Maxwell–Boltzmann distribution, which describes the distribution of velocities of particles in a gas, to derive the equation of state for an ideal gas, identifying the mean energy of a particle with $\frac{3}{2}k_{\mathrm{B}}T$, and even set up transport equations.

Despite the enormous success of these ideas, Boltzmann later abandoned these approaches and turned to ideas centered around ergodicity. In this way he responded to the fact that he could not get rid of the assumption of molecular chaos, which appeared unacceptable to him.

4.2 Ergodicity

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 can be seen as the time average over many hypothetical instantaneous measurements. Hence, if it were true that a trajectory ventured through all regions of the accessible volume in phase space, no matter where it started, the measured behavior of a system would be as if it were 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 ventured through during the observation time.

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

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

If this statement were taken for granted, it could be shown that the amount of time that the trajectory spends in a given volume is proportional to that volume [29]. This leads to another formulation of the ergodic hypothesis stating 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 the trajectory is a one dimensional line, whereas the so called energy surface is typically a very high dimensional volume, hence the trajectory cannot pass through all points of the energy surface in any finite time [29]. To circumvent this limitation, 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 which are quasi-ergodic in this sense and that their representing points actually spend equal time in equal phase space cells [14, 89]. This proof, however, cannot be generalized to the class of thermodynamic systems as a whole, and it remains unclear how exactly an entropy should be introduced. It has been suggested to divide phase space into finite size cells (*coarse-graining*) and simply count the cells that the trajectory passes through in a given time (see Fig. 4.4), or to count the cells weighted with the time the representing point spent within the cell.

Thus there is a fair amount of arbitrariness. Obviously a lot has to be introduced artificially, such as averaging (counting) time, cell size, etc.

4.3 Ensemble Approach

The term "ensemble" was introduced by Gibbs in about 1902 [41]. The idea is that, in general, a macroscopic observation will be consistent with a very large number of microscopic configurations. All these, represented by their corresponding points in phase space, form the "ensemble". The ensemble, therefore, is basically represented by a density in phase space which is normalized and non-zero everywhere where the system could possibly be found.

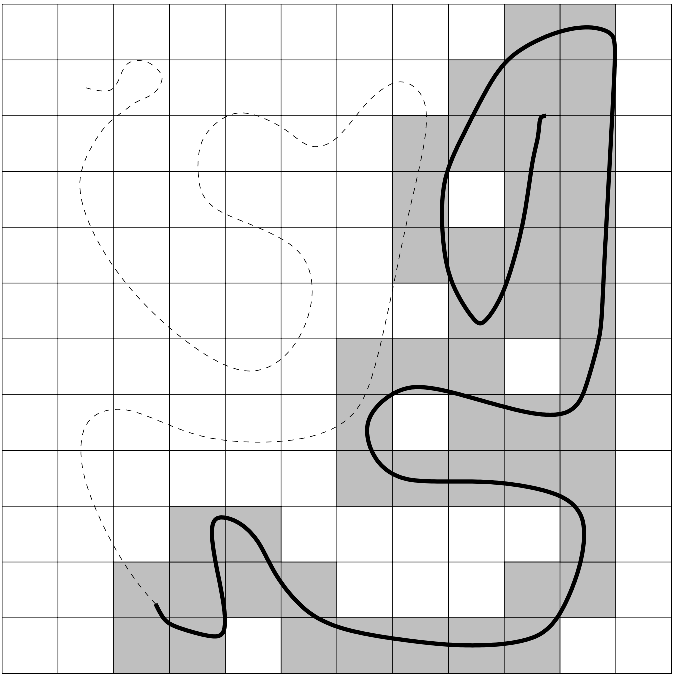
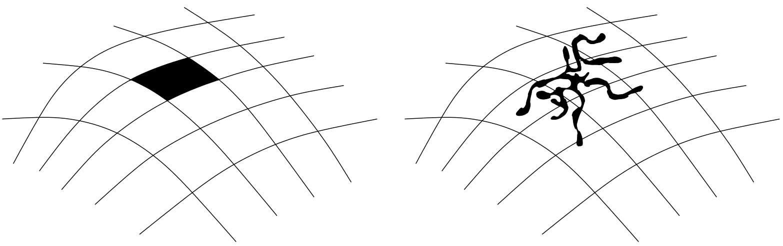


Fig. 4.4. Ergodicity approach. Entropy is defined by the volume in phase space that is occupied by the cells that the trajectory has already ventured through (gray region). This obviously depends on the observation time (solid line) and the cell size.

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 for the analysis of the evolution of such a density is Liouville's theorem. This theorem states that the volume of any region in phase 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. Therefore one could be tempted to connect the volume, in which such a density is non-zero, with the entropy. Unfortunately, the second consequence is that such a volume cannot change in time. This means that if a system does not start off in an equilibrium state, it can never reach one.

In order to save this concept, Ehrenfest and others introduced *coarse-graining* also into this idea [32]. They claimed that entropy should not be connected to the volume in which the density is non-zero, but rather to the number of cells, in which the density is non-zero somewhere. If a smooth shaped region, in which an initial density is non-zero, would then be mapped



(a) (b)

Fig. 4.5. Ensemble approach: the volume of a region in phase space cannot grow during a Hamiltonian evolution, due to Liouville's law. Nevertheless a simple initial density (**a**) can be transformed to a complicated structure (**b**) that may eventually be found in any cell, if some graining is introduced.

by the Hamiltonian evolution onto a "sponge" like structure featuring the same volume but stretched over the whole energy shell (see Fig. 4.5), such an entropy could be said to grow up to the limit, where 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 [41].

However, even if phase space played a unique role here (Liouville's theorem is only true in phase space or in 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 thermodynamic systems at all. Another objection raised against this idea is concerned with the fact that entropy seems to be here due to the observer's inability to find out in exactly which (micro) state the system is. It has been argued that this would introduce an unacceptable amount of subjectivity into this field of physics.

4.4 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 phase 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 internal energy, U, all points in phase space corresponding to a gas in this macro state would form the macroscopic cell labeled by those specific macroscopic variables, V and U. This way, phase space is not grained into equal sized Cartesian cells like in the former approaches, but into strangely shaped

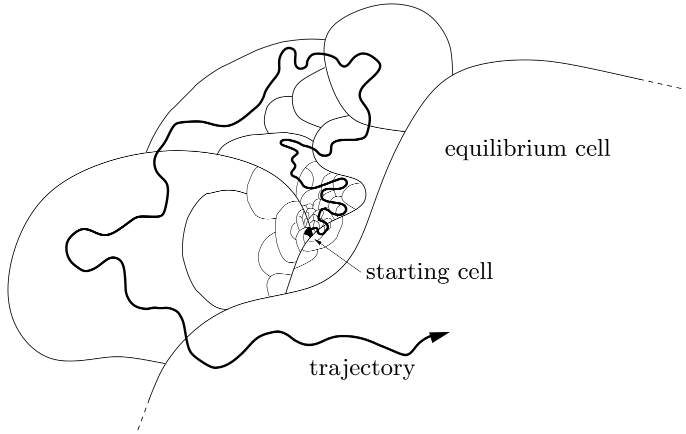


Fig. 4.6. Macroscopic cell approach: the phase space is divided into cells, according to the macro state of the system. One of these cells is much larger than all the others, the equilibrium cell. Any trajectory that is not subject to further restrictions is likely to end up in the biggest cell.

macroscopic cells that may be extremely different in size from each other (see Fig. 4.6).

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, V, would be by far the largest one, i.e., be large enough to almost fill the entire phase space. Technically this also needs to be proven for any thermodynamic system individually, but it seems much more plausible than the assumption of ergodicity or mixing.

This plausibility is connected to the so-called "law of large numbers". It is usually established by considering some abstract space (basically identified with μ-space), grained into a large number of equally sized cells, and divided into two halves, each one containing an equal number of cells (see Fig. 4.7). 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, the more drastic is this result. Transferring this result to phase space, it can be argued that almost all points in phase 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 [101].

Having established such a structure of phase space, one does not need the strict ergodicity hypothesis any longer, for if the trajectory wanders around in phase 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. 4.6).

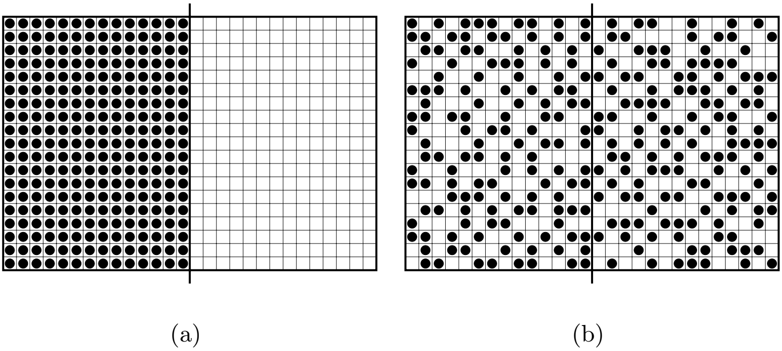


Fig. 4.7. Macroscopic cell approach: "the Law of large numbers". The vast majority of distributions of points in these cells features the same amount of points in both halves (**b**). E.g., there is only one configuration where all points are in the left half (**a**).

In such an approach, entropy is connected to the size of the cell the representing point is wandering through. However, 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. However, this only yields 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 with this idea arises if one tries to examine situations, in which the internal 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 of finding the system in a state of energy E is proportional to $\exp(-E/k_{\mathrm{B}}T)$ (Boltzmann distribution, see Sect. 3.3.3). There are attempts to explain this behavior by enlarging the phase space to contain both the system and the bath. Qualitatively the argument then is that the cells containing less energy in the considered system will have more energy in the bath (overall energy conservation) and are thus bigger than cells corresponding to higher energies in the considered system. Therefore, so the argument goes, it is more likely to find the considered system at lower energies. However, in order to derive a Boltzmann-distribution more properties are needed to make this concept work: an exponential growth of the cell sizes of the bath (see Chap. 11) and, most importantly, ergodicity of the full system. Otherwise cell sizes will not map onto probabilities. So, again, this approach only holds under the assumption of ergodicity, the very condition one tried to get rid of.

4.5 The Problem of Adiabatic State Change

One issue that has attracted much less attention in the past than the second law, is the fact that entropy should be shown to be invariant during so-called adiabatic processes [85]. As known from countless practical experiments, the state of a system controlled by the change of an external parameter, like volume or magnetic moment, proceeds in such the way that entropy is left unchanged, if the system is thermally insulated and the parameter change happens slowly enough. As obvious as this fact may seem from an experimental perspective, as surprising it is from the theoretical side. From all the possible ways in which, e.g., energy can change, exactly that one needs to be singled out that leaves a quantity as complicated as entropy unchanged! This is true for all sorts of processes that fulfill the conditions mentioned above. From a microscopical point of view this is an enormously large class of processes. For the phenomenological theory of thermodynamics it is a very important property since, otherwise, the identification of pressure with the negative derivative of energy with respect to volume under constant entropy, a basic statement of the first law, would be without justification (see Sect. 3.1.3 and especially (3.18)).

The most popular answer to this question on the basis of classical mechanics is the "law of the invariance of the phase space volume" [85]. This law is introduced by investigating a short time step of the evolution of a system while an external parameter a is changed. The considered time step is short enough to neglect the motion of any representing point of the system in phase space during that step. However, with the parameter change the "energy landscape" of the system in phase space also changes by an infinitesimal amount. This means: all points that belonged to one energy shell of energy E and thickness ΔE before the step may belong, in general, to different energy shells after the step (see Fig. 4.8).

Now the ensemble average is computed, i.e., the mean energy change $\overline{\Delta E}$ corresponding to all the points on the energy shell before the time step. Hereby it turns out (by using differential geometry) that the volume below the energy surface $\Omega_{a+\Delta a}(E+\overline{\Delta E})$ corresponding to the energy $E+\overline{\Delta E}$ defined by the "new" Hamiltonian encloses a volume in phase space that is exactly as large as the volume enclosed by the energy surface $\Omega_a(E)$ corresponding to the energy E defined by the "old" Hamiltonian,

$$\Omega_{a+\Delta a}(E+\overline{\Delta E}) = \Omega_a(E) \; . \tag{4.2}$$

For the exact definition of such phase space volumes see Sect. 3.3.2.

The idea now is that in an adiabatic process the actual representing point of the system moves much faster than the Hamiltonian changes, and that it

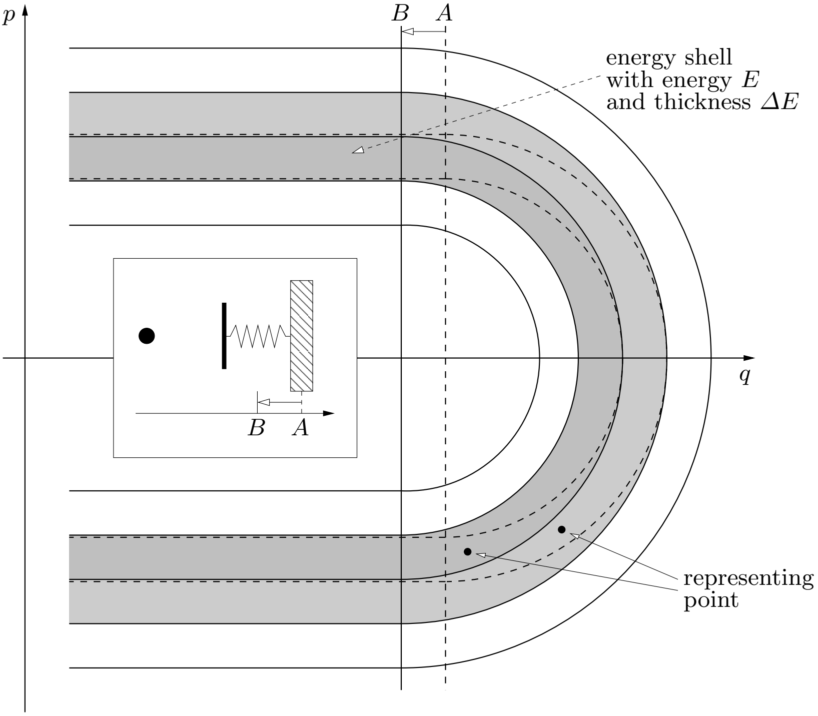


Fig. 4.8. Invariance of entropy. phase space of a particle interacting with a harmonic wall (see inset). If we were going to change an external parameter by moving the wall from position A to B, the Hamiltonian would change and therefore also the phase space. If the particle was initially within the region between the dashed lines of the phase space, it could be in the two different gray regions after the change. Two representing points are shown belonging to the same energy surface before the process, which are on different surfaces afterwards.

moves in an ergodic way. If this were the case, the system would undergo many such infinitesimal changes, while the energy changes corresponding to the points on the energy shell would remain practically the same. Within a time interval in which the representing point passes all points on the energy shell and in which the external parameter only changes by a very small amount, the system would "average itself". If the change of the external parameter were performed sufficiently slowly, the representing point would thus travel through phase space in such a way that the energy surfaces the point can be actually found on will always enclose the same volume. If one now considered the evolution of many representing points, initially on equidistant energy surfaces, it may be argued that the volumes in between two adjacent energy surfaces corresponding to two representing points cannot change, since the volumes enclosed by those energy surfaces do not change. The conclusion of this consideration is that the volume of the energy shell of a representing point does not change in a sufficiently slow process.

This reasoning faces two major shortcomings: firstly, the whole concept relies on ergodicity, a behavior that, in general, can only be postulated (see Sect. 4.2). Furthermore, the Boltzmann definition connects entropy with the volume of the energy shell (see (3.47)). This volume depends linearly on the thickness of the shell, ΔE and may be written as $\Delta E G(E)$, where $G(E)$ is the classical state density, which may be computed from the volume enclosed by some energy surface $\Omega(E)$ by $G(E) = \partial\Omega(E)/\partial E$ (see Sect. 3.3.2). The thickness ΔE is controversial, if only entropy changes are considered. However, its precise value is irrelevant, as long as it is not changed during the process. This is why entropy is usually defined simply on the basis of $G(E)$ as $S = k_{\mathrm{B}} \ln G(E)$. The latter, however, may lack invariance, as will be shown below.

Consider, e.g., a particle bound to a two dimensional rectangular surface (such a system may even be quasi-ergodic, see Fig. 4.9(a)). The state density $G(E)$ of such a system is constant, i.e., does not vary with E (see Fig. 4.9(b)). If one edge of the surface is moved out, the state density changes but remains constant with respect to E, shifted only to a higher value. Thus, although the energy changes, the system cannot have the same entropy after the process, according to the above definition. The volumes enclosed by energy surfaces corresponding to representing points at the upper and lower energy of the energy shell may nevertheless be left invariant (see Fig. 4.9(b)). The above consideration involves a change of ΔE, which is problematic, since it is usually left out of the entropy definition. This problem does not show in standard calculations simply because, contrary to our example, one has for typical systems like ideal gases $G(E) \approx \Omega(E)$. However, in principle, the problem remains unsolved (cf. Sect. 13.1).

4.6 Shannon Entropy, Jaynes' Principle

In Sect. 3.3.3 Jaynes' principle (maximization of Shannon entropy (3.49)) has been introduced basically as a recipe of how to calculate thermodynamic behavior from knowledge about microscopic structures. Nevertheless, since maximization of entropy plays a major role in this concept, it is sometimes brought forth as a justification for the second law. As explained in Sect. 3.3.3, 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 basic idea is that entropy represents 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 [56, 57]. Thus, this approach somehow radicalizes the idea that has already been underlying some of the previously described concepts, namely that thermodynamic behavior is due to the observer's inability to measure precisely and in enough detail. This

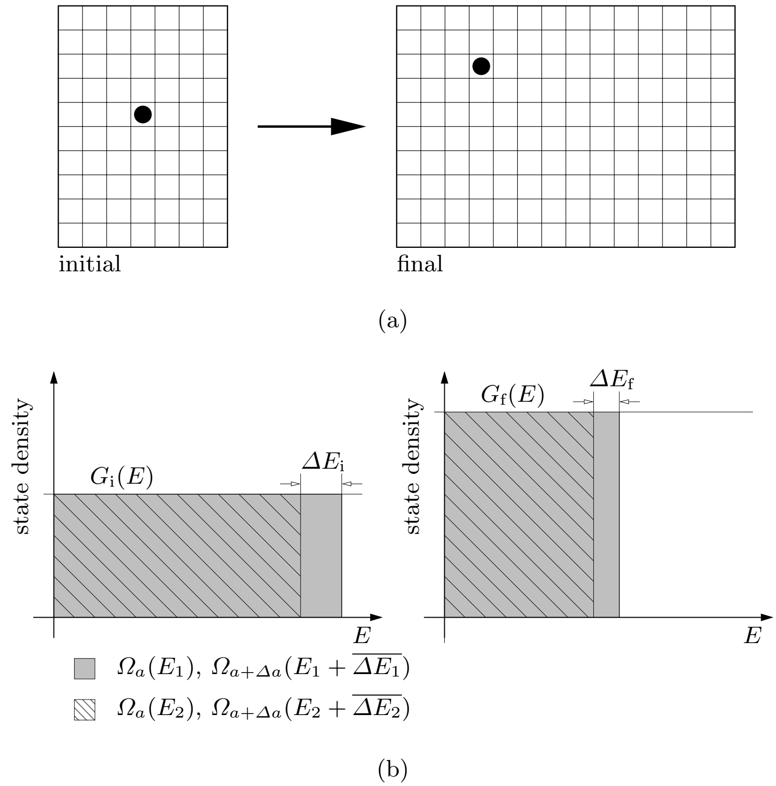


Fig. 4.9. Invariance of entropy: (**a**) Special example of a particle in a two dimensional box. (**b**) The respective state densities and the problem of different ΔE before and after the change of the external parameter (volume of a box) from a to $a + \Delta a$.

would mean that the origin of the second law was no longer searched for in the physical world but would take place more or less in the observer's brain.

And this, so the most common objection against this point of view, causes principal problems. From the fact that entropy and thus a basic thermodynamic quantity is based on the subjective lack of knowledge of the observer it must follow that, if an observer gains more information about a system, its entropy decreases and thus, e.g., its temperature decreases. This means that a change in the observer's mind could induce a macroscopic physical change of a system. This is considered unacceptable. At least as long as it is unclear whether or not there are principal limits to the observer's possibility to overcome his lack of knowledge.

All other approaches mentioned above define entropy on the basis of the dynamics of a microscopic picture, thus entropy is eventually defined as a function of time, the evolution of which is controlled by some underlying equation of motion. A reduction proper is not even attempted in Jaynes' approach, microscopic dynamics does not play any role there.

At first sight it appears to be an advantage of this approach that properties of systems under canonical (rather than microcanonical) conditions follow naturally from the basic idea. However, for this to be true in general, one has to accept that keeping an intensive quantity fixed leads to a fixed average value of the conjugate extensive quantity. For this claim no further justification is given.

Furthermore, if this principle is applied to the exchange of any other extensive quantity, the resulting states are not necessarily stationary anymore, which is inconsistent, for the resulting states should be equilibrium states.

A last objection against this theory is that the limits of its applicability are not stated clearly. Technically, this concept might be applied, e.g., to low dimensional "few body problems", for which the laws of thermodynamics are obviously not valid.

4.7 Time-Averaged Density Matrix Approach

This concept is explicitly quantum mechanical and refers to the von Neumann entropy (see Sect. 2.2.4 or [89])

$$S = -k_{\mathrm{B}} \mathrm{Tr}\left\{\hat{\rho} \ln \hat{\rho}\right\} \ . \tag{4.3}$$

Since any possible state $\hat{\rho}$ of a quantum system can be written as a weighted mixture of states that form an orthogonal set, the von Neumann 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 than the actual instantaneous one.

The elements of the time dependent density matrix read

$$\langle i|\hat{\rho}(t)|j\rangle = \exp\left(\frac{(E_i - E_j)t}{\mathrm{i}\hbar}\right) \langle i|\hat{\rho}(0)|j\rangle \ , \tag{4.4}$$

where $|i\rangle$, $|j\rangle$ are energy eigenstates and E_i, E_j the respective eigenenergies. Since all off-diagonal elements are oscillating, they will vanish if the density matrix is time-averaged [37]. Moreover, 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 thermodynamic systems. In the limit of degenerate energy eigenstates this averaging time becomes infinitely long. Thus, technically speaking, a system with a precisely given energy (microcanonical conditions) and therefore only occupying degenerate energy eigenstates, would never exhibit an increasing entropy, regardless of the averaging time.

4.8 Open System Approach and Master Equation

The open system approach is based on the fact that any description of a real system will necessarily be incomplete. This means that the subsystem selected for investigation will eventually be subject to the interaction with some unobserved and, to a large extent, uncontrollable "environment". Isolated models should thus be taken as an idealization of rather limited validity for classical as well as for quantum systems. Embeddings of the target system prevail, which do not only limit the accessible information but also drastically change the dynamics of the parts. Note, however, that the partition between system and environment does not follow from the underlying complete description; it has to be put in "by hand".

4.8.1 Classical Domain

Master equations have originally been introduced within classical many-particle physics. Here, the pertinent state space is the $6N$-dimensional phase space, in which each micro state is represented by one point with the coordinates $\{\boldsymbol{q},\boldsymbol{p}\} = \{q_1,\dots,q_{3N},p_1,\dots,p_{3N}\}$. The function of interest is the probability distribution function $\rho(\boldsymbol{q},\boldsymbol{p})$.

Following Zwanzig [137] one then assumes that a "relevant" part of ρ can be projected out according to

$$\rho_{\text{rel}}(\boldsymbol{q},\boldsymbol{p}) = \hat{P}\,\rho(\boldsymbol{q},\boldsymbol{p}) \quad \text{with} \quad \hat{P}^2 = \hat{P} \tag{4.5}$$

and

$$\rho_{\text{irrel}}(\boldsymbol{q},\boldsymbol{p}) = (1-\hat{P})\rho(\boldsymbol{q},\boldsymbol{p}) \tag{4.6}$$

denoting the "irrelevant" part. Note that this irrelevant part need not be a physically distinct subsystem: Boltzmann's relevance concept, e.g., just neglects the correlations between the considered particles. In any case, the corresponding relevant dynamics can then no longer be autonomous. The classical Liouville equation

$$\frac{\partial\rho}{\partial t} = \{H,\rho\}\ , \tag{4.7}$$

where the object on the right hand side denotes that the Poisson bracket

$$\{H,\rho\} := \sum_i \left(\frac{\partial H}{\partial q_i} \frac{\partial \rho}{\partial p_i} - \frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} \right) , \tag{4.8}$$

must be replaced by

$$\frac{\partial \rho_{\mathrm{rel}}}{\partial t} = f(\rho_{\mathrm{rel}}, \rho_{\mathrm{irrel}}) \ . \tag{4.9}$$

On the right hand side this equation contains terms depending also on the irrelevant parts of the probability distribution function. It is now of primary concern to justify approximation schemes leading back to some autonomous version of this equation of motion. Such an autonomous equation may then contain memory effects or else be local in time (Markovian),

$$\frac{\partial}{\partial t} \rho_{\mathrm{rel}}(\zeta, t) = \int G(\zeta, \zeta') \ \rho_{\mathrm{rel}}(\zeta', t) \ \mathrm{d}\zeta' \ . \tag{4.10}$$

Here, ζ denotes the respective relevant variables. For example, with the phenomenological solution

$$G(\zeta, \zeta') = W(\zeta, \zeta') + \delta(\zeta, \zeta') \int W(\zeta, \zeta'') \, \mathrm{d}\zeta'' , \tag{4.11}$$

with $W(\zeta, \zeta')$ describing transition probabilities, one finds the well known rate equation

$$\frac{\partial}{\partial t} \rho_{\mathrm{rel}}(\zeta, t) = \int \Big(W(\zeta, \zeta') \, \rho_{\mathrm{rel}}(\zeta', t) - W(\zeta', \zeta) \, \rho_{\mathrm{rel}}(\zeta, t) \Big) \, \mathrm{d}\zeta' \ . \tag{4.12}$$

This so-called *master equation* describes the time dependent change of the probability distribution function at position ζ via a transition rate from all other positions into ζ minus a transition from ζ to all other positions. Under the action of this equation of motion the entropy of the system will, in general, increase (remember Boltzmann's H-Theorem Sect. 4.1).

4.8.2 Quantum Domain

System and environment have to be treated quantum mechanically. We assume the total state to be pure and to be described by a density operator $\hat{\rho} = |\psi\rangle\langle\psi|$ defined on the corresponding Liouville space (here taken to be discrete and finite). The discrimination between relevant and irrelevant parts, again, does not follow from the underlying complete description, it is typically motivated by a certain partition between system proper and environment. In the following discussions we will label all quantities of the system with "g" and quantities of the environment with "c", according to the thermodynamic model of "gas" and "container". In this section this analogy does not have any real impact, we only use it for consistency. The respective density operators can then be identified as reduced density operators

$$\hat{\rho}_{\text{rel}} = \text{Tr}_{\text{c}}\{\hat{\rho}\} = \hat{\rho}^{\,\text{g}}\,, \tag{4.13}$$
$$\hat{\rho}_{\text{irrel}} = \text{Tr}_{\text{g}}\{\hat{\rho}\} = \hat{\rho}^{\,\text{c}}\,, \tag{4.14}$$

where the density operators of the individual parts are no longer pure and the dimension of the Liouville space of the system g is $(n^{\text{g}})^2$. Again, the Liouville–von–Neumann equation

$$\text{i}\hbar\,\frac{\partial\hat{\rho}}{\partial t} = \left[\hat{H},\hat{\rho}\right] \tag{4.15}$$

has to be replaced by a non-autonomous form,

$$\text{i}\hbar\,\frac{\partial\hat{\rho}^{\,\text{g}}}{\partial t} = \hat{f}(\hat{\rho}^{\,\text{g}},\hat{\rho}^{\,\text{c}})\;. \tag{4.16}$$

Since such an equation cannot be solved one has to get back to an autonomous form. This, in turn, cannot be achieved without specializations, approximations and/or assumptions, since in such a procedure the largest part of the degrees of freedom is simply neglected. Different authors propose slightly different schemes. All of them proceed in the interaction picture and require the interaction between the subsystems to be small. Some of them, e.g., Walls and Milburn in their book "Quantum Optics" [129], construct a Dyson series and truncate at some order due to the smallness of the interaction. This might be in conflict with evaluating the result for arbitrarily long times, or even times long enough to see the effect of the container on the gas. However, taking the container to be a set of decoupled harmonic oscillators, with an eigenfrequency density suitable to fulfill the Markov assumption (see below), and assuming a special type of interaction (rotating wave approximation), an autonomous form for a two level gas system is derived:

$$\begin{aligned}\frac{\text{d}\hat{\rho}^{\,\text{g}}}{\text{d}t} &= \frac{W_{1\to0}}{2}\left(2\,\hat{\sigma}_-\,\hat{\rho}^{\,\text{g}}\,\hat{\sigma}_+ - \hat{\rho}^{\,\text{g}}\,\hat{\sigma}_+\hat{\sigma}_- - \hat{\sigma}_+\hat{\sigma}_-\,\hat{\rho}^{\,\text{g}}\right)\\ &\quad+ \frac{W_{0\to1}}{2}\left(2\,\hat{\sigma}_+\,\hat{\rho}^{\,\text{g}}\,\hat{\sigma}_- - \hat{\rho}^{\,\text{g}}\,\hat{\sigma}_-\hat{\sigma}_+ - \hat{\sigma}_-\hat{\sigma}_+\,\hat{\rho}^{\,\text{g}}\right)\;.\end{aligned} \tag{4.17}$$

This holds true for the container being in a thermal state initially. The rates $W_{1\to0}$ and $W_{0\to1}$ are such that the gas system ends up at the same temperature in which the container initially was, so this approach somehow excludes situations in which the container temperature changes during the process, and leaves the question open, how the container developed into a thermal state to begin with.

Other authors, Scully and Zubairy [118], exploit the assumption that the joint system should approximately factorize at any time since it factorized in the beginning and the interaction is weak (Born approximation). This, again, may be challenged since even weak interactions may lead to serious quantum correlations for longer times [40]. Furthermore, for pure overall initial states it is hard to see how local entropy could rise, without developing entanglement between the gas and the container system, local entropy being an entanglement measure in this case. However, by assuming the container system to be Markovian and non-changing altogether, these approaches also lead to the above equation. It should not come as a surprise that under the non-entanglement assumption environments (linearly coupled to the system proper) could even be treated classically [43].

There are also approaches [16, 79] that do not introduce any specific form of the container system and the interaction but rely on the Markov assumption, i.e., the assumption that bath correlations decay faster than the state of the gas system is changing. However, they also require a stationary bath as well as a factorizable joint system.

So all these techniques come with the remarkable advantage of being able to describe not only equilibrium states but also the way to equilibrium. However, although there are extensions, these approaches typically rely on the Born approximation and on systems being Markovian. Furthermore, state changes of the container system are not included, which introduces a certain asymmetry between system and environment. It is the environment already being in an equilibrium state that induces equilibrium in the considered system.

The most general form of an autonomous equation for the gas system, i.e., an equation that guarantees the density operator to be positive semidefinite at all times, has been introduced by Lindblad [72],

$$\frac{\partial \hat{\rho}^{\mathrm{g}}(t)}{\partial t} = -\frac{\mathrm{i}}{\hbar}\left[\hat{H}^{\mathrm{g}}, \hat{\rho}^{\mathrm{g}}\right] + \frac{1}{2}\sum_{i,j=1}^{(n^{\mathrm{g}})^2-1} A_{ij}\left(2\hat{L}_i\hat{\rho}^{\mathrm{g}}\hat{L}_j^{\dagger} - \hat{\rho}^{\mathrm{g}}\hat{L}_j^{\dagger}\hat{L}_i - \hat{L}_j^{\dagger}\hat{L}_i\hat{\rho}^{\mathrm{g}}\right) . \tag{4.18}$$

The first term is just the normal coherent time evolution of the system on its own. It could contain, and does as a rule, energy shifts that result from the interaction. The rest describes the incoherent influence of the environment, defined by the environment operators $\hat{L}_i$. These $(n^{\mathrm{g}})^2 - 1$ traceless operators together with the $\hat{1}$-operator form a complete orthogonal basis of the $(n^{\mathrm{g}})^2$ dimensional Liouville space of the relevant system and act such as to approximate the influence of the environment c. A_{ij} is a Hermitian, positive definite parameter matrix. For a complete derivation of the Lindblad equation for open systems and pertinent examples, see [73, 79].

In general, the number of independent damping parameters A_{ij} rapidly increases with n^{g}, e.g., with the number of subsystems. For a simple spin we have two rates. For a two spin system, e.g., the so-called transverse relaxation rates are based on the three Hermitian environment operators $\hat{L}_1 = \hat{\sigma}_3^{(1)}$, $\hat{L}_2 = \hat{\sigma}_3^{(2)}$, $\hat{L}_3 = \hat{\sigma}_3^{(1)} \otimes \hat{\sigma}_3^{(2)}$, which already lead to three "autocorrelated" rates κ_{11} , κ_{22} , κ_{33} and three "cross correlated" rates κ_{12} , κ_{13} , κ_{23}. The inclusion of "longitudinal rates" in terms of $\hat{\sigma}_1^{(\mu)}$, $\hat{\sigma}_2^{(\mu)}$ would further increase this number considerably, but can be neglected for typical nuclear spin pairs [61].

The master equation is invariant under the simultaneous transformation of these environment operators

$$\hat{L}_i \to \hat{\tilde{L}}_i = \sum_j \mathsf{U}_{ij} \hat{L}_i \ , \tag{4.19}$$

together with the parameter matrix

$$A_{ij} \mapsto \tilde{A}_{ll'} = \sum_{i,j} \mathsf{U}^*_{li} A_{ij} \mathsf{U}_{l'j} \ , \tag{4.20}$$

where U_{ik} is a unitary matrix. This means that different sets of damping channels can produce the same dynamics for $\hat{\rho}^{\mathrm{g}}$.

It is always possible to diagonalize A_{ij}; the diagonal terms then represent rates. For which environment operators this should be the case has to be decided by further analysis of the total system under consideration. This can be done by a direct derivation of the corresponding master equation from the complete Hamiltonian for system and environment. It then becomes clear that the coherent and incoherent part of (4.18) are not really independent, despite the simple, additive form.

The right hand side of the above master equation can be divided into several parts, the effect of a coherent and two incoherent super-operators on $\hat{\rho}^{\mathrm{g}}$:

$$\hat{\mathcal{L}}_{\mathrm{coh}}\, \hat{\rho}^{\mathrm{g}} = -\frac{\mathrm{i}}{\hbar}\left[\hat{H}^{\mathrm{g}}, \hat{\rho}^{\mathrm{g}}\right] \ , \tag{4.21}$$

$$\hat{\mathcal{L}}^{(1)}_{\mathrm{inc}}\, \hat{\rho}^{\mathrm{g}} = \sum_{i,j} A_{ij} \hat{L}_i \hat{\rho}^{\mathrm{g}} \hat{L}^\dagger_j \ , \tag{4.22}$$

$$\begin{aligned} \hat{\mathcal{L}}^{(2)}_{\mathrm{inc}}\, \hat{\rho}^{\mathrm{g}} &= -\frac{1}{2} \sum_{i,j} A_{ij} \left(\hat{\rho}^{\mathrm{g}} \hat{L}^\dagger_j \hat{L}_i + \hat{L}^\dagger_j \hat{L}_i \hat{\rho}^{\mathrm{g}} \right) \\ &= -\frac{\mathrm{i}}{\hbar} \left(\Delta\hat{H} \hat{\rho}^{\mathrm{g}} - \hat{\rho}^{\mathrm{g}} \Delta\hat{H}^\dagger \right) \end{aligned} \tag{4.23}$$

with

$$\Delta\hat{H} = -\frac{\mathrm{i}\hbar}{2} \sum_{i,j} A_{ij} \hat{L}^\dagger_j \hat{L}_i \ . \tag{4.24}$$

Equation (4.23) can thus be included in (4.21) by allowing the effective Hamiltonian to be non-Hermitian. Due to the incoherent parts there is, in general, an increase of entropy for system g. The quality of this approximation scheme – including the non-entanglement assumption – will depend on the partition chosen.

It has become popular to interpret the solution of the Lindblad equation as an ensemble of pure state trajectories for the system under consideration. This is essentially a "classical" idea. It is assumed that the non-pure state $\hat{\rho}^{\mathrm{g}}$ was due to subjective ignorance of not knowing the actual pure state. (In general, there need not exist a pure state, the apparent mixing could be due to entanglement.) However, the mere possibility of carrying out this program confirms the fact, that there is no entanglement between the involved system and environment. If the environment state $\hat{\rho}^{\mathrm{c}}$ was independent of $\hat{\rho}^{\mathrm{g}}$, which typically holds, the resulting mixture of products would remain a simple product at all times. The pure state trajectories are not unique for a given Lindblad equation; they follow from a procedure called "stochastic unraveling" [28].

5 The Program for the Foundation of Thermodynamics

> *Nobody knows, what entropy really is.*
> — J. von Neumann quoted in [128]

For a foundation of thermodynamics it is tempting to give an abstract but, nevertheless, intuitively appealing definition of entropy, such as: entropy is the number of micro realizations for a given macro state, entropy is some volume in phase space, entropy is the lack of knowledge, etc. However, then one is left with the task of showing that the quantity so defined indeed behaves in a way that is empirically observed (see Chap. 4).

Thus, thermodynamics is often introduced on the basis of some axiomatic structure spelling out how thermodynamic quantities are expected to behave (see Sect. 3.1). Usually some of the axioms deal with the notorious irreversibility of thermodynamic systems, others with the possibility of formulating the internal energy U as a state function in terms of the entropy S and other extensive variables. Starting from first principles it is evident that much of the behavior of thermodynamic quantities (including their mere existence) needs explanation.

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

The set of rules given here should neither be considered unique nor irreducible. If, e.g., ergodicity (and mixing) were assumed, the validity of the second law (Checklist 2.) and the equality of intensive variables (Checklist 4.) would follow, thus one could replace those two items by demanding ergodicity. This, however, would not be suitable for the approach at hand. The choice of properties given here is expected to be complete and appropriate for starting from the theory of quantum mechanics.

5.1 Basic Checklist: Equilibrium Thermodynamics

1. *Definition of Thermodynamic Quantities:*
   All thermodynamic 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

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 61–64 (2004)
`http://www.springerlink.com/`

of those variables. If introduced in this way, the thermodynamic quantities would "always" be well-defined, i.e., even outside equilibrium, or for systems that are not thermodynamic (see Sect. 5.2). Definitions of thermodynamic quantities are given in Chaps. 6, 12 and 13.

2. *Second Law of Thermodynamics (Maximum Principle for Entropy):*
   This axiom establishes the irreversibility of some thermodynamic processes (i.e., which processes are possible, which are not). It postulates the existence of a certain stationary equilibrium state, into which a thermodynamic system will evolve eventually. Under given constraints this equilibrium state is stable with respect to perturbations. It should be shown that the system considered reaches a state for which the fluctuations of all well defined thermodynamic quantities are negligible. This state has to be controllable by macroscopic constraints. Since those constraints can be imposed 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 distribution of the initial state. (This behavior is established in Sect. 9.1.)
   b) *Canonical conditions* (temperature T kept fixed):
      Since under these conditions the equilibrium state of the system is controlled by the contact with a heat bath, the only specifying parameter of which is its temperature, the equilibrium state should only depend on this temperature, regardless of its initial state. (This behavior is established in Sect. 9.2.)
3. *Gibbsian Fundamental Form (Energy Conservation, State Functions):*
   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 V and (equilibrium) entropy S, take on certain values, the internal energy U necessarily has a corresponding value, regardless of the path by which the state has been reached.
   b) *Temperature as a Conjugate Variable:*
      It should be shown that there are processes (heating, cooling, etc.), in which all extensive variables are kept fixed except for energy and entropy, which then should be shown to depend on each other, according to

$$\left(\frac{\partial U}{\partial S}\right)_{V=\text{const.}} = T \,, \tag{5.1}$$

where of course the same definition of temperature T as above has to be used. (This behavior is established in Sect. 12.3.)

  c) *Pressure as a Conjugate Variable:*
  It should be shown that there are processes (isentropic), in which the extensive variable volume V, changes, while all others, including especially entropy, remain constant. The analysis of such a process then has to yield,

$$\left(\frac{\partial U}{\partial V}\right)_{S=\text{const.}} = -p\ , \tag{5.2}$$

  where p is the pressure. (This behavior is established in Sect. 13.1.)
  d) *Other Conjugate Variables:*
  It should be shown that there may be processes, in which some additional extensive variable changes, while all others remain constant. The derivative of U should yield the respective conjugate intensive variable.

4. *Classes of Thermodynamic Variables:*
  a) *Extensive Variables:*
  It should be shown that thermodynamic variables that are claimed to be extensive, in particular the entropy S, are indeed extensive quantities. (This is shown in Chap. 11)
  b) *Intensive Variables:*
  It should be shown that two systems allowed to exchange some extensive quantity will end up in an equilibrium state having the same conjugate intensive variable, for which, of course, the same definitions as used in 3. have to be valid. (This behavior is established in Sect. 12.2 and Sect. 13.2.)

Those properties of thermodynamic 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 from quantum mechanics, the field of thermodynamics can, in some sense, be considered as reducible to quantum mechanics.

Such a reconstruction, 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 as its connection to the standard classical theory. This is the subject of the supplementary checklist.

5.2 Supplementary Checklist: Quantum Mechanical Versus Classical Aspects

1. *Thermodynamic Systems:*
   It is necessary to explain and clarify the relationship between the emerging theory and its underlying theory. If the emerging properties would inevitably result from the underlying theory, one could discard the latter completely. In the text 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, while Schrödinger-type quantum theory is believed to underlie all sorts of non-relativistic systems. Thus, a fairly precise definition of the class of systems that 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 thermodynamic systems", but in a criterion that can be checked with acceptable efforts.
2. *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 would eventually amount to showing that quantum state density and the volume of energy shells in classical phase space are proportional for large classes of systems (that have a classical analog). Such a relation would prove a kind of "correspondence principle".

6 Outline of the Present Approach

> *One person's mystery is another person's explanation.*
> — N. D. Mermin quoted in [103]

As already indicated we want to derive the properties of thermodynamic quantities from non-relativistic quantum mechanics, i.e., from starting with a wave function to describe some isolated quantum system the evolution of which 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 classical Hamiltonian dynamics supply only the background for the considerations described in Chap. 4, the analysis of structures in phase space. Similarly, our present approach will essentially be based on the analysis of structures in Hilbert space.

6.1 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 (Sect. 2.2.4; [89]), which is defined for a system on the basis of its density operator.

As already explained in Sect. 2.4 this entropy is invariant with respect to unitary transformations. Since the von Neumann equation (2.49) gives rise to a unitary evolution for the density operator $\hat{\rho}$, this entropy can never change and is thus not a candidate for the thermodynamic entropy with all its required properties.

If one deals with bipartite systems, i.e., systems that can "naturally" be decomposed into two parts, the observables come in two classes: those defined locally (energy in subsystem 1, energy in subsystem 2, positions of the particles of subsystem 1, positions of the particles of subsystem 2, etc.), and those defined globally (total energy etc.). It is convenient to organize the full Hilbert space of the compound system as a product Hilbert space as explained in Sect. 2.2.5. In this case a valid and complete description of, say, subsystem 1 (in terms of any of its local observables) is given by the reduced density operator $\hat{\rho}_1$, rather than by the density operator of the full system $\hat{\rho}$. Contrary to the entropy of the total system, the entropy of subsystem 1 defined according to (2.42) on the basis of $\hat{\rho}_1$ can very well change under unitary transformations of the compound system. It would not change if the unitary transformation generated by the von Neumann equation factorized as $\hat{U}(t) = \hat{U}^{(1)}(t) \otimes \hat{U}^{(2)}(t)$ which, in turn, would be the case if the Hamiltonian

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 65–69 (2004)
`http://www.springerlink.com/`

of the full system did not contain any interactions between the subsystems. However, if the Hamiltonian contains such interactions, regardless of how small they might be, the subsystem entropy is no longer a conserved quantity and will, in general, change in time [39].

For our considerations it will be assumed that the state of the compound and isolated system (our "quantum universe") is always a pure state, i.e., a state of zero entropy. This implies that (for us) the state does not represent subjective knowledge but has ontological qualities like the (often unknown) micro states of classical few- or many-partite systems (cf. Sect. 6.2). If for such a state the entropy of a subsystem also vanished, the state would have to be a product state, i.e., a state that could be written as the product of two separate wave functions: $|\psi\rangle = |\psi^{(1)}\rangle \otimes |\psi^{(2)}\rangle$. If the subsystem entropy did not vanish, the state could not have this product form, which means it would be entangled.

In our theory the von Neumann entropy of the reduced density operator, which describes the system under consideration, will be identified with 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 the environment of the thermodynamic system, which must be considered indispensable. So one of the basic ideas is that there is no thermodynamic 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, e.g., one analyzes a gas, the gas could only be a thermodynamic system if confined to some volume by a container, which thus represents the environmental subsystem. Hence, if a gas relaxes towards equilibrium, it would do so due to increasing entanglement with the container (or other external 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) in the case of an isolated system. However, in the context of thermodynamics this isolation only means 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 at all between the system and its environment. In particular, this is not to be confounded with a microscopically closed system. Gas particles within a thermally insulating container nevertheless interact with the particles that make up the container walls, otherwise the gas 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 [39]. Thus even for a thermally closed system the existence of an environment and the interaction with it is indispensable.

6.2 Fundamental and Subjective Lack of Knowledge

It is often stated that entropy should somehow be a measure for the lack of knowledge. However, 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 indeed influence the entropy and the resulting thermodynamic behavior of real physical systems.

Within classical mechanics lack of knowledge may always be considered subjective: in principle, any observable could be known with "unlimited" precision. This is fundamentally 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 is a pure state, hence the fundamental lack of knowledge does not grow. However, for compound systems there are pure states for which 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, in the latter case, the fundamental lack of local knowledge is considerably larger than in the former case. Those states are the entangled states mentioned in Sect. 6.1 [9, 16]. 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 local observable [39].

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 thermodynamic 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.

6.3 The Natural Cell Structure of Hilbert Space

Within the context of classical mechanics, it has often been pointed out that it is 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 been impossible to formulate the von Neumann entropy as an observable in the sense of an entropy operator for which

$$S = \langle\psi|\hat{S}|\psi\rangle \quad \text{or} \quad S = \mathrm{Tr}\{\hat{S}\hat{\rho}\} \; . \tag{6.1}$$

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

$$S = -k_{\mathrm{B}} \operatorname{Tr}_{\mathrm{g}} \Big\{ \operatorname{Tr}_{\mathrm{c}} \{ |\psi\rangle\langle\psi| \} \ln \big(\operatorname{Tr}_{\mathrm{c}} \{ |\psi\rangle\langle\psi| \} \big) \Big\} . \tag{6.2}$$

Here the index c denotes the "container", i.e., the environment subsystem and g the considered system or "gas". 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 Hilbert space of the full system into different cells, each characterized by some local entropy. All states featuring, according to the above definition (6.2), the same entropy are grouped together to form one cell. This cell structure is thus uniquely and unambiguously specified.

Just like the point representing a classical system wanders around in phase space, the state vector of a quantum mechanical system wanders around in Hilbert space. And just like the point in phase space is confined to some energy shell, the state vector in Hilbert space 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, etc., but also depends 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 for establishing thermodynamic 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: almost all states within an accessible region feature the maximum entropy consistent with the macroscopic constraints.

The main idea here is that stationary equilibrium is not reached because the motion of the state vector in Hilbert space will cease to move at some point (on the contrary, it will be shown to move always with constant velocity), but because the state vector will eventually almost always venture inside the cell with the maximum local 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, Sect. 4.4, for which Penrose argues [101]:

> "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 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 an extremely small volume, if it is not confined to it. The system is not treated as if it were at any point at the same time or with the same probability, it is always treated as being at some absolutely concrete point in Hilbert space. Which point this exactly may be depends on the precise knowledge of the initial state, the details of the Hamiltonian of the system, and can only be calculated by (in the macroscopic case) solving the Schrödinger equation of some 10^{23} interacting particles.

Thus there will be a huge subjective lack of knowledge about where exactly the system might be found in Hilbert space. However, this is not what gives rise to the second law, and this subjective lack of knowledge is not to be confounded with the thermodynamic entropy. The entropy would not be any smaller if one knew the exact location in Hilbert space. 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 Hilbert space forever, but if both trajectories eventually wander through the same compartment of maximum entropy, the local subsystems 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 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 system considered the maximum entropy state is an attractor.

7 System and Environment

> *To understand quantum dynamics it is essential to realize that observable systems are always subsystems of a larger one – ultimately of the whole universe. The notion of an isolated system makes sense only as an approximation depending on the concrete situation.*
>
> — J. Kupsch [68]

In a typical thermodynamic situation we consider a bipartite system with the larger part being called "environment" or "container", c, whereas the smaller part is the observed "system" or "gas", g. Mostly we think of a very large environment so that its energy does not change much when the system g feeds, say, energy into it during the time evolution. However, also for smaller environments we are looking for a thermal behavior of the considered system, inquiring about the constraints for such a behavior.

7.1 Partition of the System and Basic Quantities

We assume that the full Hamiltonian can be partitioned in the following way

$$\hat{H} = \hat{H}^{\mathrm{g}} + \hat{H}^{\mathrm{c}} + \hat{I} \,, \tag{7.1}$$

where $\hat{H}^{\mathrm{g}}$ and $\hat{H}^{\mathrm{c}}$ are the local Hamiltonians of the system (gas g) and the environment (container c), respectively, which act on two different parts of a product Hilbert space (see Fig. 7.1). Thus these two Hamiltonians commute,

$$\left[\hat{H}^{\mathrm{g}}, \hat{H}^{\mathrm{c}}\right] = 0 \,. \tag{7.2}$$

The interaction $\hat{I}$ between the two systems constitutes some sort of weak coupling. It may allow for energy transfer or for some sort of dephasing only. It thus specifies the macroscopic or even microscopic constraints, as will be seen later. Such a partition with a weak coupling might require a reorganization of the pertinent description. In the following we write the Hamiltonian of the whole system in the product eigenbasis of the gas and the container.

Let us introduce our nomenclature for such a bipartite system as shown in Fig. 7.1. The energy eigenvalues of the gas (container) system are E_A^{g} (E_B^{c}), where the indices A and B, respectively, specify the different eigenenergies of the respective subsystem. To account for degeneracy, we introduce a subindex counting the states in each degenerate subspace A (B),

$$a = 1, \ldots, N^{\mathrm{g}}(E_A^{\mathrm{g}}) \quad \text{and} \quad b = 1, \ldots, N^{\mathrm{c}}(E_B^{\mathrm{c}}) \,, \tag{7.3}$$

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 71–77 (2004)
`http://www.springerlink.com/`

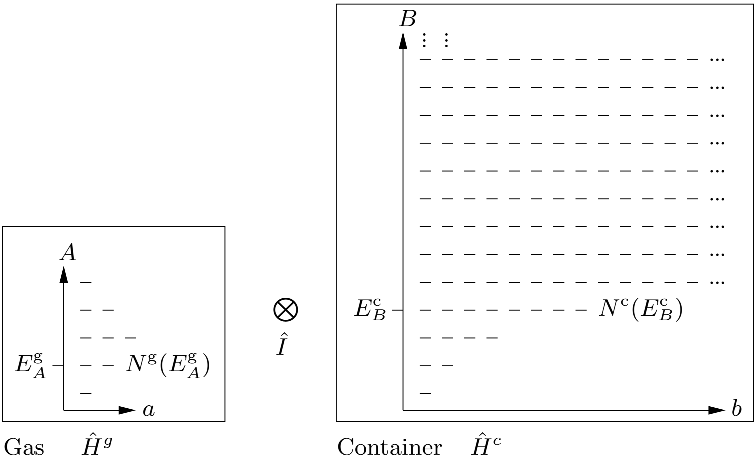


Fig. 7.1. Energy eigenvalues of a bipartite system: schematic representation of nomenclature.

where $N^{\mathrm{g}}(E^{\mathrm{g}}_A)$ $(N^{\mathrm{c}}(E^{\mathrm{c}}_B))$ are the respective degrees of degeneracy of the subspaces A (B). Sometimes we need this detailed notation of the degeneracies, but mostly it is enough to write N_A (N_B) for the degeneracy of the subspace A (B) of the gas (container) system. So $|A,a\rangle$ $(|B,b\rangle)$ denotes the a-th (b-th) energy eigenstate with the energy eigenvalue E^{g}_A (E^{c}_B) of the gas (container) system, as depicted in Fig. 7.1.

A pure state of the full system will be denoted as a superposition of product eigenstates $|A,a\rangle \otimes |B,b\rangle$,

$$|\psi\rangle = \sum_{A,B} \sum_{a,b} \psi^{AB}_{ab} |A,a\rangle \otimes |B,b\rangle \ , \tag{7.4}$$

with the complex amplitudes ψ^{AB}_{ab}. These amplitudes have to fulfill the normalization condition

$$\sum_{A,B} \sum_{a,b} \left|\psi^{AB}_{ab}\right|^2 = 1 \ . \tag{7.5}$$

In general it is not possible to write the state of one subsystem as a pure state. This is due to the entanglement between the gas and the container system (cf. Sect. 2.2.5), which will emerge during the time evolution, even if we start with a separable state in the beginning. The state of a single subsystem is thus a mixed state and must be described by a density matrix. From the density operator of the whole system $\hat{\rho} = |\psi\rangle\langle\psi|$, the reduced density operator of the subsystem, g, is found by tracing over the container system (see Sect. 2.2.5)

$$\hat{\rho}^{\mathrm{g}} := \sum_{A,A',B} \sum_{a,a',b} \psi_{ab}^{AB} \left(\psi_{a'b}^{A'B}\right)^{*} |A,a\rangle\langle A',a'| \, . \tag{7.6}$$

Analogously, it is possible to evaluate the density operator of the environment, but mostly we are not interested in this state.

The diagonal elements of the density operator are the probabilities of finding the system in the respective eigenstate. Since they are frequently needed, we introduce these quantities now. The joint probability of finding the gas system at energy E_A^{g} and at the same time the container system at the energy E_B^{c} is given by

$$W_{AB} = W(E_A^{\mathrm{g}}, E_B^{\mathrm{c}}) := \sum_{a,b} \left|\psi_{ab}^{AB}\right|^2 \, . \tag{7.7}$$

In the special case of a product state of the whole system, the joint probability is a product of the single probabilities W_A (W_B) of finding the gas (container) at the energy E_A^{g} (E_B^{c}),

$$\begin{aligned} W_{AB} = W(E_A^{\mathrm{g}}, E_B^{\mathrm{c}}) &:= \sum_{a,b} \left|\psi_a^{A}\right|^2 \left|\psi_b^{B}\right|^2 \\ &= W(E_A^{\mathrm{g}}) W(E_B^{\mathrm{c}}) = W_A W_B \, . \end{aligned} \tag{7.8}$$

Another important quantity is the probability of finding the complete system at the energy E. This probability is a summation of all possible joint probabilities $W_{AB} = W(E_A^{\mathrm{g}}, E_B^{\mathrm{c}})$ under the subsidiary condition of overall energy conservation,

$$W(E) = \sum_{A,B/E} W_{AB} \, , \tag{7.9}$$

here $A, B/E$ stands for a summation over all A, B such that $E_A^{\mathrm{g}} + E_B^{\mathrm{c}} = E$.

Another important quantity that will be considered in the following is the purity (see Sect. 2.2.4). In the case at hand the purity of the considered subsystem can be written as a function of the full system state, just like entropy (see (6.2)) and reads

$$P^{\mathrm{g}} = \sum_{A,B,C,D} \sum_{a,b,c,d} \psi_{ab}^{AB} \left(\psi_{cb}^{CB}\right)^{*} \psi_{cd}^{CD} \left(\psi_{ad}^{AD}\right)^{*} \, . \tag{7.10}$$

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

7.2 Weak Coupling

In the last section we have introduced a bipartite system – an observed system or gas and an environment or container. This partition scheme is found also in standard thermodynamics, where one typically considers gases or other systems in contact with their environment. In the standard example of an ideal gas the walls of the vessel containing the gas usually appear in the calculations only as a boundary condition for the motion in phase space. The container is not treated as a system by itself. However, this is what is done in the following. Thus thermodynamic scenarios are modeled here within the scheme described in Sect. 7.1.

Typically, *weak coupling* between system and environment is assumed in standard thermodynamic considerations. Already the concept of energy being either in the system or the environment, i.e., the idea that the total energy of the full system may be computed as a sum of the energy contained in the system and the energy contained in the environment makes the assumption of a weak coupling indispensable. Otherwise none of the above ideas would hold. Furthermore, it can be shown that the concept of intensive and extensive variables relies on a weak coupling limit (cf. [30] and also [4]). In our case, the weak coupling limit also guarantees that a picture of the energy schemes of the uncoupled systems (like Fig. 7.1) still remains reasonably close to the energy scheme of the actually coupled system. As long as the interaction remains weak, the full spectrum of the coupled system will not look significantly different from the one that results from a mere convolution of the two spectra of the uncoupled systems (joint spectrum).

To quantify the weak coupling pre-condition we have to require that the mean energy contained in the interaction $\hat{I}$ be much smaller than the energy in the individual subsystem, gas and container, separately,

$$\sqrt{\langle \hat{I}^2 \rangle} \ll \langle \hat{H}^g \rangle, \langle \hat{H}^c \rangle \,. \tag{7.11}$$

This inequality for the expectation values must hold for all states that the system can possibly evolve into under given (macroscopic) constraints. If a partition according to this weak coupling scheme were impossible, the idea of system proper and surrounding would be meaningless.

7.3 Effective Potential, Example for a Bipartite System

Consider, just as an example, an ideal gas in a container. Then the sum over the kinetic energies of all gas particles μ (mass m, momentum operator $\hat{\boldsymbol{p}}^{\mathrm{g}}_{\mu}$) is the only part of the Hamiltonian that acts on the gas subspace alone,

$$\hat{H}^{\mathrm{g}} = \sum_{\mu} \frac{\hbar^2}{2m} (\hat{\boldsymbol{p}}^{\mathrm{g}}_{\mu})^2 \,. \tag{7.12}$$

$\hat{H}^{\mathrm{c}}$, the Hamiltonian of the container provides the environment that has to be present to make the gas particles a thermodynamic system. It reads

$$\hat{H}^{\mathrm{c}} = \sum_{\mu} \frac{\hbar^2}{2m} (\hat{\boldsymbol{p}}^{\mathrm{c}}_{\mu})^2 + \frac{1}{2} \sum_{\mu,\nu} \hat{V}^{\mathrm{c}}(\boldsymbol{q}^{\mathrm{c}}_{\mu}, \boldsymbol{q}^{\mathrm{c}}_{\nu}) \,, \tag{7.13}$$

where $\hat{V}^{\mathrm{c}}(\boldsymbol{q}^{\mathrm{c}}_{\mu}, \boldsymbol{q}^{\mathrm{c}}_{\nu})$ are the interactions that bind the container particles (mass m) at position $\boldsymbol{q}^{\mathrm{c}}_{\mu}$ and $\boldsymbol{q}^{\mathrm{c}}_{\nu}$ to each other to form a solid, and acts exclusively in the container subspace. Thus, as required, $\hat{H}^{\mathrm{g}}$ and $\hat{H}^{\mathrm{c}}$ commute.

Now, $\hat{I}$ contains the interactions of all gas particles with all container particles and reads

$$\hat{I} = \sum_{\mu,\nu} \hat{V}^{\mathrm{gc}}(\boldsymbol{q}^{\mathrm{g}}_{\mu}, \boldsymbol{q}^{\mathrm{c}}_{\nu}) \,. \tag{7.14}$$

This part contains the repelling interactions between the gas particles and the container particles and establishes the container as a boundary for the gas particles from which they cannot escape. Starting from first principles, the Hamiltonian has to be written in this way, especially the last part is indispensable (see Fig. 7.2).

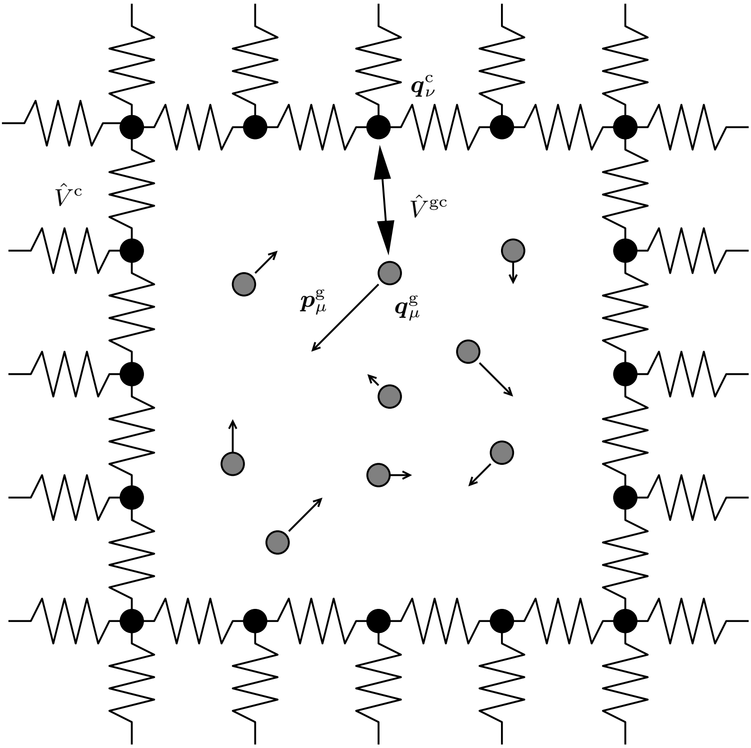


Fig. 7.2. Bipartite system: gas in a container, represented by an interacting net of particles (black dots).

Unfortunately, any stationary state of $\hat{H}^{\mathrm{g}}$, and it is such a state we want to see the system evolve into, is unbounded and thus not confined to any volume that might be given by the container. This is due to the fact that such a state is a momentum eigenstate and therefore not localized in position space. This means the expectation value of $\hat{I}$, for an energy eigenstate of the uncoupled problem, $\hat{H}^{\mathrm{g}} + \hat{H}^{\mathrm{c}}$ would definitely not be small, and thus the system would miss a fundamental prerequisite, the weak coupling, for a thermodynamic system accessible from our method.

The standard way to overcome the above mentioned deficiency is to define an effective potential for the gas particles generated by the container, in which all gas particles are trapped. Fortunately, an effective local Hamiltonian and an effective interaction can be defined so that the weak coupling limit is fulfilled, by

$$\hat{H}^{\mathrm{g}\prime} := \hat{H}^{\mathrm{g}} + \hat{V}^{\mathrm{g}} \; , \quad \hat{I}' := \hat{I} - \hat{V}^{\mathrm{g}} \; , \quad \hat{V}^{\mathrm{g}} = \hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}) \; . \tag{7.15}$$

Here, $\hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\})$ is some potential for the gas particles alone and depends on all position vectors $\{\boldsymbol{q}_\mu^{\mathrm{g}}\}$. $\hat{V}^{\mathrm{g}}$ will be chosen to minimize the weak coupling pre-condition (7.11). Substituting the real parts 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 (see Sect. 7.1). A good candidate for $\hat{V}^{\mathrm{g}}$ will be some sort of effective "box" potential for each gas particle, comprising the mean effect of all container particles. This makes $\hat{I}'$, the deviation of the true "particle-by-particle" wall interaction from the "effective box" wall interaction, likely to be small. The eigenstates of the gas system alone are then simply the well known bound eigenstates of particles in a box of corresponding size.

The exact mathematical minimization of $\hat{I}'$ for a given wave function of the full system $\psi = \psi(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\})$, using the Euler–Lagrange method, reads

$$\begin{aligned}\langle\psi|(\hat{I} - \hat{V}^{\mathrm{g}})^2|\psi\rangle = \int \Big[&\Big(\hat{I}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\})\Big)^2 - 2\hat{I}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\})\hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}) \\ &+ \Big(\hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\})\Big)^2\Big] \left|\psi(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\}\right|^2 \prod_\nu \mathrm{d}\boldsymbol{q}_\nu^{\mathrm{c}} \prod_\mu \mathrm{d}\boldsymbol{q}_\mu^{\mathrm{g}} \; .\end{aligned} \tag{7.16}$$

Since the minimization has to be done with respect to $\hat{V}^{\mathrm{g}}$, which only depends on the gas variables $\{\boldsymbol{q}_\mu^{\mathrm{g}}\}$, we find

$$\begin{aligned}-2\int \hat{I}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\}) \left|\psi(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\})\right|^2 \prod_\nu \mathrm{d}\boldsymbol{q}_\nu^{\mathrm{c}} \\ + 2\,\hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}) \int \left|\psi(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}, \{\boldsymbol{q}_\nu^{\mathrm{c}}\})\right|^2 \prod_\nu \mathrm{d}\boldsymbol{q}_\nu^{\mathrm{c}} = 0 \; ,\end{aligned} \tag{7.17}$$

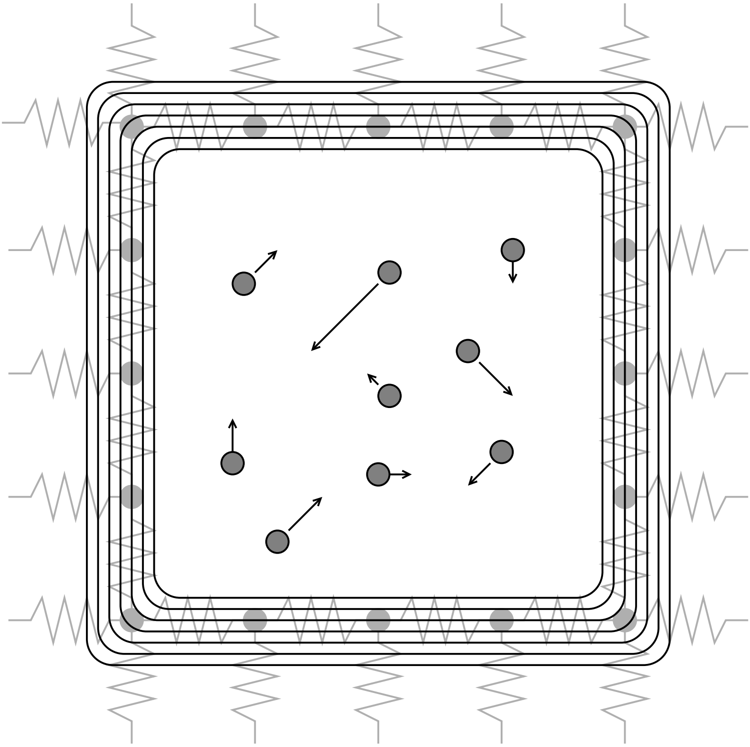


Fig. 7.3. Bipartite system: gas in a container, represented by an effective single particle potential (indicated by equipotential lines).

from which we get, solving for $\hat{V}^{\mathrm{g}}$,

$$\hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\}) = \frac{\int \hat{I}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\},\{\boldsymbol{q}_\nu^{\mathrm{c}}\}) \ \left|\psi(\{\boldsymbol{q}_\mu^{\mathrm{g}}\},\{\boldsymbol{q}_\nu^{\mathrm{c}}\})\right|^2 \ \prod_\nu \mathrm{d}\boldsymbol{q}_\nu^{\mathrm{c}}}{\int \ \left|\psi(\{\boldsymbol{q}_\mu^{\mathrm{g}}\},\{\boldsymbol{q}_\nu^{\mathrm{c}}\})\right|^2 \ \prod_\nu \mathrm{d}\boldsymbol{q}_\nu^{\mathrm{c}}} \ . \tag{7.18}$$

The effective potential $\hat{V}^{\mathrm{g}}(\{\boldsymbol{q}_\mu^{\mathrm{g}}\})$ is now a sort of a weighted summation over all interactions of the gas particle with all particles of the container (see Fig. 7.3). According to the pre-condition of these considerations, this potential will indeed be a weak interaction. The Hamiltonian is now reorganized so that a partition into system and environment with a weak interaction is definitely possible. For a concrete situation with given $\hat{I}$, one can evaluate the effective potential with the aid of (7.18).

In general, 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 specific example of this partition scheme see [40].)

8 Structure of Hilbert Space

> *It is not true that we can pursue science completely by using only those concepts which are directly subject to experiment.*
>
> — R. P. Feynman, R. B. Leighton and M. Sands [36]

Before we can actually start to analyze the structure of compound Hilbert spaces in a similar way in which the topology of phase space has been analyzed in Chap. 4, we have to establish a representation of Hilbert space. This representation is singled out by being appropriate for our purposes just as momentum-position phase space has been established as an appropriate representation in the classical domain.

Once we have chosen some representation, any quantity defined as a function of the total system state $|\psi\rangle$ may be visualized as a “landscape” over the variables defined by the representation. To analyze now the structure of such landscapes, i.e., to find out whether they are essentially “flat” or very “hilly”, what their mean altitude is, etc., we use the mathematical methods called Hilbert space average and Hilbert space variance. These can be used as measures for the above properties and will be introduced rather formally in this chapter.

8.1 Representation of Hilbert Space

Contrary to the real configuration space, Hilbert space (see Sect. 2.2), 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 to parametrize, i.e., to specify by a set of numbers, quantum mechanical states. Obviously one has to choose a basis $\{|i\rangle\}$ such that

$$|\psi\rangle = \sum_i \psi_i |i\rangle \; . \tag{8.1}$$

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 Hilbert space really is, and this is a very important question (see Sect. 6.3), the only way is to calculate the size of the region that the corresponding specifying parameters

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 79–89 (2004)
`http://www.springerlink.com/`

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 Hilbert spaces, it needs to be answered for classical phase 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 formulations 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 (see Sect. 4.4) this parametrization remains without justification.

If one wants to use the relative sizes of compartments to guess in which compartment the representing point will most likely be found, the crucial variable is the effective velocity. In the case of, e.g., two compartments, one being much bigger than the other, the guess that the representing point will preferably be found in the bigger one might be wrong, if the dynamics were such that the point moved with extremely high velocity in the big compartment and very slowly in the small one. Unfortunately, the effective velocity of the representing point on his trajectory through the representing Cartesian parameter space depends on the 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 Hilbert space) were constant throughout one accessible region. Fortunately this is feasible.

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

$$|\psi\rangle = \sum_i (\eta_i + \mathrm{i}\,\xi_i)\,|i\rangle \ . \tag{8.2}$$

If the η_i and ξ_i are organized in a Cartesian parameter space, a Hilbert space with a real, regular, Cartesian metric is defined. All vectors that represent physical states, i.e., that are normalized, lie on a hypersphere of unit radius,

$$\langle\psi|\psi\rangle = \sum_i (\eta_i - \mathrm{i}\,\xi_i)\,(\eta_i + \mathrm{i}\,\xi_i) = \sum_i \left(\eta_i^2 + \xi_i^2\right) = 1 \ , \tag{8.3}$$

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

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

$$v = \sqrt{\sum_i \left[\left(\frac{\mathrm{d}\eta_i}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}\xi_i}{\mathrm{d}t}\right)^2\right]} \ . \tag{8.4}$$

The square of this velocity may be written as

$$v^2 = \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\eta_i + \mathrm{i}\,\xi_i\right)\frac{\mathrm{d}}{\mathrm{d}t}\left(\eta_i - \mathrm{i}\,\xi_i\right) = \left|\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle\right|^2 , \tag{8.5}$$

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:

$$\mathrm{i}\hbar\,\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle = \left(\hat{H} - E_0\right)|\psi\rangle , \tag{8.6}$$

where E_0 is an arbitrary real constant zero-point adjustment of the energy that results just in an overall phase factor of $\exp(\mathrm{i}E_0 t/\hbar)$ for any solution. It has, however, an influence on the Hilbert space velocity, for it could, e.g., make a stationary state be represented by a moving point in Hilbert space. One thus wants to choose E_0 such as to make the Hilbert space velocity as small as possible, since any motion that is due to E_0 just reflects a changing overall phase which has no physical significance.

Inserting (8.6) into (8.5) we find for the square of the Hilbert space velocity v^2:

$$\begin{aligned} v^2 &= \langle\psi|\left(\hat{H} - E_0\right)\left(\hat{H} - E_0\right)|\psi\rangle \\ &= \langle\psi|\hat{H}^2|\psi\rangle - 2E_0\langle\psi|\hat{H}|\psi\rangle + E_0^2\langle\psi|\psi\rangle . \end{aligned} \tag{8.7}$$

Obviously v is constant, since all terms in (8.7) 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 E_0 now yields

$$E_0 = \langle\psi|\hat{H}|\psi\rangle = \langle\psi(0)|\hat{H}|\psi(0)\rangle , \tag{8.8}$$

which, inserted into (8.7) gives the Hilbert space velocity

$$\begin{aligned} v &= \sqrt{\langle\psi|\hat{H}^2|\psi\rangle - (\langle\psi|\hat{H}|\psi\rangle)^2} \\ &= \sqrt{\langle\psi(0)|\hat{H}^2|\psi(0)\rangle - (\langle\psi(0)|\hat{H}|\psi(0)\rangle)^2} , \end{aligned} \tag{8.9}$$

which is just the energy uncertainty, or the variance of the energy probability distribution of the corresponding state. Accordingly, stationary states, i.e., energy eigenstates, are represented by non-moving points in Hilbert space. 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 Schrödinger dynamics through one accessible region do so with the same constant velocity.

8.2 Hilbert Space Average

This and the remaining sections of Chap. 8 are meant to introduce the mathematical ideas behind the methods used in Chap. 9 on a rather abstract level. (These methods are explained in full detail in the Appendix.) Though we strongly recommend that the reader go through these sections, they may be skipped by the reader who is primarily interested in results.

In the following we will often be interested in the average of a certain quantity within a subregion of the complete Hilbert space called the accessible region (AR). Considering the partition scheme of Sect. 7.1 (the full system consists of a gas system g and a container system c) this special quantity is mostly the purity P^{g} of the gas system (see Chap. 9). Of course, the purity itself depends on the complete state, the state of system and environment together. The state of the full system is constrained to the accessible region within the high-dimensional Hilbert space, which results from the interaction model of system and environment, as will be seen later. Within this accessible region we will then have to calculate the respective average of the purity P^{g}. Furthermore, there are some other functions of the state of the total system, for which we will evaluate the mean value within some subregion of the respective Hilbert space. Before we can compute these mean values, we need to know how to evaluate such an average of a quantity in Hilbert space in general.

Let f be a function of the complete state $|\psi\rangle$ of the system in the accessible region, AR, of the whole Hilbert space. To calculate the Hilbert space average $[\![f]\!]$ of f over AR we use the parametrization for a state $|\psi\rangle$ introduced in the last section, the real and imaginary parts $\{\eta_i, \xi_i\}$ of the complex amplitude ψ_i. The Hilbert space is now represented by a $2n_{\mathrm{tot}}$-dimensional Cartesian space, in which the Hilbert space average over AR is defined as

$$[\![f]\!] = \frac{\int_{\mathrm{AR}} f(\{\eta_i, \xi_i\}) \prod_{i=1}^{n_{\mathrm{tot}}} \mathrm{d}\eta_i \mathrm{d}\xi_i}{\int_{\mathrm{AR}} \prod_{i=1}^{n_{\mathrm{tot}}} \mathrm{d}\eta_i \mathrm{d}\xi_i} , \tag{8.10}$$

where the integral in the denominator is just the area $\mathcal{O}(\mathrm{AR})$ of the accessible region we are integrating over. The Hilbert space average meets all properties of standard averages

$$[\![c f]\!] = c[\![f]\!] \quad \text{with} \quad c \in \mathbb{C} , \tag{8.11}$$

$$[\![f + f']\!] = [\![f]\!] + [\![f']\!] , \tag{8.12}$$

$$[\![f^*]\!] = [\![f]\!]^* . \tag{8.13}$$

The accessible region will always be defined by some contraints for the Cartesian coordinates $\{\eta_i, \xi_i\}$. The most basic constraint is the normalization of a state in Hilbert space,

$$\mathrm{sph}(2n_{\mathrm{tot}}) : \sum_{i=1}^{n_{\mathrm{tot}}} \left(\eta_i^2 + \xi_i^2\right) = 1 , \tag{8.14}$$

obviously a hypersphere $\mathrm{sph}(2n_{\mathrm{tot}})$ with radius 1 in the parameter space (for more details about hyperspheres see App. A).

Firstly, we now calculate the Hilbert space average of the quantity f over the whole Hilbert space – with only one condition, the normalization, restricting the parameters $\{\eta_i, \xi_i\}$. Thus, the accessible region (AR) is the Hilbert space itself, a hypersphere in the $2n_{\mathrm{tot}}$-dimensional Cartesian space with radius one. All allowed quantum mechanical states of the system are on the surface of this hypersphere in the parameter space. Instead of directly integrating only over this accessible region, we can integrate over the whole space $\mathbb{R}^{(2n_{\mathrm{tot}})}$, introducing a δ-function to restrict the integration to the respective accessible region,

$$[\![f]\!] = \frac{1}{\mathcal{O}(\mathrm{AR})} \int_{\mathbb{R}} f(\{\eta_i, \xi_i\})\, \delta\Big(\sqrt{\textstyle\sum_{i=1}^{n_{\mathrm{tot}}}(\eta_i^2+\xi_i^2)}-1\Big) \prod_{i=1}^{n_{\mathrm{tot}}} \mathrm{d}\eta_i \mathrm{d}\xi_i \ . \tag{8.15}$$

Since the constraint (8.14) defines a hypersphere, it is convenient to use generalized spherical coordinates for the integration: a radius r and $2n_{\mathrm{tot}} - 1$ angle coordinates ϕ_i (see App. A.1 and especially (A.3)). Based on this coordinate transformation

$$\{\eta_i, \xi_i\} \mapsto \{r, \phi_1, \dots, \phi_{2n_{\mathrm{tot}}-1}\} \tag{8.16}$$

with the appropriate functional matrix (Jacobian matrix)

$$\mathsf{F} = \begin{pmatrix} \frac{\partial \eta_1}{\partial r} & \frac{\partial \eta_1}{\partial \phi_1} & \cdots & \frac{\partial \eta_1}{\partial \phi_{2n_{\mathrm{tot}}-1}} \\ \vdots & \ddots & \vdots & \\ \frac{\partial \eta_{n_{\mathrm{tot}}}}{\partial r} & \frac{\partial \eta_{n_{\mathrm{tot}}}}{\partial \phi_1} & \cdots & \frac{\partial \eta_{n_{\mathrm{tot}}}}{\partial \phi_{2d-1}} \\ \frac{\partial \xi_1}{\partial r} & \frac{\partial \xi_1}{\partial \phi_1} & \cdots & \frac{\partial \xi_1}{\partial \phi_{2n_{\mathrm{tot}}-1}} \\ \vdots & \ddots & \vdots & \\ \frac{\partial \xi_{n_{\mathrm{tot}}}}{\partial r} & \frac{\partial \xi_{n_{\mathrm{tot}}}}{\partial \phi_1} & \cdots & \frac{\partial \xi_{n_{\mathrm{tot}}}}{\partial \phi_{2n_{\mathrm{tot}}-1}} \end{pmatrix} \tag{8.17}$$

and its functional determinant $\det \mathsf{F}$, the integral can be transformed into a spherical integration, finding

$$[\![f]\!] = \frac{1}{\mathcal{O}(\mathrm{AR})} \int f(\{r, \phi_1, \dots, \phi_{2n_{\mathrm{tot}}-1}\})\, \delta(r-1)\, |\det \mathsf{F}|\, \mathrm{d}r \prod_{i=1}^{2n_{\mathrm{tot}}-1} \mathrm{d}\phi_i \ . \tag{8.18}$$

The integration over the radius can be done immediately according to the δ-function. For the remaining integrals

$$[\![f]\!] = \frac{1}{\mathcal{O}(\mathrm{AR})} \int f(\{1, \phi_1, \dots, \phi_{2n_{\mathrm{tot}}-1}\})\, |\det \mathsf{F}| \prod_{i=1}^{2n_{\mathrm{tot}}-1} \mathrm{d}\phi_i \ , \tag{8.19}$$

there are no further restrictions. These integrals can be solved directly for concrete situations, especially for a polynomial function f (see App. A.2).

In many cases there are more restrictions to the Cartesian coordinates, besides the normalization. These follow from the coupling model of the system and its environment. Think, e.g., of a coupling model, where the system is not allowed to exchange any energy with the environment, then the energy is conserved in both parts of the full system separately. The additional conditions restrict the state of the full system to a subregion within the full hypersphere generated by the normalization condition.

As will be seen later, all these conditions, labeled by J, also define hyperspheres, but with a dimension n^J small compared to the dimension n_{tot} of the normalization hypersphere. A very important property of the set of additional conditions is that all of them depend only on a subset of the parameters $\{\eta_i, \xi_i\}$, where every parameter enters just one condition J. Therefore we could label each coordinate with the number of the condition J to which the parameter belongs. The condition itself reads

$$\text{sph}(2n^J) : \sum_{j=1}^{n^J} \left((\eta_j^J)^2 + (\xi_j^J)^2 \right) = R^J \,, \tag{8.20}$$

with dimension n^J and radius R^J.

Because each parameter now enters one further condition, we find according to the normalization condition (8.14)

$$\sum_J \sum_{j=1}^{n^J} \left((\eta_j^J)^2 + (\xi_j^J)^2 \right) = \sum_J R^J = 1 \,, \tag{8.21}$$

restricting the possible radii R^J. Later we will find that the R^J's are probabilities and therefore this normalization condition is automatically fulfilled.

Due to these considerations, the whole parameter space may be decomposed into completely independent subspaces J. Since each subspace is defined by a hypersphere equation, it is convenient to switch, again, to generalized spherical coordinates,

$$\{\eta_j^J, \xi_j^J\} \mapsto \{r^J, \phi_j^J\} \tag{8.22}$$

in each subspace J. This obviously leads to a functional matrix F with block diagonal form

$$\mathsf{F}(\{\eta_i, \xi_i\}) = \begin{pmatrix} \tilde{\mathsf{F}}^1(\{r^1, \phi_j^1\}) & & & 0 \\ & \ddots & & \\ & & \tilde{\mathsf{F}}^J(\{r^J, \phi_j^J\}) & \\ 0 & & & \ddots \end{pmatrix} = \prod_J \mathsf{F}^J(\{r^J, \phi_j^J\}) \,, \tag{8.23}$$

where F^J has at the position J the block $\tilde{\mathsf{F}}^J$ and else the $\hat{1}$-operator. Because the determinant of a product is the product of the individual determinants, we find for the functional determinant

$$\det \mathsf{F} = \det \left(\prod_J \mathsf{F}^J \right) = \prod_J \det \mathsf{F}^J \ . \tag{8.24}$$

Like in the case of only a single condition, the normalization condition, the whole integration is now simplified. Each condition leads to a δ-function in the integration, which, after transformation to spherical coordinates in each subspace, just reads $\delta(r^J - R^J)$. This allows for a trivial integration over all radius variables r^J, leaving all integrations over angle variables of all subspaces J without any further restrictions. Finally we find for the Hilbert space average of the quantity f over the accessible region, defined by the additional conditions J,

$$[\![f]\!] = \frac{1}{\mathcal{O}(\mathrm{AR})} \int f(\{R^J, \phi_i^J\}) \prod_J \left|\det \mathsf{F}^J\right| \prod_j \mathrm{d}\phi_j^J \ . \tag{8.25}$$

Therefore the average over the region in Hilbert space decomposes into a product of single averages over subspaces J.

In a concrete situation it remains difficult to calculate these Hilbert space averages, but based on these technical considerations we will be able to evaluate the Hilbert space average of the purity P^{g} in the next chapter.

8.3 Hilbert Space Variance

In order to analyze the properties of a certain region in Hilbert space by calculating the Hilbert space average of a special quantity, it is furthermore necessary to investigate the deviation from the mean value introduced in the last section. The respective quantity we are interested in could be seen as a landscape over the Hilbert space, since it is a scalar quantity dependent on the state of the full system. We are not only interested in something like the mean "height" of the landscape in a certain region of the Hilbert space (the accessible region), but also in how "hilly" this landscape really is. For an arbitrary state of the accessible region we would like to estimate a quantity, say the purity P^{g}, by its mean value in the respective region. If the landscape were indeed very flat, we would find that the estimated purity is very close to the actual P^{g} else, for very "hilly" landscapes, this would usually not be the case. In fact, our estimation could then be very bad. Therefore we need the Hilbert space variance of the respective quantity to additionally estimate the expected deviation of our estimated value.

Like for any standard mean value it is possible to define a second moment. One can thus define the Hilbert space variance as

$$\Delta^2_{\mathrm{H}}(f) := [\![f^2]\!] - [\![f]\!]^2 \,. \tag{8.26}$$

Together with all the techniques introduced in the last section, this quantity can in principle be evaluated, even if the concrete integration is in most cases a very complicated calculation.

In Chap. 9 we intend to analyze the purity landscape over the Hilbert space of the bipartite system in the accessible region with the aid of the powerful tools – the Hilbert space average and the Hilbert space variance.

8.4 Purity and Local Entropy in Product Hilbert Space

In this section we would like to investigate the properties – purity and local entropy – of a complete Hilbert space without further restrictions. We first consider a distribution of pure states over the whole Hilbert space, which is invariant under any infinitesimal unitary transformations. Having found such a distribution, we can compute the distribution of certain quantities within the whole Hilbert space.

8.4.1 Unitary Invariant Distribution of Pure States

For simplicity we start by considering a two-dimensional Hilbert space. According to (8.2), any normalized state vector $|\psi\rangle$ can be represented by the real as well as imaginary parts $\{\eta_i, \xi_i\}$, basis $|i\rangle$, of the complex amplitudes, and fulfill the condition (8.3). In spite of this constraint let us assume for the moment that all these parameters are independent. We are looking for the probability distribution,

$$W(\eta_1, \eta_2, \xi_1, \xi_2) = W(\eta_1)W(\eta_2)W(\xi_1)W(\xi_2) \,, \tag{8.27}$$

which is invariant under the unitary transformation

$$\hat{U} = \begin{pmatrix} 1 + \mathrm{i}\varepsilon_3 & \mathrm{i}\varepsilon_1 - \varepsilon_2 \\ \mathrm{i}\varepsilon_1 + \varepsilon_2 & 1 - \mathrm{i}\varepsilon_3 \end{pmatrix} \,. \tag{8.28}$$

This transformation leads to

$$|\psi'\rangle = \hat{U}|\psi\rangle = \sum_{i=1}^{2} (\eta'_i + \mathrm{i}\xi'_i)\,|i\rangle \tag{8.29}$$

with the coordinate transformation

$$\eta'_1 = \eta_1 - \xi_2\varepsilon_1 - \eta_2\varepsilon_2 - \xi_1\varepsilon_3 \,, \tag{8.30}$$

$$\eta'_2 = \eta_2 - \xi_1\varepsilon_1 + \eta_1\varepsilon_2 + \xi_2\varepsilon_3 \,, \tag{8.31}$$

$$\xi'_1 = \xi_1 + \eta_2\varepsilon_1 - \xi_2\varepsilon_2 + \eta_1\varepsilon_3 \,, \tag{8.32}$$

$$\xi'_2 = \xi_2 + \eta_1\varepsilon_1 + \xi_1\varepsilon_2 - \eta_2\varepsilon_3 \,. \tag{8.33}$$

It suffices to consider infinitesimal changes in the transformed probability distribution of a single coordinate:

$$W(\eta_i') = W(\eta_i) + (\eta_i' - \eta_i)\,\frac{\partial W(\eta_i)}{\partial \eta_i} \tag{8.34}$$

and in the same way for the distribution of the imaginary parts of the coordinates $W(\xi_i')$. Keeping only terms of first order in ε, we obtain for the completely transformed probability distribution

$$\begin{aligned}
&W(\eta_1')\,W(\eta_2')\,W(\xi_1')\,W(\xi_2') \\
&\approx W(\eta_2)\,W(\xi_1)\,W(\xi_2)\,\frac{\partial W(\eta_1)}{\partial \eta_1}\,(-\xi_2\varepsilon_1 - \eta_2\varepsilon_2 - \xi_1\varepsilon_3) \\
&\quad + W(\eta_1)\,W(\xi_1)\,W(\xi_2)\,\frac{\partial W(\eta_2)}{\partial \eta_2}\,(-\xi_1\varepsilon_1 + \eta_1\varepsilon_2 + \xi_2\varepsilon_3) \\
&\quad + W(\eta_1)\,W(\eta_2)\,W(\xi_2)\,\frac{\partial W(\xi_1)}{\partial \xi_1}\,(\eta_2\varepsilon_1 - \xi_2\varepsilon_2 + \eta_1\varepsilon_3) \\
&\quad + W(\eta_1)\,W(\eta_2)\,W(\xi_1)\,\frac{\partial W(\xi_2)}{\partial \xi_2}\,(\eta_1\varepsilon_1 + \xi_1\varepsilon_2 + \eta_2\varepsilon_3)\;.
\end{aligned} \tag{8.35}$$

Postulating

$$\frac{\partial W(\eta_i)}{\partial \eta_i} \approx -\eta_i\,W(\eta_i)\;, \qquad \frac{\partial W(\xi_i)}{\partial \xi_i} \approx -\xi_i\,W(\xi_i)\;, \tag{8.36}$$

leads to

$$W(\eta_1')\,W(\eta_2')\,W(\xi_1')\,W(\xi_2') \approx \text{const.} \cdot W(\eta_1)\,W(\eta_2)\,W(\xi_1)\,W(\xi_2)\;, \tag{8.37}$$

as required. Equation (8.36) implies the normalized solution

$$W(\eta_i) = \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{e}^{\eta_i^2/2\sigma^2}\;, \qquad W(\xi_i) = \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{e}^{\xi_i^2/2\sigma^2}\;. \tag{8.38}$$

As long as the complete probability distribution is a Gaussian distribution of the single coordinates, it is invariant under unitary transformations.

Generalizing this result for two-dimensional Hilbert spaces to any finite Hilbert space of dimension n_{tot}, we thus end up with the Gaussian (cf. App. A)

$$W(\{\eta_i, \xi_i\}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{2n_{\text{tot}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n_{\text{tot}}}\left(\eta_i^2 + \xi_i^2\right)\right)\;. \tag{8.39}$$

The normalization condition for the wave function, though, requires that the sum of the squares of the coordinates is one (see (8.3)), i.e., the parameters are not independent, contrary to our assumption. However, for large n_{tot} the central limit theorem tells us that $W(\{\eta_i, \xi_i\})$ is indeed approximate a Gaussian provided we choose [92]

$$\sigma = \frac{1}{\sqrt{2n_{\text{tot}}}} . \tag{8.40}$$

The above unitary invariant distribution holds for an n_{tot}-dimensional Hilbert space without further constraints. It characterizes an ensemble in Hilbert space, from which to pick "typical" pure states.

8.4.2 Application

The results of the preceding section are now applied to a bipartite system of dimension $n_{\text{tot}} = n^{\text{g}} \cdot n^{\text{c}}$, with $n^{\text{g}} \leq n^{\text{c}}$. Let $f = f(|\psi\rangle) = f(\{\eta_i, \xi_i\})$ be some function of the state vector. Then we can define its Hilbert space average $[\![f]\!]$ and its Hilbert space distribution $\{f\}$, respectively, as

$$[\![f]\!] = \int W(\{\eta_i, \xi_i\}) f(\{\eta_i, \xi_i\}) \prod_{i=1}^{n_{\text{tot}}} \mathrm{d}\eta_i \mathrm{d}\xi_i , \tag{8.41}$$

$$\{f\} = \int \delta(f(\{\eta_i, \xi_i\}) - f)\, W(\{\eta_i, \xi_i\}) \prod_{i=1}^{n_{\text{tot}}} \mathrm{d}\eta_i \mathrm{d}\xi_i . \tag{8.42}$$

Here we restrict ourselves to the local purity P^{g} and local entropy S^{g}. The resulting distribution $\{P^{\text{g}}\}$ is shown in Fig. 8.1 for $n^{\text{g}} = 2$ and varying n^{c}. We see that this distribution tends to peak at the minimum value $P^{\text{g}} = 1/n^{\text{g}} = 1/2$. Its average is given by (see [94])

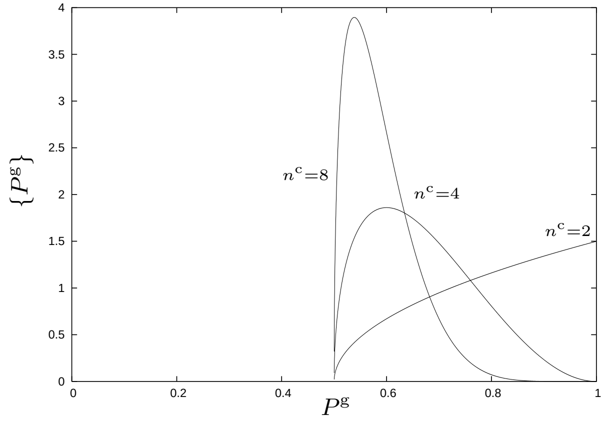


Fig. 8.1. Purity distribution $\{P^{\text{g}}\}$ for $n^{\text{g}} = 2$ and several container dimensions n^{c}.

$$[\![P^{\mathrm{g}}]\!] = \frac{n^{\mathrm{g}} + n^{\mathrm{c}}}{n^{\mathrm{g}} n^{\mathrm{c}} + 1} \ . \tag{8.43}$$

In Fig. 8.2 we show the Hilbert space average of S^{g} for $n^{\mathrm{g}} = 2$ as a function of n^{c}. Again we see that S^{g} rapidly approaches its maximum value $S^{\mathrm{g}}_{\mathrm{max}}$ for large embedding n^{c}. For $1 \ll n^{\mathrm{g}} \leq n^{\mathrm{c}}$ one obtains [94]

$$[\![S^{\mathrm{g}}]\!] \approx \ln n^{\mathrm{g}} - \frac{n^{\mathrm{g}}}{2n^{\mathrm{c}}} \ . \tag{8.44}$$

Both results indicate that for $n^{\mathrm{g}} \ll n^{\mathrm{c}}$ a typical state of subsystem g is totally "mixed": all its local properties have maximum uncertainty.

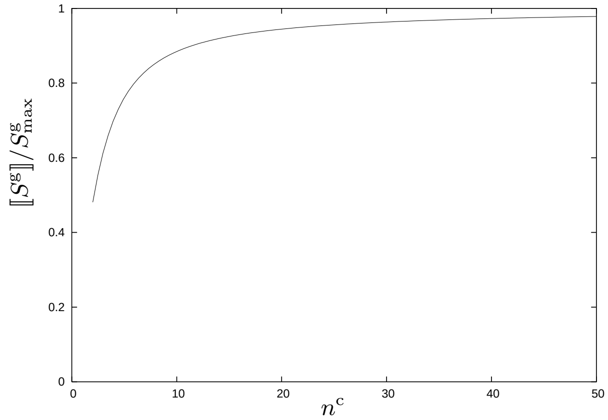


Fig. 8.2. Average entropy $[\![S^{\mathrm{g}}]\!]/S^{\mathrm{g}}_{\mathrm{max}}$ of subsystem g ($n^{\mathrm{g}} = 2$), depending on the dimension n^{c} of the environment.

9 Quantum Thermodynamic Equilibrium

> *The results in decoherence theory strongly suggest that interactions with the environment are crucial in the emergence of quasi-classical and thermodynamic behavior.*
>
> — M. Hemmo and O. Shenker [53]

A very important characterization of thermodynamic systems concerns the kind of contact between the observed system and its environment. In classical thermodynamics there is a large variety of different contact scenarios, because of the large variety of thermodynamic experiments. These scenarios are determined by the constraints under which the experiments are performed. Not so much for practical, but for theoretical reasons, the most important contact conditions are the *microcanonical* and the *canonical* conditions. In the microcanonical contact scenario no energy transfer between system and environment is allowed, whereas in the canonical contact such an energy exchange is possible. In this chapter we will analyze these two important constraints.

9.1 Microcanonical Conditions

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. [12, 15]). 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 Chap. 7. 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. As will be seen later, this does not mean that such an interaction has no effect on the considered system, a fact that might seem counterintuitive from a classical point of view but is, nevertheless, true in the quantum regime. This constraint, however, leads to an immense reduction of the region in Hilbert space which the wave vector is allowed to enter. This reduced area is called the “accessible region” of the system.

9.1.1 Accessible Region (AR)

We focus now on a bipartite system under the special constraints given in Chap. 7 and with the Hamiltonian (7.1). If the energies contained in the gas and the environment, respectively,

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 91–108 (2004)
`http://www.springerlink.com/`

$$E^{\mathrm{g}} := \langle \hat{H}^{\mathrm{g}} \rangle \,, \quad E^{\mathrm{c}} := \langle \hat{H}^{\mathrm{c}} \rangle \tag{9.1}$$

are to be conserved, which means that these two energies are constants of motion, the following commutator relations should hold

$$\left[\hat{H}^{\mathrm{g}}, \hat{H}\right] = 0 \,, \quad \left[\hat{H}^{\mathrm{c}}, \hat{H}\right] = 0 \,. \tag{9.2}$$

It then follows from

$$\left[\hat{H}^{\mathrm{g}}, \hat{H}\right] = \left[\hat{H}^{\mathrm{g}}, \hat{H}^{\mathrm{g}}\right] + \left[\hat{H}^{\mathrm{g}}, \hat{H}^{\mathrm{c}}\right] + \left[\hat{H}^{\mathrm{g}}, \hat{I}\right] = 0 \tag{9.3}$$

that

$$\left[\hat{H}^{\mathrm{g}}, \hat{I}\right] = 0 \,, \quad \left[\hat{H}^{\mathrm{c}}, \hat{I}\right] = 0 \,. \tag{9.4}$$

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 these conditions and create entanglement and therefore give rise to the increase of local entropy.

Due to (9.2) and the considerations of Sect. 2.4 the local energy projectors $\hat{P}^{\mathrm{g}}_A$ of the gas system and $\hat{P}^{\mathrm{c}}_B$ of the container

$$\hat{P}^{\mathrm{g}}_A = \sum_a |A,a\rangle\langle A,a| \,, \quad \hat{P}^{\mathrm{c}}_B = \sum_b |B,b\rangle\langle B,b| \tag{9.5}$$

commute with the full Hamiltonian,

$$\left[\hat{P}^{\mathrm{g}}_A, \hat{H}\right] = \left[\hat{P}^{\mathrm{c}}_B, \hat{H}\right] = 0 \,. \tag{9.6}$$

Thus, because the system is not allowed to exchange energy with the environment the joint probability W_{AB} introduced in (7.7) must be conserved (see Sect. 2.4)

$$\langle\psi|\hat{P}^{\mathrm{g}}_A \hat{P}^{\mathrm{c}}_B|\psi\rangle = \sum_{a,b} \left|\psi^{AB}_{ab}(t)\right|^2 = \sum_{a,b} \left|\psi^{AB}_{ab}(0)\right|^2 = W_{AB} \,, \tag{9.7}$$

and is set by the initial state. This means that the energy probability distribution $\{W_{AB}\}$ is a constant of motion in this system. Vice versa, any state that features the same energy probability distribution as the initial state belongs to the accessible region and could possibly be reached during microcanonical dynamics.

In the following, we mainly consider initial product states, states that have zero local entropy in the beginning and for which

$$\sum_{a,b} \left|\psi^{AB}_{ab}(0)\right|^2 = \sum_{a,b} \left|\psi^{A}_{a}(0)\right|^2 \left|\psi^{B}_{b}(0)\right|^2 = W_A W_B \,. \tag{9.8}$$

This is the only constraint that microcanonical conditions impose on the accessible region of Hilbert space. Note that this does not mean that local entropy is constant.

9.1.2 The "Landscape" of P^{g} in the Accessible Region

To demonstrate that the accessible region really has the cell structure mentioned in Sect. 6.3, namely that the biggest part of it is filled with states of almost minimum purity (maximum entropy), we proceed as follows:

1. First we compute the (unique) state with the lowest possible purity, $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ (with purity $P(\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}) = P^{\mathrm{g}}_{\mathrm{min}}$) that is consistent with the given initial state and the microcanonical conditions, consequently with a given energy probability distribution $\{W_A\}$, (see Sect. 9.1.3).
2. In Sect. 9.1.4 we compute the average of P^{g} over the total accessible Hilbert space region as introduced in Sect. 8.2.
3. We will show that this average purity is very close to the purity of the lowest possible purity state $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ for a large class of systems. Considering only these systems, which then define the class of thermodynamic systems, we can conclude that $P^{\mathrm{g}} \approx P^{\mathrm{g}}_{\mathrm{min}}$ for almost all states within the accessible region. Note that this conclusion is only possible because of the fact that the purity of $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ is the absolute minimal purity which can be reached at all in the system. A quantity with a mean value close to a boundary cannot vary very much. Thus it is not possible that the distribution of the purity within the accessible region is something else but a very flat "lowland", with a "soft depression" at $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ (see Fig. 9.1) and a "peak" with $P^{\mathrm{g}} = 1$.
4. Since all states from the accessible region have the same energy probability distribution $\{W_A\}$ (remember (9.8)) and the minimum purity state $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ is consistent with this distribution, all other states within the accessible region that feature $P^{\mathrm{g}} \approx P^{\mathrm{g}}_{\mathrm{min}}$ must yield reduced local states that are very close to $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ (in this context close means in terms of the distance measure $\mathrm{Tr}\left\{(\hat{\rho}^{g} - \hat{\rho}^{\mathrm{g}}_{\mathrm{min}})^2\right\}$ defined in Sect. 2.2.3 (2.19)). Thus, as long as the trajectory keeps wandering through the compartment filled with those states, the gas system is locally in a stationary state, i.e., equilibrium is reached.

We are now going to work out these steps.

9.1.3 The Minimum Purity State

As will be shown below, the minimum purity state is established, if all states a that belong to the same energy eigenspace of the gas E^{g}_A are equally likely. Thus the minimum purity state consistent with the microcanonical conditions (9.7) and (9.8) and its corresponding purity are

$$\hat{\rho}^{\mathrm{g}}_{\mathrm{min}} = \sum_{A,a} \frac{W_A}{N_A} |A,a\rangle\langle A,a| \,, \qquad P^{\mathrm{g}}_{\mathrm{min}} = \sum_{A} \frac{W_A^2}{N_A} \,. \tag{9.9}$$

To check that this is, indeed, the state with the smallest purity consistent with the given energy probability distribution $\{W_A\}$, we introduce a deviation D

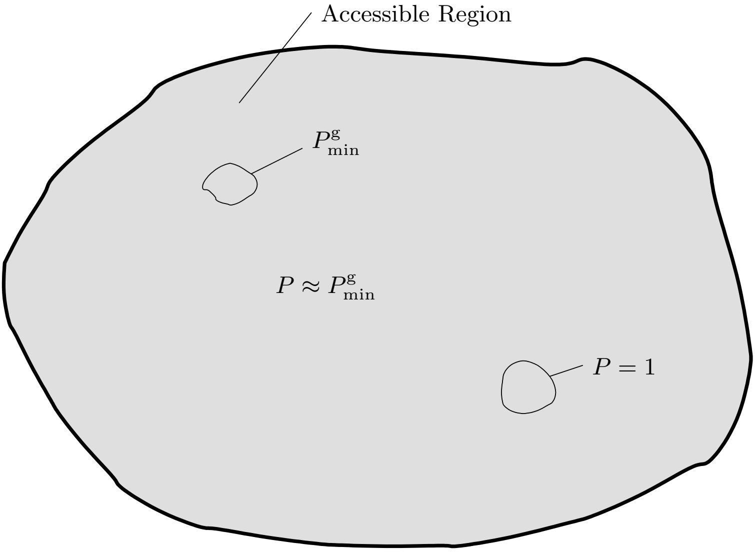


Fig. 9.1. Qualitative picture of the purity landscape in the microcanonical case. The biggest part of the accessible region is at $P \approx P^{\mathrm{g}}_{\mathrm{min}}$ or at $P = P^{\mathrm{g}}_{\mathrm{min}}$. There is however only a small zone featuring P significantly above $P^{\mathrm{g}}_{\mathrm{min}}$ or at the extreme $P = 1$. The only topological property this rough picture refers to is the relative size of different regions.

of the diagonal elements and a deviation E of the off-diagonal elements such that the resulting state is still consistent with $\{W_A\}$ and compute its purity. E is thus introduced as a matrix that does not have any diagonal elements. For the deviation

$$\mathsf{D} = \sum_{A,a} \mathsf{D}_{A,a} |A, a\rangle\langle A, a| \tag{9.10}$$

the partial trace over one degenerate subspace A has to vanish

$$\mathrm{Tr}_a \{\mathsf{D}\} = \sum_{a} \mathsf{D}_{A,a} = 0 \;, \tag{9.11}$$

because under microcanonical conditions the total probability distribution $\{W_A\}$ introduced by the initial state is fixed. The deviation D only redistributes the probability within a subspace A. E and D of course have to be Hermitian. Now with

$$\hat{\rho} = \hat{\rho}^{\mathrm{g}}_{\mathrm{min}} + \mathsf{D} + \mathsf{E} \;, \tag{9.12}$$

we find

$$\begin{aligned} P(\hat{\rho}) &= \mathrm{Tr}\left\{(\hat{\rho}^{\mathrm{g}}_{\mathrm{min}} + \mathsf{D} + \mathsf{E})^2\right\} \\ &= \mathrm{Tr}\left\{(\hat{\rho}^{\mathrm{g}}_{\mathrm{min}})^2\right\} + \mathrm{Tr}\left\{\mathsf{D}^2\right\} + \mathrm{Tr}\left\{\mathsf{E}^2\right\} \\ &\quad + 2\mathrm{Tr}\left\{\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}\mathsf{D}\right\} + 2\mathrm{Tr}\left\{\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}\mathsf{E}\right\} + 2\mathrm{Tr}\left\{\mathsf{D}\mathsf{E}\right\} \ . \end{aligned} \tag{9.13}$$

Due to the properties of E and the diagonality of $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ and D the last two terms vanish. Using the definitions (9.9) and (9.10) we compute the term

$$\mathrm{Tr}\left\{\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}\mathsf{D}\right\} = \sum_{A,a} \frac{W_A}{N_A} \mathsf{D}_{A,a} = \sum_{A} \frac{W_A}{N_A} \sum_{a} \mathsf{D}_{A,a} = 0 \ . \tag{9.14}$$

So we find

$$P(\hat{\rho}) = \mathrm{Tr}\left\{(\hat{\rho}^{\mathrm{g}}_{\mathrm{min}})^2\right\} + \mathrm{Tr}\left\{\mathsf{D}^2\right\} + \mathrm{Tr}\left\{\mathsf{E}^2\right\} \ . \tag{9.15}$$

Since

$$\mathrm{Tr}\left\{\mathsf{D}^2\right\} \geq 0 \ , \quad \mathrm{Tr}\left\{\mathsf{E}^2\right\} \geq 0 \ , \tag{9.16}$$

the smallest purity is reached for

$$\mathsf{E} = 0 \quad \text{and} \quad \mathsf{D} = 0 \ . \tag{9.17}$$

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

$$\hat{\rho} = \hat{\rho}^{\mathrm{g}}_{\mathrm{min}} \ . \tag{9.18}$$

Due to the uniqueness of $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ the following holds. 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.

9.1.4 The Hilbert Space Average of P^{g}

Now we have to evaluate the Hilbert space average of the purity of the gas system $[\![P^{\mathrm{g}}]\!]$ within the accessible region of the state using the techniques of Sect. 8.2. For these considerations we use the parametrization of the Hilbert space as before. The whole space is represented by the real and imaginary parts of a complex amplitude of the basis states ψ^{AB}_{ab}, introduced by the set of real Cartesian coordinates $\{\eta^{AB}_{ab}, \xi^{AB}_{ab}\}$. In each degeneracy subspace AB we find $N_{AB} = N_A N_B$ coordinates.

The accessible region (AR) of the system is defined by conditions (9.7) and (9.8), respectively, and obviously consists in each degeneracy subspace AB of hyperspheres with radii $R^{AB} = \sqrt{W_{AB}}$ or in the case of an initial product state $R^{AB} = \sqrt{W_A W_B}$. (Some information about hyperspheres in high-dimensional spaces can be found in App. A.)

According to its definition in (7.10) the purity of the gas system P^{g} is a function of the Cartesian coordinates $\{\eta_{ab}^{AB}, \xi_{ab}^{AB}\}$. In fact it is a polynomial of fourth order in all coordinates $\{\eta_{ab}^{AB}, \xi_{ab}^{AB}\}$. As explained in Sect. 8.2 the Hilbert space average of a function, here the purity P^{g}, within an accessible region (AR) can be evaluated by the integral

$$[\![P^{\mathrm{g}}]\!] = \frac{\int_{\mathrm{AR}} P^{\mathrm{g}}(\{\eta_{ab}^{AB}, \xi_{ab}^{AB}\}) \prod_{ABab} \mathrm{d}\eta_{ab}^{AB} \mathrm{d}\xi_{ab}^{AB}}{\int_{\mathrm{AR}} \prod_{ABab} \mathrm{d}\eta_{ab}^{AB} \mathrm{d}\xi_{ab}^{AB}} \,. \tag{9.19}$$

The concrete calculation of these integrals is a rather complicated task and of no physical relevance. Therefore we just give a flavor of how to integrate P^{g} over the accessible region and some ideas of the structure of the considered Hilbert space. The complete mathematical procedure can be found in App. B.

Since we have one hypersphere in each subspace AB, it is straightforward to switch to generalized spherical coordinates for each subspace. It is quite evident that the functional determinant $\det \mathsf{F}$ of this coordinate transformation decomposes into a product of determinants $\prod_{AB} \det \mathsf{F}^{AB}$ in the different subspaces AB because of the total independence of subspaces, as already mentioned in Sect. 8.2. As a consequence, the average over each term of P^{g} according to (9.19) factorizes and reduces to a product of averages over the degeneracy subspaces, to which the amplitudes in that term correspond.

The restriction to the accessible region is just a product of δ-functions, which restrict the radius r^{AB} of the generalized spherical coordinates to the radius of the respective hypersphere R^{AB}. Finally, the integral reduces to an integral over all angle variables of each subspace,

$$[\![P^{\mathrm{g}}]\!] = \frac{\int P^{\mathrm{g}}(\{\phi_i^{AB}\}) \prod_{AB} \left|\det \mathsf{F}^{AB}\right| \prod_j \phi_j^{AB}}{\prod_{AB} \int \left|\det \mathsf{F}^{AB}\right| \prod_j \phi_j^{AB}} \,. \tag{9.20}$$

These remaining integrals of a polynomial function over several hyperspheres are not easy to get but with a mathematical trick (see App. B) one is able to integrate P^{g}, finding the Hilbert space average of the quantity in the accessible region (cf. (B.29))

$$\begin{aligned} [\![P^{\mathrm{g}}]\!] = & \sum_A \frac{W_A^2}{N_A}\left(1 - \sum_B W_B^2\right) + \sum_B \frac{W_B^2}{N_B}\left(1 - \sum_A W_A^2\right) \\ & + \sum_{A,B} \frac{W_A^2 W_B^2 (N_A + N_B)}{N_A N_B + 1} \,. \end{aligned} \tag{9.21}$$

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

For a single degenerate subspace AB, the summations over A, B break down and we get ($N_A = n^{\mathrm{g}}$, $N_B = n^{\mathrm{c}}$)

$$[\![P^{\mathrm{g}}]\!] = \frac{n^{\mathrm{g}} + n^{\mathrm{c}}}{n^{\mathrm{g}} n^{\mathrm{c}} + 1} \tag{9.22}$$

in agreement with Sect. 8.4.

9.1.5 Microcanonical Equilibrium

If the degeneracy of the occupied energy levels is large enough so that

$$N_A N_B + 1 \approx N_A N_B \,, \tag{9.23}$$

which should hold true for typical thermodynamic systems, (9.21) reduces to

$$[\![P^{\mathrm{g}}]\!] \approx \sum_A \frac{(W_A)^2}{N_A} + \sum_B \frac{(W_B)^2}{N_B} \,. \tag{9.24}$$

The first sum in this expression is obviously exactly $P^{\mathrm{g}}_{\mathrm{min}}$ (9.9), 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 $P^{\mathrm{g}} \approx P^{\mathrm{g}}_{\mathrm{min}}$. The second sum will be small if the container system occupies highly degenerate states typical for thermodynamic systems, in which the surrounding is much larger than the considered system. This is the set of cases mentioned already in Sect. 9.1.2: all systems fulfilling this pre-condition are now called *thermodynamic systems*. Thus we can conclude that all states within the accessible region are very close to $\hat{\rho}^{\mathrm{g}}_{\mathrm{min}}$ and have approximately the purity $P^{\mathrm{g}}_{\mathrm{min}}$. The density operator, which has $P^{\mathrm{g}} = P^{\mathrm{g}}_{\mathrm{min}}$ and $S^{\mathrm{g}} = S^{\mathrm{g}}_{\mathrm{max}}$, and which is consistent with the microcanonical conditions, is unique. The density operators with $P^{\mathrm{g}} \approx P^{\mathrm{g}}_{\mathrm{min}}$ should not deviate much from this one and should therefore also have $S^{\mathrm{g}} \approx S^{\mathrm{g}}_{\mathrm{max}}$, the latter being

$$S^{\mathrm{g}}_{\mathrm{max}} = -k_{\mathrm{B}} \sum_A W_A \ln \frac{W_A}{N_A} \,, \tag{9.25}$$

which reduces for sharp energy probability distribution $\{W_A\} = \delta_{AA'}$ to

$$S^{\mathrm{g}}_{\mathrm{max}} = k_{\mathrm{B}} \ln N_{A'} \,. \tag{9.26}$$

For a numerical demonstration of several aspects of these considerations we refer to Sect. 18.1.

9.2 Energy Exchange Conditions

In the last sections we only considered a contact scenario, for which no energy transfer between the gas and the container was allowed. However, many systems do exchange energy with the environment, and therefore it is necessary to allow also for this possibility in our considerations.

For environments with a special kind of degeneracy structure, i.e., an exponential increase of the degeneracy with energy, the system will reach the canonical equilibrium state. This special scenario is called *canonical contact.* However, first of all let us consider here the more general situation of an energy exchange contact condition, without any assumptions for the spectrum of the environment.

9.2.1 The Accessible and the Dominant Regions

Our approach to the "energy exchange conditions" will be based on similar techniques as before. The possibility of a partition according to Sect. 7.1 is still assumed. However, now there is no further constraint on 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 of finding the total system at some energy E,

$$W(E) := \sum_{A,B/E} W_{AB} = \sum_{A,B/E} \sum_{a,b} \left|\psi_{ab}^{AB}\right|^2 , \tag{9.27}$$

should be conserved (see (7.9) where $A, B/E$ stands for: all A, B such that $E_A^{\mathrm{g}} + E_B^{\mathrm{c}} = E$). This constraint is nothing but the overall energy conservation.

One could try to repeat the above calculation under the energy conservation constraint, but now it turns out that the average purity over the accessible region is no longer close to the actual minimum purity. Furthermore, the energy probability distribution of the individual system considered is no longer a constant of motion. Thus, we proceed in a slightly different way:

1. Contrary to the microcanonical case, the probability of finding the gas (container) subsystem at some given energy is no longer a constant of motion here. However, we are going to prove that there is a predominant distribution, $\{W_{AB}^{\mathrm{d}}\}$, which almost all states within the allowed region have in common. The subregion formed by these states will be called the "dominant region" (see Sect. 9.2.2).
2. Having identified the "dominant region" we will demonstrate that this region is by far the biggest subregion in the accessible region of the system (see Sect. 9.2.3 and Fig. 9.2).
3. 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, which is specified by a fixed energy probability distribution for the considered system, feature the maximum local entropy that is consistent with the predominant distribution. Out of this analysis we get the equilibrium state of the considered system (see Sect. 9.2.4).

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

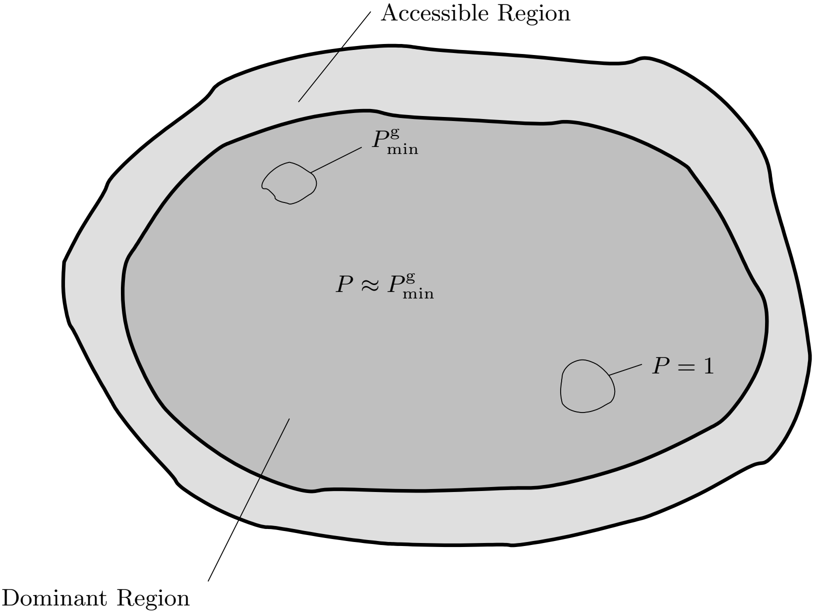


Fig. 9.2. Qualitative picture of the purity landscape. In the canonical case the accessible region contains a dominant region which almost entirely fills the accessible region. Within the dominant region, all states feature the same energy probability distribution. Thus all topology from the microcanonical case (cf. Fig. 9.1) transfers to the dominant region.

reduced local state $\hat{\rho}^{\mathrm{g}}(t)$ as a function of the full state $|\psi(t)\rangle$ should always evolve into a state with a fixed probability distribution W_A, 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 (8.7) in Hilbert space.

9.2.2 Identification of the Dominant Region

First, we calculate the size of a region in Hilbert space that is associated with a certain energy probability distribution $\{W_{AB}\}$. In order to find the predominant distribution $\{W^{\mathrm{d}}_{AB}\}$, 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.

To calculate the size of the region associated with the energy distribution $\{W_{AB}\}$ according to (7.8), we again use the set of independent parameters $\{\phi_i\}$. Due to the condition (9.27) the set $\{\phi_i\}$ can be rearranged with respect to AB-subspaces ($\{\phi_i^{AB}\}$) and therefore the functional matrix F factorizes again. For the size of the respective region with the energy distribution $\{W_{AB}\}$ we get

$$\mathcal{V}\left(\{W_{AB}\}\right) = \prod_{AB} \int \det \mathsf{F}^{AB} \prod_{i} \mathrm{d}\phi_i^{AB} \; . \tag{9.28}$$

These integrals are just surfaces of hyperspheres and can be evaluated using the techniques described in the App. A.3:

$$\mathcal{V} = \prod_{AB} \left(W_{AB}\right)^{N_A N_B - 1/2} \mathcal{O}(1, 2N_A N_B) \; . \tag{9.29}$$

Here $\mathcal{O}(1, 2N_A N_B)$ is the surface area of a $2N_A N_B$-dimensional hypersphere of radius $R = 1$.

Instead of maximizing $\mathcal{V}$ directly we choose to maximize $\ln \mathcal{V}$; this is equivalent, since the logarithm is a monotonous function. Additionally we set $N_A N_B - 1/2 \approx N_A N_B$, an approximation that is not necessary but simplifies the calculation and is definitely valid for large degrees of degeneracy:

$$\ln \mathcal{V} = \sum_{A,B} \big(N_A N_B \ln W_{AB} + \ln \mathcal{O}(1, 2N_A N_B) \big) \, . \tag{9.30}$$

Furthermore we drop all terms that do not depend on $\{W_{AB}\}$, since they are of no relevance for the maximum. Introducing the Lagrange multipliers $\{\lambda_E\}$ to account for the condition of overall energy conservation, the function we want to maximize with respect to the $\{W_{AB}\}$ reads

$$\ln \mathcal{V} = \sum_{A,B} N_A N_B \ln W_{AB} - \sum_{E} \lambda_E \left(\sum_{A,B/E} W_{AB} - W(E) \right) \; . \tag{9.31}$$

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

$$\frac{\partial \ln \mathcal{V}}{\partial W_{AB}} = 0 \tag{9.32}$$

and yields

$$W_{AB}^{\mathrm{d}} = \frac{N_A N_B}{\lambda_{E = E_A^{\mathrm{g}} + E_B^{\mathrm{c}}}} \; . \tag{9.33}$$

Finally, using (9.27) we find for the Lagrange multipliers

$$\lambda_E = \frac{N(E)}{W(E)} \quad \text{with} \quad N(E) = \sum_{A,B/E} N_A N_B \; . \tag{9.34}$$

To check the character of the extremum one can additionally consider the second derivative of (9.31),

$$\frac{\partial^2 \ln \mathcal{V}}{\partial W_{AB}^2} = -\frac{N_A N_B}{W_{AB}^2} < 0 \,. \tag{9.35}$$

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, $\{W_{AB}^{\mathrm{d}}\}$.

9.2.3 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 one. However, for our argument we need to show that this region $\mathcal{V}^{\mathrm{d}}$ is, indeed, much 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 $\{W_{AB}^{\mathrm{d}}\}$. Therefore we consider $W_{AB} =: W_{AB}^{\mathrm{d}} + \epsilon_{AB}$, where the ϵ_{AB}'s are supposed to be small and

$$\sum_{A,B/E} \epsilon_{AB} = 0 \tag{9.36}$$

to guarantee that the new W_{AB} still belongs to the accessible region. For $\ln \mathcal{V}$ we then find

$$\ln \mathcal{V} = \sum_{A,B} N_A N_B \ln(W_{AB}^{\mathrm{d}} + \epsilon_{AB}) \,. \tag{9.37}$$

It is possible to replace the sum over A, B by a sum over all E and the respective subspaces $A, B/E$. Additionally expanding the logarithm to second order we get

$$\ln \mathcal{V} \approx \sum_{E} \sum_{A,B/E} N_A N_B \left(\ln(W_{AB}^{\mathrm{d}}) + \frac{\epsilon_{AB}}{W_{AB}^{\mathrm{d}}} - \frac{1}{2} \left(\frac{\epsilon_{AB}}{W_{AB}^{\mathrm{d}}} \right)^2 \right) \,. \tag{9.38}$$

Since the expansion is around an extremum the linear term should vanish. Indeed, using (9.36) the second summation over this term yields

$$\sum_{A,B/E} N_A N_B \frac{\epsilon_{AB}}{W_{AB}^{\mathrm{d}}} = \sum_{A,B/E} \lambda_E \, \epsilon_{AB} = \lambda_E \sum_{A,B/E} \epsilon_{AB} = 0 \,. \tag{9.39}$$

Thus, using (9.33) and (9.34), we finally find

$$\mathcal{V} \approx \mathcal{V}^{\mathrm{d}} \prod_{AB} \exp\left(-\frac{\left(\sum_{A,B/E} N_A N_B \right)^2}{2 N_A N_B W(E)} \, \epsilon_{AB}^2 \right) \tag{9.40}$$

i.e., regions, $\mathcal{V}$, that correspond to energy probability distributions that deviate from the dominant one are smaller than the dominant region, $\mathcal{V}^{\mathrm{d}}$. Since the smallest factor that can appear in the exponent of (9.40) for given N_A and N_B is $(N_A N_B)/(2W(E))$, the regions $\mathcal{V}$ will be much smaller already for very small deviations, if the corresponding $N_A N_B$ is large. This is another prerequisite for a system to be thermodynamical.

9.2.4 The 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

$$W_A^{\mathrm{d}} = \sum_B W_{AB}^{\mathrm{d}} = \sum_{B/E} \frac{N_A N_B}{\lambda_E} = N^{\mathrm{g}}(E_A^{\mathrm{g}}) \sum_{B/E} \frac{N^{\mathrm{c}}(E_B^{\mathrm{c}}) W(E)}{N(E)} , \tag{9.41}$$

where again the sum over B/E denotes a summation under the condition $E = E_A^{\mathrm{g}} + E_B^{\mathrm{c}}$, and $N_A = N^{\mathrm{g}}(E_A^{\mathrm{g}})$ and $N_B = N^{\mathrm{c}}(E_B^{\mathrm{c}})$ the respective degeneracies. Since $E_B^{\mathrm{c}} = E - E_A^{\mathrm{g}}$ is a function of E for fixed E_A^{g} we switch from a summation over B to a summation over E,

$$W_A^{\mathrm{d}} = N^{\mathrm{g}}(E_A^{\mathrm{g}}) \sum_E \frac{N^{\mathrm{c}}(E - E_A^{\mathrm{g}}) W(E)}{N(E)} . \tag{9.42}$$

This is the energy probability distribution for the gas system that one will find with overwhelming probability for a thermodynamic 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

$$W_B^{\mathrm{d}} = N^{\mathrm{c}}(E_B^{\mathrm{c}}) \sum_E \frac{N^{\mathrm{g}}(E - E_B^{\mathrm{c}}) W(E)}{N(E)} . \tag{9.43}$$

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. The equilibrium state is still undetermined. Once the trajectory has entered the dominant region, we can assume that the trajectory will practically never leave it, because this region fills almost the whole accessible region of the system. However, since all states within the dominant region feature the same energy probability distribution, motion within the dominant region will never give rise to any further energy exchange between the subsystems. As a consequence the situation is as if the system were controlled by microcanonical conditions.

Therefore, we can take the arguments from Sect. 9.1.5 to identify the equilibrium state. Following this idea, we can start with (9.24) and use the dominant energy distribution W^{d}_{AB}, finding for the Hilbert space average of the purity of the gas

$$[\![P^{\mathrm{g}}]\!] \approx \sum_A \frac{(W^{\mathrm{d}}_A)^2}{N_A} + \sum_B \frac{(W^{\mathrm{d}}_B)^2}{N_B} \,. \tag{9.44}$$

Again it is possible to conclude that the second term (due to the environment) is much smaller than the first one for a sufficiently degenerate environment. The first term is exactly the minimum purity of the gas system within the dominant region. Thus, almost all states from the dominant region will also yield approximately the same local gas state. This equilibrium state $\hat{\rho}^{\mathrm{g}}_{\mathrm{eq}}$ is again the state of minimum purity (maximum entropy, see Sect. 9.1.3) that is consistent with the dominant energy distribution,

$$\hat{\rho}^{\mathrm{g}}_{\mathrm{eq}} \approx \sum_{A,a} \frac{W^{\mathrm{d}}_A}{N_A} |A,a\rangle\langle A,a| \,. \tag{9.45}$$

One problem remains: the dominant energy probability distribution W^{d}_A (9.42) 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 (9.42) and thus even (9.45). Normally the canonical contact of standard thermodynamics leads to an equilibrium state, which does not depend on the initial state. Here we have considered a more general contact scenario from which the canonical contact seems to be a special subclass, as we will demonstrate in the next section.

9.3 Canonical Conditions

In the last section we have investigated a situation, for which an energy transfer between the system and its environment was allowed. With these constraints alone it does not seem possible to end up in an equilibrium state that does not depend on the initial state. For a canonical situation the gas system should be found in the canonical equilibrium state, independent of the initial conditions. This behavior, however, can be found, if we take a further condition into account: a special form of the degeneracy of the environment N_B.

So, for the moment, we assume an exponential increase of the container degeneracy

$$N_B = N^{\mathrm{c}}_0 \, \mathrm{e}^{\alpha E^{\mathrm{c}}_B} \,, \tag{9.46}$$

where α, N^{c}_0 are some constants. Let us start again from (9.42) using (9.46) for the degeneracy of the environment

$$W_A^{\mathrm{d}} = N_A \, \mathrm{e}^{-\alpha E_A^{\mathrm{g}}} \sum_E \frac{N_0^{\mathrm{c}} \, \mathrm{e}^{\alpha E} W(E)}{N(E)} \,. \tag{9.47}$$

Obviously, the sum does not depend on A at all. Since W_A^{d} has been constructed as some probability distribution it is still normalized by definition. Therefore the sum has to reduce to a normalizing factor. This could also be shown in a rather lengthy calculation, which we skip here. Finally we get for the dominant energy probability distribution of the gas system

$$W_A^{\mathrm{d}} = \frac{N_A \, \mathrm{e}^{-\alpha E_A^{\mathrm{g}}}}{\sum_A N_A \mathrm{e}^{-\alpha E_A^{\mathrm{g}}}} \,. \tag{9.48}$$

This result is no longer dependent on the initial state! The justification for the degeneracy structure of the environmental system (9.46) for thermodynamic systems will be discussed later in Sect. 11.2 and Sect. 11.4.

The energy probability distributions of almost all states from the accessible region consistent with the constraints is then the canonical distribution. Since the argumentation for the minimal purity state (state of maximal entropy) remains unchanged, the equilibrium state now reads

$$\hat{\rho}_{\mathrm{eq}}^{\mathrm{g}} \approx \frac{1}{\sum_A N_A \mathrm{e}^{-\alpha E_A^{\mathrm{g}}}} \sum_{A,a} \mathrm{e}^{-\alpha E_A^{\mathrm{g}}} |A, a\rangle\langle A, a| \,. \tag{9.49}$$

Obviously, this is the well known canonical equilibrium state with the inverse temperature $\beta = \alpha$.

For some more concrete illustrations of the implications, which the rather abstractly derived principles in this chapter bear on the dynamics of adequate systems, see Sect. 18.2.

9.4 Fluctuations of Occupation Probabilities W_A

Unfortunately the term “fluctuations” has various meanings in the field of physics. In the context of thermostatistics one speaks of thermal fluctuations, meaning that some extensive variable, e.g., energy defined as an average over a distribution with fixed intensive variable, e.g., temperature (see Sect. 3.3.2) might not be exactly sharp. Instead one gets an energy probability distribution peaked at some value, having a certain width. This width is taken to characterize “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 (“uncertainties”). The so-called “vacuum fluctuations” refer to the fact that the probability to measure some finite electromagnetic field intensity, say, does not vanish even in vacuum (i.e., a pure state). Nevertheless again the probability distribution itself is constant in time.

The fluctuations we want to discuss in this section are of a different kind. In our approach all occupation probabilities are explicit functions of time, which reflects the eternal motion of the pure state of the total system within its accessible region. Since the "probability landscape" W_A is not entirely flat, probabilities will vary in time as the state vector wanders around in Hilbert space. These fluctuations in time will be studied here by analyzing the respective probability landscapes.

Fortunately it is possible to calculate precisely the size of regions that are associated with a given W_A, this being equivalent to the probability to find a state featuring this W_A, if states from the accessible region were picked at random. The size of those regions as a function of the concrete probabilities will be called $\mathcal{V}(W_A)$. Any peak of $\mathcal{V}(W_A)$ will then be the most likely value for W_A, and the sharper this peak, the smaller the fluctuations.

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

$$\{W(E)\} = \delta_{E,U} \ . \tag{9.50}$$

Because of this constraint and the overall energy conservation $E = E_A^{\mathrm{g}} + E_B^{\mathrm{c}}$, the index B of the container system can be written as a function of A, $B = B(A)$. Therefore we can write all states of the accessible region in terms of the state of the gas system with the parameter A, as

$$|\psi_{\mathrm{acc}}\rangle = \sum_A \sum_{a,b} \psi_{ab}^{AB(A)} |A, a\rangle \otimes |B(A), b\rangle \ . \tag{9.51}$$

For the dominant probability distribution W_A^{d} we find here from (9.42)

$$\begin{aligned} W_A^{\mathrm{d}} &= N_A \sum_E \frac{N_{B(A)} W(E)}{N(E)} = N_A \sum_E \frac{N_{B(A)}\, \delta_{E,U}}{N(E)} = \frac{N_A N_{B(A)}}{N(U)} \\ &= \frac{N_A N_{B(A)}}{\sum_{A,B/U} N_A N_B} \ , \end{aligned} \tag{9.52}$$

where we have used (9.34). If we take the overall energy conservation into account, the summation over $A, B/U$ in the denominator reduces to a summation over A leading to

$$W_A^{\mathrm{d}} = \frac{N_A N_{B(A)}}{\sum_A N_A N_{B(A)}} \ . \tag{9.53}$$

Since the states of the accessible region have to be normalized, we require

$$\sum_A \sum_{a,b} \left(\left(\eta_{ab}^{AB(A)}\right)^2 + \left(\xi_{ab}^{AB(A)}\right)^2 \right) = 1 \ , \tag{9.54}$$

where $\{\eta_{ab}^{AB(A)}, \xi_{ab}^{AB(A)}\}$ again denote the real and imaginary parts of the amplitudes $\psi_{ab}^{AB(A)}$. Note that this condition restricts the parameter space to a hypersphere with radius one, as has already been discussed several times. In this parameter space the dimension, d, of this hypersphere is

$$d = 2 \sum_A N_A \, N_{B(A)} \,. \tag{9.55}$$

The total Hilbert space is partitioned here into subspaces A, filled with states for which the gas system is to be found at energy E_A^{g}. The dimension d_A of such a subspace is given by

$$d_A = 2 N_A \, N_{B(A)} \,. \tag{9.56}$$

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 (e.g., the partition introduced in Sect. 7.1), it is only a subset of these parameters that determines the probability of the gas system to be found at some energy E_A^{g}. Namely all parameters belonging to the respective subspace A of the Hilbert space with dimension d_A. The probability W_A of finding the gas system at energy E_A^{g} can now be evaluated by summing all probabilities (squares of amplitudes) of the states belonging to the respective subspace A

$$W_A = \sum_{a,b} \left(\left(\eta_{ab}^{AB(A)} \right)^2 + \left(\xi_{ab}^{AB(A)} \right)^2 \right) \,. \tag{9.57}$$

Here we are interested in the size of the corresponding region $\mathcal{V}(W_A)$. We ask: how big is the zone on the surface of the hypersphere of radius one and dimension d, which consists of points for which (9.57) is fulfilled? The size of the zone can be calculated by parametrizing the surface in such a way that W_A is one of the parameters. If now the surface integral is computed and all other parameters are integrated out except for W_A, the remaining function will describe the size of the zones as a function of W_A. We have transferred the complicated integration on the hypersphere to App. A.3, citing here only the result for the size of such a subspace A

$$\mathcal{V}(W_A) = C \left(\sqrt{W_A} \right)^{d_A - 2} \left(\sqrt{1 - W_A} \right)^{d - d_A - 2} \,, \tag{9.58}$$

with some normalization constant C. The function describes the sizes of the respective subspaces A, which is directly proportional to the relative frequency of states featuring W_A. By adequate normalization of $\mathcal{V}(W_A)$, choosing C in such a way that the integral over $\mathcal{V}(W_A)$ is one (see App. A.3), it is now possible to estimate the probability of finding a state that features W_A. In Fig. 9.3 we show the distribution function in dependence on the dimensions d and d_A.

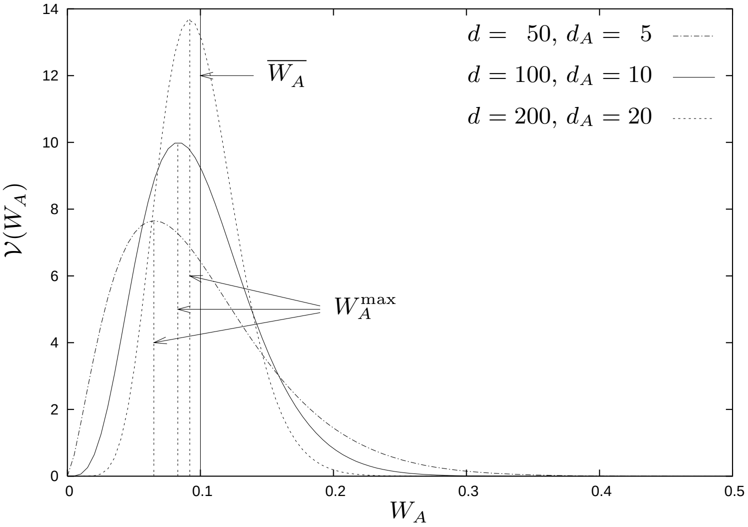


Fig. 9.3. Probability distribution function of finding a state featuring W_A, for three different dimensions d and d_A. The vertical solid line at position 0.1 marks the respective mean value of each distribution.

First of all we find for the W_A, for which the respective subspace A has maximal volume $\mathcal{V}(W_A)$ (see (A.34)),

$$W_A^{\max} = \frac{d_A - 2}{d - 4} = \frac{N_A N_{B(A)} - 1}{\sum_A N_A N_{B(A)} - 2} \ . \tag{9.59}$$

Furthermore, it is possible to evaluate the mean value $\overline{W_A}$ (see (A.35))

$$\overline{W_A} = \int_0^1 W_A\, \mathcal{V}(W_A)\ \mathrm{d}W_A = \frac{d_A}{d} = \frac{N_A N_{B(A)}}{\sum_A N_A N_{B(A)}} \ . \tag{9.60}$$

For large d and d_A, i.e., for large total systems, $W_A^{\max}$ as well as the mean value $\overline{W_A}$ turn out to be identical with W_A^{d} (cf. (9.53)). Thus the energy probabilities of the dominant distribution are indeed the most likely ones, also for each single probability (see Fig. 9.3).

Now, the variance of this distribution, $\mathcal{V}(W_A)$, can be calculated as well, yielding (see (A.36))

$$\Delta W_A = \sqrt{\int_0^1 W_A^2\, \mathcal{V}(W_A)\ \mathrm{d}W_A - \overline{W_A}} = \sqrt{\frac{2\, d_A(d - d_A)}{d^2(d+2)}} \ . \tag{9.61}$$

If the trajectories in Hilbert space were ergodic, ΔW_A would be a direct measure of the fluctuations. Since we know that the trajectories are not ergodic, such a simple connection is not necessarily true. Nevertheless, if ΔW_A

is small, almost all states within the accessible region feature a W_A very near to W_A^{d}, such that fluctuations can be expected to be small. In that sense ΔW_A can be considered a measure of the fluctuations. If the environment gets bigger, both d and d_A, will scale with some factor, say, α, and then ΔW_A will scale like

$$\Delta W_A \propto \sqrt{\frac{1}{\alpha}} \,. \tag{9.62}$$

Typically, this should apply to the fluctuations as well.

Coming back to the case of the total 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 d_A^E of its subset that corresponds to the gas system being at energy E_A^{g}. Note that the considerations so far hold true for each energy subspace E. Thus they all have their own relative frequency $\mathcal{V}^E\big(W_A^E\big)$ of the probability W_A^E (where the superscript E denotes the restriction to the respective energy subspace E). This leads to a set of variances ΔW_A^E, one for each subspace E.

We are interested in the variance of the relative frequency of W_A, the overall probability of finding the gas system at E_A^{g}. This probability W_A now consists of a sum over the product of the probability to find the whole system at energy E and at the same time the gas system in the state E_A^{g}, W_A^E,

$$W_A = \sum_E W(E)\, W_A^E \,. \tag{9.63}$$

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

$$\Delta W_A = \sqrt{\sum_E \big(W(E)\, \Delta W_A^E\big)^2} \,. \tag{9.64}$$

Thus, the full width of the distribution is essentially the geometrical mean of the widths corresponding to the pertinent energy subspaces. This means that the principle stated in (9.62) still applies: the more states there are within the energy subspaces, either in the gas or in the container system, the smaller the fluctuations of the occupation probabilities. In particular, in the limit of the environment becoming arbitrarily large the fluctuations will practically vanish.

For some illustrations of these results, see Sect. 18.3.

10 Interim Summary

> *Experiments cannot be extrapolated, only theories.*
>
> — D. J. Raine and E. G. Thomas [106]

Before we continue with our checklist for a foundation of thermodynamics, we review some of the main ideas and results of the last chapters. These should be seen in connection with former efforts towards a foundation based on classical theories.

10.1 Equilibrium Properties

So far we have developed a theory to predict the behavior of a subsystem, without following the complete time evolution. The considered model is a bipartite quantum system, a small one, the gas g, which we want to observe and a big one, the environment or container c, weakly coupled to each other. Depending on the type of interaction (with or without energy exchange) and on the structure of the environment, we can predict the equilibrium state of the gas system and its purity. The possibility to make predictions is a consequence of the respective Hilbert space and its accessible region only. We do not need any further assumptions like ergodicity, etc. The complete system is eternally in pure state and evolves according to a unitary evolution with constant velocity in Hilbert space, a reversible Schrödinger dynamics.

10.1.1 Microcanonical Contact

We have first considered a microcanonical interaction between a small system (gas g) and an environment (container c) (no energy exchange). The whole system (gas and container) is confined to an accessible region in Hilbert space, restricted by the conservation of energy in each subsystem separately.

Without any additional assumptions, such as those used in damping models, it is then possible to make a prediction about the equilibrium state of the system (g). In spite of the eternal unitary evolution of the state of the total system this subsystem alone can be found at approximately maximum entropy (minimum purity) at any time. Any state belonging to the accessible region of the total Hilbert space has approximately maximum entropy with respect to the subsystem. This type of behavior is found to be generic for a very large class of even small quantum systems and agrees with the classical

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 109–112 (2004)
http://www.springerlink.com/

thermodynamic expectations according to the second law of thermodynamics. All subsystems g, which meet these conditions, are called “thermodynamic systems”. Note that this thermodynamic behavior is not enforced by an ad hoc introduced relaxation rate or something like that but is just a property of fundamental quantum mechanics in such a bipartite system under very weak restrictions. For some more concrete illustrations, see 18.1.

Equations for Microcanonical Contact (cf. Sect. 9.1):

- Accessible region:

$$\sum_{a,b} \left|\psi_a^A\right|^2 \left|\psi_b^B\right|^2 = W_A W_B \ . \tag{10.1}$$

- Hilbert space average of P^{g} in the accessible region:

$$[\![P^{\mathrm{g}}]\!] \approx \sum_A \frac{(W_A)^2}{N_A} + \sum_B \frac{(W_B)^2}{N_B} \ . \tag{10.2}$$

- Equilibrium state, minimum purity state:

$$\hat{\rho}^{\mathrm{g}}_{\mathrm{min}} = \sum_{A,a} \frac{W_A}{N_A} |A,a\rangle\langle A,a| \ , \tag{10.3}$$

- Minimum purity, maximum entropy:

$$P^{\mathrm{g}}_{\mathrm{min}} = \sum_A \frac{W_A^2}{N_A} \ , \qquad S^{\mathrm{g}}_{\mathrm{max}} = -k_{\mathrm{B}} \sum_A W_A \ln \frac{W_A}{N_A} \ . \tag{10.4}$$

10.1.2 Energy Exchange Contact, Canonical Contact

If we allow the system g and the environment c to exchange energy, the considerations are a little bit more complicated. The accessible region is no longer filled with states of maximum entropy only, there are several different areas, where the entropy takes on even bigger values than on average. However, we discover a very large region, which fills almost the whole accessible state space of the system, called the “dominant region”. Since the dominant region is exponentially bigger than all other regions and the state wanders through the whole accessible region with constant velocity, we can predict that the state should be found mostly in the dominant region. Within the dominant region each state features the same energy distribution and thus the system no longer exchanges energy with the environment. This situation is now “quasi” microcanonical and we can thus transcribe the respective results to the dominant region. Therefore we can predict the entropy in the dominant region to be maximal (the system has minimum purity) and the equilibrium state of the system is similar to the microcanonical equilibrium state but with the energy distribution of the dominant region.

However, the dominant energy distribution is not necessarily the canonical energy distribution, wherefore we firstly talk about “energy exchange contact”. However, there is a large variety of systems, which do have a canonical energy distribution in their dominant region. For this to happen the environment must have an exponential increase of the degeneracy with energy. This exponential increase is quite intuitive for a lot of physical situations, as will be discussed in the next chapter. For a concrete model exemplifying these considerations see Sect. 18.2.

Equations for Energy Exchange Contact (cf. Sect. 9.2):

- Accessible region:

$$W(E) := \sum_{A,B/E} W_{AB} = \sum_{A,B/E} \sum_{a,b} \left|\psi_{ab}^{AB}\right|^2 . \tag{10.5}$$

- Dominant probability distribution:

$$W_{AB}^{\mathrm{d}} = \frac{N_A N_B W(E)}{\sum_{A,B/E} N_A N_B} . \tag{10.6}$$

- Dominant probability distribution for the gas system:

$$W_A^{\mathrm{d}} = N^{\mathrm{g}}(E_A^{\mathrm{g}}) \sum_E \frac{N^{\mathrm{c}}(E - E_A^{\mathrm{g}}) W(E)}{N(E)} . \tag{10.7}$$

- Hilbert space average of P^{g} in the dominant region:

$$[\![P^{\mathrm{g}}]\!] \approx \sum_A \frac{(W_A^{\mathrm{d}})^2}{N_A} + \sum_B \frac{(W_B^{\mathrm{d}})^2}{N_B} . \tag{10.8}$$

- Equilibrium state, minimum purity state:

$$\hat{\rho}_{\mathrm{eq}}^{\mathrm{g}} \approx \sum_{A,a} \frac{W_A^{\mathrm{d}}}{N_A} |A,a\rangle\langle A,a| . \tag{10.9}$$

Equations for Canonical Contact (cf. Sect. 9.3):

- Dominant probability distribution for the gas system:

$$W_A^{\mathrm{d}} = \frac{N_A \, \mathrm{e}^{-\alpha E_A^{\mathrm{g}}}}{\sum_A N_A \mathrm{e}^{-\alpha E_A^{\mathrm{g}}}} . \tag{10.10}$$

- Equilibrium state, minimum purity state:

$$\hat{\rho}_{\mathrm{eq}}^{\mathrm{g}} \approx \frac{1}{\sum_A N_A \mathrm{e}^{-\alpha E_A^{\mathrm{g}}}} \sum_{A,a} \mathrm{e}^{-\alpha E_A^{\mathrm{g}}} |A,a\rangle\langle A,a| . \tag{10.11}$$

10.2 Local Equilibrium States and Ergodicity

As explained in Sect. 6.3, ergodicity is not needed in our approach, while 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 thermodynamic system, the considered local system will under microcanonical conditions most likely be found in the momentary equilibrium state (state of minimum purity and maximum entropy)

$$\hat{\rho}^{\mathrm{g}}_{\mathrm{eq}} \approx \sum_{A,a} \frac{W_A}{N_A} |A,a\rangle\langle A,a| \, . \tag{10.12}$$

Obviously all states $|A,a\rangle$, which belong to one "energy shell" E^{g}_A, have the same probability W_A/N_A. These states are the same as those one would have found, if one had applied the "a priori postulate" (Sect. 3.3.1). They are also the same states as obtained, 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 $\hat{\rho}^{\mathrm{g}}_{\mathrm{eq}}$ it is even irrelevant in which space one assumes this ergodicity. One could take as the possible states of the system, in each of which the system will be found for the same time, the orthogonal energy basis states or one could allow for superpositions, in both cases one would find $\hat{\rho}^{\mathrm{g}}_{\mathrm{eq}}$ as a result.

To sum up, one can say that the effect of some adequate, microcanonically coupled environment is that the considered system behaves locally as if it were perfectly ergodic, within arbitrarily short time periods. Alternatively one could describe the result by an ensemble, i.e., many copies of the same system, the same number of copies being in every pure state accessible to the system. However, in the present picture the single considered system is simply highly entangled with its environment.

For a thermodynamic system under energy exchange contact conditions, the momentary equilibrium state is

$$\hat{\rho}^{\mathrm{g}}_{\mathrm{eq}} \approx \sum_{A,a} \sum_{E} \frac{N^c(E - E^{\mathrm{g}}_A)\, W(E)}{N(E)} \, |A,a\rangle\langle A,a| \, . \tag{10.13}$$

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

Further examples for weakly coupled bipartite systems have been studied using also entirely different techniques [17, 58, 124]. The results are similar to those described in this chapter.

11 Typical Spectra of Large Systems

> *... the positions and velocities of every particle of a classical plasma or a perfect gas cannot be observed, nor can they be in an atom nor in a molecule; the theoretical requirement now is to find the gross macroscopic consequences of states that cannot be observed in detail.*
>
> — A. Cook [27]

11.1 The Extensivity of Entropy

If a set of axioms is formulated as a basis of thermodynamics, one is usually told that entropy has to be an extensive quantity. This basically means that if two identical systems with entropy S are brought in contact such as to form a system of twice the size of the original systems, the entropy S^{tot} of the joint system should double,

$$S^{\text{tot}} = 2S \; . \tag{11.1}$$

Formulated more rigorously this means that entropy should be a homogeneous function of the first order, or that it should be possible to write it as a function of the other extensive variables, say, energy U, volume V and particle number N as

$$S = N \; s(U/N, V/N) \; , \tag{11.2}$$

where $s(U/N, V/N)$ is the entropy of a single particle. This is obviously an important property, since it guarantees, e.g., that temperature defined in the usual way (see (3.17))

$$T = \frac{\partial U}{\partial S} \tag{11.3}$$

remains the same under this procedure, i.e., temperature is an intensive quantity.

However, this basic requirement faces severe problems for the standard definition of entropy as considered in the following. The classical definition of entropy for the microcanonical ensemble (see Sect. 3.3.2) reads

$$S = k_{\mathrm{B}} \ln m \approx k_{\mathrm{B}} \ln G(U) \; , \tag{11.4}$$

where m denotes the number of micro states consistent with the energy U, i.e., the volume of the corresponding energy shell in phase space, divided by the volume of some elementary cell, and $G(U)$ the state density. In our approach the same formula holds (for a sharp energy probability distribution) for the

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 113–122 (2004)
`http://www.springerlink.com/`

equilibrium entropy (see (9.26)), except $G(U)$ being the quantum mechanical energy state density at the energy U.

Regardless of whether we are following classical or quantum mechanical ideas, if one assumes that the thermal contact of two identical systems, while containing only negligible energy by itself, allows for energy exchange between the systems, the entropy S^{tot} of the doubled system at the double energy could be calculated from the state density by the convolution

$$S^{\mathrm{tot}} = k_{\mathrm{B}} \ln \int_0^{2U} G(E) G(2U - E) \, \mathrm{d}E \, . \tag{11.5}$$

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

$$k_{\mathrm{B}} \ln \int_0^{2U} G(E) G(2U - E) \, \mathrm{d}E \neq 2 k_{\mathrm{B}} \ln G(U) \, . \tag{11.6}$$

This could only be true, if the function $G(E)G(2U - E)$ were extremely peaked at $E = U$. In general, however, there is no reason to assume this, even if $G(E)$ were a rapidly growing function. If $G(E)$ grows exponentially, the integrand of the convolution is flat, rather than peaked. The identity of (11.6) is often claimed in standard textbooks by referring to the ideal gas, for which it happens to be approximately true, or by complicated considerations based on the canonical ensemble [22]. All this, however, is not a straightforward, general extensivity proof for the microcanonical case. So, according to those definitions, one cannot claim without further study that entropy is an extensive quantity. (This problem is not to be confused with Gibbs' paradox that can be solved by using Boltzmann statistics of identical particles; here dividing the left hand side of (11.6) by some function of N will not fix the problem [127].)

Finally one is often referred to Shannon entropy

$$S^{(\mu)} = -k_{\mathrm{B}} \sum_i W_i^{(\mu)} \ln W_i^{(\mu)} \, , \tag{11.7}$$

which appears to be extensive, since (11.1) holds, if $W_l^{(12)} = W_i^{(1)} W_j^{(2)}$. However, this means that the probabilities of finding the systems in their individual states should be uncorrelated. This is clearly 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 $U - E$.

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

11.2 Spectra of Modular Systems

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

Let us, as an example, consider the sequential build-up of some piece of solid material. First, we have one atom with some energy spectrum. If we bring two atoms together, the spectrum of the resulting molecule will be substantially different from the energy spectrum of the two separate atoms. The energy resulting from the binding can be as large as typical level splitting within the spectrum of the separate atoms. However, 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 broad energy spectra, typically containing bands that stretch over a considerable energy region with a smooth state density. If we now combine these parts again, 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 of the separate spectra caused by the binding will be negligible. This is the limit beyond which the weak coupling limit applies. This limit is always assumed to hold in thermodynamics. For the contact between a system and its environment it is thus assumed that the spectra of the separate systems are almost undistorted by the contact. So, this principle should apply to the different identical parts of one system above some size. Here we assume that there are a lot of parts above this limit to make up a macroscopic system, as is the case in our example, where there are a lot of parts containing some 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 are strongly interacting. In the case of a gas no additional consideration is necessary, for its spectrum can naturally be understood as the combined spectrum of all the individual gas particles.

We finally analyze the properties of spectra that result from very many identical non-interacting systems. Just as 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, $g(E)$. Defining

$$\mathcal{C}_N\{g(E)\}(U) := \big(g(E) * g(E) * \cdots * g(E)\big)(U) \tag{11.8}$$

as the convolution of N identical functions $g(E)$, where the convolution labeled by "$*$" is mathematically defined by the integration

$$\mathcal{C}_N\{g(E)\}(U) := \int \cdots \int g(E_1)g(E_2 - E_1) \cdots g(E_i - E_{i-1}) \cdots g(U - E_{N-1}) \prod_{j=1}^{N-1} \mathrm{d}E_j \ . \tag{11.9}$$

Thus we can write

$$G(U) = \mathcal{C}_N\{g(E)\}(U) \ . \tag{11.10}$$

To evaluate this convolution, we start by considering another convolution. We define

$$r(E) := \frac{\mathrm{e}^{-\alpha E} g(E)}{\int \mathrm{e}^{-\alpha E} g(E) \mathrm{d}E} \tag{11.11}$$

and the quantities

$$R := \int \mathrm{e}^{-\alpha E} g(E) \, \mathrm{d}E \ , \tag{11.12}$$

$$\overline{r} := \int E \, r(E) \, \mathrm{d}E \ , \tag{11.13}$$

$$\sigma^2 := \int E^2 \, r(E) \, \mathrm{d}E - \overline{r}^2 \ . \tag{11.14}$$

If the increase of $g(E)$ with energy is not faster than exponential, which we have to assume here, then all these quantities are finite and, since $r(E)$ is normalized, $\overline{r}$ is the mean value of $r(E)$ and σ^2 is the variance of $r(E)$. Now, consider the convolution of all $r(E)$ written as

$$\mathcal{C}_N\{r(E)\}(U) = \frac{\mathrm{e}^{-\alpha U} G(U)}{R^N} \ . \tag{11.15}$$

To evaluate $\mathcal{C}_N\{r(E)\}(U)$ we exploit properties typical for a convolution. Since the integral over a convolution equals the product of the integrals of the convoluted functions, we have

$$\int \mathcal{C}_N\{r(E)\}(U) \, \mathrm{d}U = 1 \ . \tag{11.16}$$

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

$$M := \int U \mathcal{C}_N\{r(E)\}(U) \, \mathrm{d}U = N \, \overline{r} \ . \tag{11.17}$$

As the square of the variance of a convolution of normalized functions is the sum of the squares of the convoluted functions, we finally get

$$\varSigma^2 := \int U^2 \mathcal{C}_N\{r(E)\}(U)\,\mathrm{d}U - N^2\,\overline{r}^2 = N\,\sigma^2\ . \tag{11.18}$$

The Fourier transform of two convoluted functions equals the product of the Fourier transforms of the convoluted functions. If for simplicity we define the Fourier transform of a function $r(E)$ as $\mathcal{F}\{r(E)\}$, we thus find

$$\mathcal{F}\{\mathcal{C}_N\{r(E)\}\} = \big(\mathcal{F}\{r(E)\}\big)^N\ . \tag{11.19}$$

If $r(E)$ is integrable, $\mathcal{F}\{r(E)\}$ is integrable as well and it is very likely that the function $\mathcal{F}\{r(E)\}$ has a single global maximum somewhere. This maximum should become much more predominant, if the function is multiplied very many times with itself, regardless of how strongly peaked the maximum originally was. This means that the function $\mathcal{F}\{\mathcal{C}_N\{r(E)\}\}$ should get extremely peaked at some point, if N becomes large enough. One can show (see App. D) that this peak, containing almost all of the area under the curve, is approximately Gaussian. One can now split $\mathcal{F}\{\mathcal{C}_N\{r(E)\}\}$ up into two parts, the Gaussian and the rest. Since a Fourier transform is additive leaving the area under the square of the curve invariant, and transforming a Gaussian into a Gaussian, $\mathcal{C}_N\{r(E)\}$ should again mainly consist of a 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, $\mathcal{F}\{\mathcal{C}_N\{r(E)\}\}$ should be almost entirely dominated by the Gaussian part. At the edges, where the Gaussian vanishes, the small remainder may dominate. If we assume that the integral, the mean value and the variance of $\mathcal{F}\{\mathcal{C}_N\{r(E)\}\}$ are entirely dominated by its Gaussian part, we can, using (11.16), (11.17) and (11.18), give a good approximation for $\mathcal{F}\{\mathcal{C}_N\{r(E)\}\}$ that should be valid at the peak, i.e., around $U = N\,\overline{r}$

$$\mathcal{C}_N\{r(E)\}(U) \approx \frac{1}{\sqrt{2\pi N\sigma^2}}\,\exp\!\left(-\frac{(U - N\,\overline{r})^2}{2N\sigma^2}\right)\ . \tag{11.20}$$

Solving (11.15) for $G(U)$ and inserting (11.20), evaluated at the peak, we thus find

$$G(N\,\overline{r}) \approx \frac{R^N \mathrm{e}^{\alpha N\,\overline{r}}}{\sqrt{2\pi N\sigma^2}}\ , \tag{11.21}$$

where $\overline{r}, R$ and σ are all functions of α. Thus, we have expressed G as a function of α. Since we want G as a function of the internal energy U, we define

$$U := N\,\overline{r}(\alpha) \quad \text{or} \quad \frac{U}{N} = \overline{r}(\alpha)\ . \tag{11.22}$$

Solving formally for α we get

$$\alpha = \alpha(U/N) = \overline{r}^{-1}(U/N)\ . \tag{11.23}$$

Now R, σ and α are all functions of the argument (U/N) and we can rewrite (11.21) as

$$G(U) \approx \frac{\big(R(U/N)\big)^N \mathrm{e}^{\alpha(U/N)U}}{\sqrt{2\pi N}\sigma(U/N)} \,, \tag{11.24}$$

or, by taking the logarithm

$$\ln G(U) \approx N \left(\ln R(U/N) + \frac{U}{N}\,\alpha(U/N) \right) - \frac{1}{2}\ln(2\pi N) - \ln\sigma(U/N) \,. \tag{11.25}$$

If we keep U/N fixed, but let $N \gg 1$, which amounts to a simple upscaling of the system, we can neglect everything except for the first part on the right hand side of (11.25) to get

$$\ln G(U) \approx N \left(\ln R(U/N) + \frac{U}{N}\,\alpha(U/N) \right) \,. \tag{11.26}$$

This is obviously a homogeneous function of the first order and thus an extensive quantity. Therefore, (11.2) is finally confirmed.

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

11.3 Entropy of an Ideal Gas

To check (11.26) we consider a classical ideal gas, just taking the spectrum of a free particle in one dimension as the function to be convoluted. The total energy of a classical gas depends on $3N$ degrees of freedom, corresponding to the components of the momenta of all N particles. From the dispersion relation of a classical free particle confined to one dimension

$$E = \frac{1}{2m}\,p^2 \,, \tag{11.27}$$

where m is the mass of a single particle, we find

$$\frac{\mathrm{d}p}{\mathrm{d}E} = \frac{m}{p} = \sqrt{\frac{m}{2E}} \,. \tag{11.28}$$

Since there are two momenta corresponding to one energy and taking h as the volume of an elementary cell, we get for a particle restricted to the length L the state density

$$g(E) = \frac{L}{h}\sqrt{\frac{2m}{E}} \; . \tag{11.29}$$

With this state density we find, using some standard table of integrals, for the quantities defined in Sect. 11.2

$$R = \frac{L}{h}\sqrt{\frac{2\pi m}{\alpha}} \; , \quad \overline{r} = \frac{1}{2\alpha} \; . \tag{11.30}$$

Setting $\overline{r} = \frac{U}{N'}$ and writing α and R as functions of this argument we get

$$\alpha = \frac{1}{2}\left(\frac{U}{N'}\right)^{-1} \; , \quad R = \frac{L}{h}\left(4m\pi\left(\frac{U}{N'}\right)\right)^{\frac{1}{2}} \; . \tag{11.31}$$

Inserting these results into (11.26) yields

$$\ln G(U) = N'\left(\ln\frac{L}{h} + \frac{1}{2}\ln\left(4m\pi\left(\frac{U}{N'}\right)\right) + \frac{1}{2}\right) \; . \tag{11.32}$$

Relating the number of degrees of freedom N' to the number of particles N by $N' = 3N$ we eventually find

$$\ln G(U) = N\left(3\ln\frac{L}{h} + \frac{3}{2}\ln\left(\frac{4}{3}m\pi\left(\frac{U}{N}\right)\right) + \frac{3}{2}\right) \; . \tag{11.33}$$

This is exactly the standard textbook result (without the corrected Boltzmann statistics, see e.g., [127]), which is usually calculated by evaluating the surface area of hyperspheres and using the Stirling formula.

11.4 The Boltzmann Distribution

In Sect. 9.2.4 we found an equilibrium energy probability distribution for a canonical situation, i.e., a contact, which allows for energy exchange with a large surrounding. This distribution is given by (9.42) and reads, written now in terms of state densities

$$W^{\mathrm{d}}(E^{\mathrm{g}}) = G^{\mathrm{g}}(E^{\mathrm{g}}) \int \frac{G^{\mathrm{c}}(E - E^{\mathrm{g}})W(E)}{G(E)} \, \mathrm{d}E \; . \tag{11.34}$$

Here, $G^{\mathrm{g}}(E^{\mathrm{g}})$ and $G^{\mathrm{c}}(E^{\mathrm{c}})$ are the state densities of the gas and the container system, respectively, whereas $G(E)$ is the state density of the total combined system. This is obviously not the familiar Boltzmann distribution. Instead of the Boltzmann factor $\exp(-E^{\mathrm{g}}/k_{\mathrm{B}}T)$ one has here a factor depending on the state density of the environment. We are now going to analyze this factor under the assumption that the spectrum of the environment has the typical structure as established in Sect. 11.2.

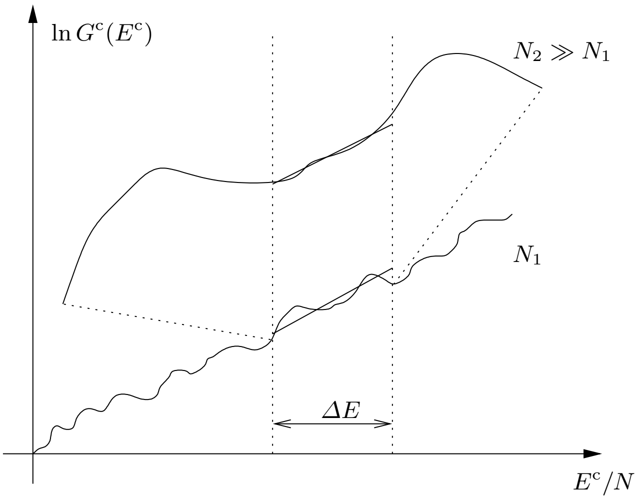


Fig. 11.1. Upscaling of the graph $\ln G^{\mathrm{c}}(E^{\mathrm{c}})$ with increasing N; the original section within ΔE gets stretched. With respect to the same ΔE the new graph gets closer to a linear approximation (straight line).

If the environment is a large system, it should be possible to write the logarithm of its state density according to (11.2) as

$$\ln G^{\mathrm{c}}(E^{\mathrm{c}}) = N s^{\mathrm{c}}(E^{\mathrm{c}}/N) \;, \tag{11.35}$$

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 clearly seen that increasing N just amounts to an upscaling of the whole picture. This means that the graph becomes smoother and smoother within finite energy intervals (see Fig. 11.1).

This can be stated in a more mathematical form by checking the expansion of $\ln G^{\mathrm{c}}(E^{\mathrm{c}})$ around some point of fixed energy per unit, $E^{\mathrm{c}}/N = \epsilon$

$$\ln G^{\mathrm{c}}(E^{\mathrm{c}}) \approx N s^{\mathrm{c}}\big|_{\epsilon} + \left.\frac{\mathrm{d}s^{\mathrm{c}}}{\mathrm{d}E^{\mathrm{c}}}\right|_{\epsilon} (E^{\mathrm{c}} - N\epsilon) + \frac{1}{2N}\left.\frac{\mathrm{d}^2 s^{\mathrm{c}}}{\mathrm{d}(E^{\mathrm{c}})^2}\right|_{\epsilon} (E^{\mathrm{c}} - N\epsilon)^2 + O(\epsilon^3) \;. \tag{11.36}$$

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 after the linear term will be a valid approximation over a wide energy range, with the range of validity becoming larger with increasing N. Without $\ln G^{\mathrm{c}}(E^{\mathrm{c}})$ being a homogeneous function of the first order such a truncation would remain without justification, although it is often routinely used [22].

In (11.34) the function $W(E)/G(E)$ is multiplied by 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 $\ln G^{\mathrm{c}}(E^{\mathrm{c}})$ is valid, we might replace the state density of the environment in (11.34) by an exponential that the linearization (11.36) gives rise to. The expansion then has to be around the energy, where $W(E)/G(E)$ is peaked, i.e., $N\epsilon$ has to be chosen to lie in the center of the peak. This replacement yields

$$\begin{aligned} W^{\mathrm{d}}(E^{\mathrm{g}}) &\approx G^{\mathrm{g}}(E^{\mathrm{g}}) \\ &\int \frac{1}{G(E)} \exp\left(N s^{\mathrm{c}}\big|_{\epsilon} + \frac{\mathrm{d}s^{\mathrm{c}}}{\mathrm{d}E^{\mathrm{c}}}\bigg|_{\epsilon} \left((E - E^{\mathrm{g}}) - N\epsilon\right)\right) W(E)\,\mathrm{d}E \,, \end{aligned} \tag{11.37}$$

or, with the terms depending on E^{g} taken out of the integral

$$\begin{aligned} W^{\mathrm{d}}(E^{\mathrm{g}}) &\approx G^{\mathrm{g}}(E^{\mathrm{g}}) \exp\left(- \frac{\mathrm{d}s^{\mathrm{c}}}{\mathrm{d}E^{\mathrm{c}}}\bigg|_{\epsilon} E^{\mathrm{g}} \right) \\ &\int \frac{1}{G(E)} \exp\left(N s^{\mathrm{c}}\big|_{\epsilon} + \frac{\mathrm{d}s^{\mathrm{c}}}{\mathrm{d}E^{\mathrm{c}}}\bigg|_{\epsilon} (E - N\epsilon)\right) W(E)\,\mathrm{d}E \,. \end{aligned} \tag{11.38}$$

To simplify this even further we define

$$\beta = \frac{\mathrm{d}s^{\mathrm{c}}}{\mathrm{d}E^{\mathrm{c}}}\bigg|_{\epsilon} \tag{11.39}$$

and rewrite (11.38) as

$$W^{\mathrm{d}}(E^{\mathrm{g}}) \propto G^{\mathrm{g}}(E^{\mathrm{g}}) \mathrm{e}^{-\beta E^{\mathrm{g}}} \,, \tag{11.40}$$

which is exactly the well known Boltzmann distribution.

11.5 Beyond the Boltzmann Distribution?

In the last years the standard limits of thermodynamics have been challenged by exploiting the laws of quantum mechanics [4, 5, 117]. It should be pointed out here that within the framework 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 Jaynes' principle. Rather, it is due to the special structure of the spectra of the systems that represent the environment. If a system is in contact with a system, which is not built according to the scheme described in Sect. 11.2, it can have a stable equilibrium energy probability distribution that significantly differs from the Boltzmann distribution. In fact, any distribution described by (9.42), must be considered stable, as long as the state density of the container system is large enough. Thus, if one could build a system with a high state density, but not of modular origin, one could get a non-standard equilibrium distribution.

However, realizing such a system is probably very hard; it would either have to be impossible to split up into identical parts, or, alternatively, 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 rather unrealistic, such effects might be seen in some future experiments.

12 Temperature

> *All concepts . . . have a definite and limited applicability . . . Such a case is that of temperature, defined as the mean kinetical energy of the random linear motion of the component particles of a many-particle system in thermal equilibrium. This notion is difficult to apply if there are too few particles in the system, or if the temperature is so low that thermal equilibrium takes a long time to establish itself, or if the temperature is so high that the nature of particles changes with small changes of the temperature.*
>
> — Th. Brody [23]

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 do so for the temperature. One could claim that temperature should only be defined for equilibrium and thus there is no need to define it as a function of the micro state. Based on this reasoning temperature would then simply be defined as

$$\frac{1}{k_{\mathrm{B}}T} = \frac{\partial S}{\partial E} = \frac{\partial}{\partial E}\ln G(E) = \frac{1}{G(E)}\,\frac{\partial G(E)}{\partial E}\ , \tag{12.1}$$

with $G(E)$ being the state density cf. (3.48). In this way one would neglect all dynamical aspects (see [131]), since this definition is based on the Hamiltonian of the system rather than on its state. Strictly speaking, this definition would exclude all situations in which temperature appears as a function of time or space, because those are non-equilibrium situations. To circumvent this restriction it would, at least, be convenient to be able to express temperature as a function of the micro state. There have been several attempts in this direction.

As already explained in Chap. 5, 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 $k_{\mathrm{B}}T$ in this case. Thus, if ergodicity is assumed, i.e., if the time average equals the ensemble average, temperature may indeed be defined as the time averaged kinetic energy. Similar approaches have been proposed on the basis of the microcanonical ensemble [107, 110]. However, temperature is eventually not really given by an observable (cf. Sect. 18.6),

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 123–131 (2004)
`http://www.springerlink.com/`

but by a time average over an observable, leaving open the question of the averaging time and thus the question on what minimum timescale temperature may be defined. Furthermore, the definition is entirely based on ergodicity. Nevertheless, it allows, at least to some extent, for an investigation of processes, in which temperature varies in time and/or space, since that definition is not necessarily restricted to full equilibrium.

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

12.1 Definition of Spectral Temperature

We define the inverse of spectral temperature as

$$\frac{1}{k_{\mathrm{B}}T} := -\left(1 - \frac{W_0 + W_M}{2}\right)^{-1} \sum_{i=1}^{M} \left(\frac{W_i + W_{i-1}}{2}\right) \frac{\ln W_i - \ln W_{i-1} - (\ln N_i - \ln N_{i-1})}{E_i - E_{i-1}} \,, \tag{12.2}$$

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

$$\frac{W_1 N_0}{W_0 N_1} = \exp\left(-\frac{E_1 - E_0}{k_{\mathrm{B}}T}\right) . \tag{12.3}$$

The definition (12.2) results if one groups the energy levels of a multi-level system into neighboring pairs, to each of which a “temperature” is assigned via the above formula, weighted by the average probability 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 should be many systems or situations with such a temperature, which do not exhibit thermodynamic properties at all. The latter will, as explained in the following, only show up in equilibrium situations or close to those.

If the spectrum of a system is very dense and if it is possible to describe the energy probability distribution, $\{W_i\}$, as well as the degrees of degeneracy, $\{N_i\}$, by smooth continuous functions $(W(E), N(E))$ with a well defined derivative, (12.2) could be approximated by

$$\frac{1}{k_{\mathrm{B}}T} \approx -\int_0^{E_{\max}} W(E) \left(\frac{\mathrm{d}}{\mathrm{d}E} \ln W(E) - \frac{\mathrm{d}}{\mathrm{d}E} \ln N(E)\right) \mathrm{d}E \, . \tag{12.4}$$

This can further be simplified by integrating the first term to yield

$$\frac{1}{k_{\mathrm{B}}T} \approx W(0) - W(E_{\mathrm{max}}) + \int_0^{E_{\mathrm{max}}} W(E) \frac{\mathrm{d}}{\mathrm{d}E} \ln N(E)\, \mathrm{d}E$$

$$\approx W(0) - W(E_{\mathrm{max}}) + \int_0^{E_{\mathrm{max}}} \frac{W(E)}{N(E)} \frac{\mathrm{d}N(E)}{\mathrm{d}E}\, \mathrm{d}E\ . \tag{12.5}$$

Since for larger systems typically neither the lowest nor the highest energy level is occupied with considerable probability (if the spectra are finite at all), it is the last term on the right hand side of (12.5) that basically matters. This term can be interpreted as the average over the standard, system based, rather than micro state based definition of the inverse temperature.

12.2 The Equality of Spectral Temperatures in Equilibrium

The equality of temperatures in equilibrium is usually shown based on entropy being extensive, i.e., additive for two systems in contact, on entropy approaching a maximum in equilibrium, and on the standard definition of temperature as given by (12.1). If we were exclusively dealing with large modular systems as described in Chap. 11, we could also introduce the equality this way, exploiting the corresponding properties derived so far. In the following, however, it will be demonstrated that the concept of equal equilibrium temperatures holds for even more general situations if based on spectral temperatures.

If two systems are in heat contact at the total energy $E = E^{\mathrm{g}} + E^{\mathrm{c}}$, we expect their energy probability distributions to be those corresponding to the dominant region (see Sect. 9.2.2), i.e. to be defined by

$$W^{\mathrm{d}}(E^{\mathrm{g}}) = N^{\mathrm{g}}(E^{\mathrm{g}}) \sum_E \frac{N^{\mathrm{c}}(E - E^{\mathrm{g}}) W(E)}{N(E)}\ , \tag{12.6}$$

$$W^{\mathrm{d}}(E^{\mathrm{c}}) = N^{\mathrm{c}}(E^{\mathrm{c}}) \sum_E \frac{N^{\mathrm{g}}(E - E^{\mathrm{c}}) W(E)}{N(E)}\ . \tag{12.7}$$

We check now if and under what circumstances those dominant energy probability distributions yield the same temperature according to the definition (12.2) or (12.5).

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 Sect. 11.2. For such a joint system the factor

$$N^{\mathrm{c}}(E^{\mathrm{c}})\, N^{\mathrm{g}}(E - E^{\mathrm{c}}) \tag{12.8}$$

will always be peaked near $E^{\mathrm{c}} \approx E$, since N^{c} grows, by definition, much faster with energy than N^{g} does. Thus, calculating $W^{\mathrm{d}}(E^{\mathrm{c}})$ and assuming the situation in Sect. 11.4, namely that $W(E)/N(E)$ takes on considerable values in some finite energy region only, we find that $W^{\mathrm{d}}(E^{\mathrm{c}})$ will also take on considerable values within and slightly below the very same energy region only. This intuitive result means that most of the total energy is in the larger system.

Since by definition, the state density of the large system can be well described by an exponential, i.e.,

$$N^{\mathrm{c}}(E^{\mathrm{c}}) \propto \mathrm{e}^{\beta E^{\mathrm{c}}} \tag{12.9}$$

in the region where $W^{\mathrm{d}}(E^{\mathrm{c}})$ is peaked, we find, applying (12.5), for the inverse temperature of the large system

$$\frac{1}{k_{\mathrm{B}} T^{\mathrm{c}}} \approx \beta \; . \tag{12.10}$$

For the same situation we infer for the small system, as explained in Sect. 11.4 (11.40)

$$W^{\mathrm{d}}(E_i^{\mathrm{g}}) \approx \frac{N^{\mathrm{g}}(E_i^{\mathrm{g}})\,\mathrm{e}^{-\beta E_i^{\mathrm{g}}}}{\sum_j N^{\mathrm{g}}(E_j^{\mathrm{g}})\,\mathrm{e}^{-\beta E_j^{\mathrm{g}}}} \; . \tag{12.11}$$

Inserting this into (12.2) yields

$$\frac{1}{k_{\mathrm{B}} T^{\mathrm{g}}} \approx \beta \; , \tag{12.12}$$

so that

$$T^{\mathrm{c}} \approx T^{\mathrm{g}} \; . \tag{12.13}$$

For a large modular continuous system in contact with a small discrete system and a reasonably peaked energy probability distribution of the combined system, we thus find 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 environment or not. The temperatures are the same, although entropy is definitely not additive with respect to the chosen partition.

Now we 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 (12.6) for a continuous spectrum yields

$$W^{\mathrm{d}}(E^{\mathrm{g}}) = N^{\mathrm{g}}(E^{\mathrm{g}}) \int_0^{\infty} \frac{N^{\mathrm{c}}(E - E^{\mathrm{g}}) W(E)}{N(E)} \, \mathrm{d}E \; . \tag{12.14}$$

Applying (12.5) to this distribution yields

$$\frac{1}{k_{\mathrm{B}}T^{\mathrm{g}}} = W^{\mathrm{d}}(0) - W^{\mathrm{d}}(E_{\mathrm{max}}) + \int_0^{E_{\mathrm{max}}} \int_0^{\infty} \frac{N^{\mathrm{c}}(E-E^{\mathrm{g}})W(E)}{N(E)} \, \frac{\mathrm{d}N^{\mathrm{g}}(E^{\mathrm{g}})}{\mathrm{d}E^{\mathrm{g}}} \, \mathrm{d}E \, \mathrm{d}E^{\mathrm{g}} \, . \tag{12.15}$$

The smallest energy value for which $N^{\mathrm{c}}(E^{\mathrm{c}})$ takes on non-zero values at all, is $E^{\mathrm{c}} = 0$. Thus we can, after reversing the order of integrations, replace E_{max} as a boundary by E. Furthermore we assume both the probability densities to find the system in the ground state $W^{\mathrm{d}}(0)$ and at the highest possible energy (if there is one), $W^{\mathrm{d}}(E_{\mathrm{max}})$, to vanish. We can then rewrite (12.15) as

$$\frac{1}{k_{\mathrm{B}}T^{\mathrm{g}}} = \int_0^{\infty} \frac{W(E)}{N(E)} \int_0^{E} \frac{\mathrm{d}N^{\mathrm{g}}(E^{\mathrm{g}})}{\mathrm{d}E^{\mathrm{g}}} N^{\mathrm{c}}(E-E^{\mathrm{g}}) \, \mathrm{d}E^{\mathrm{g}} \, \mathrm{d}E \, , \tag{12.16}$$

and apply product integration to the inner integral, to find

$$\frac{1}{k_{\mathrm{B}}T^{\mathrm{g}}} = \int_0^{\infty} \frac{W(E)}{N(E)} \Big(N^{\mathrm{g}}(E)N^{\mathrm{c}}(0) - N^{\mathrm{g}}(0)N^{\mathrm{c}}(E) - \int_0^{E} N^{\mathrm{g}}(E^{\mathrm{g}}) \, \frac{\mathrm{d}N^{\mathrm{c}}(E-E^{\mathrm{g}})}{\mathrm{d}E^{\mathrm{g}}} \, \mathrm{d}E^{\mathrm{g}} \Big) \, \mathrm{d}E \, . \tag{12.17}$$

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

$$\frac{1}{k_{\mathrm{B}}T^{\mathrm{g}}} = \int_0^{\infty} \frac{W(E)}{N(E)} \int_0^{E} \frac{\mathrm{d}N^{\mathrm{c}}(E^{\mathrm{c}})}{\mathrm{d}E^{\mathrm{c}}} \, N^{\mathrm{g}}(E-E^{\mathrm{c}}) \, \mathrm{d}E^{\mathrm{c}} \, \mathrm{d}E \, . \tag{12.18}$$

One would have obtained exactly this result, if one had applied (12.5) to the container system. This may be seen from a comparison with (12.16), obviously only the subsystem indices are reversed.

If two large systems with continuous spectra are in heat contact, almost all micro states accessible to the full system yield the same local spectral temperatures for the subsystems, regardless of whether the spectra are typical for modular systems, or how broad the energy probability distribution of the full system is.

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

As already explained, we do not only expect the 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 Gibbsian fundamental form states. Evidently, there are situations in which the temperature as defined by (12.2) will not show this behavior. If, e.g., one considered an isolated system controlled by a time-dependent Hamiltonian, one would find 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. However, this is probably not the situation one has in mind, when trying to apply the Gibbsian fundamental form. Here we want to distinguish two processes, for which the first law should be applicable. Firstly, we investigate the process of transferring energy into an arbitrarily small system by bringing it into contact with, according to our definition, a hotter environment, and, secondly, the case of slowly depositing energy into a large system by any kind of procedure (running current through it, stirring it, etc.).

12.3.1 Contact with a Hotter System

In this case we consider a discrete system in equilibrium, the entropy of which is given by

$$S = -k_{\mathrm{B}} \sum_i W_i \ln W_i \quad \text{with} \quad W_i = \frac{W(E_i)}{N(E_i)} \,, \tag{12.19}$$

where W_i is now the probability of finding the system in one of the $N(E_i)$ energy eigenstates of the respective energy level E_i, not the probability of finding the system somewhere at the energy E_i. The internal energy of the system is now given by

$$U = \sum_i W_i E_i \,. \tag{12.20}$$

The energy probability distribution of the system in contact with a larger system reads, according to (11.40)

$$W_i = \frac{\exp\Big(-\frac{E_i}{k_{\mathrm{B}}T}\Big)}{\sum_j \exp\Big(-\frac{E_j}{k_{\mathrm{B}}T}\Big)} \,, \tag{12.21}$$

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

$$\frac{\partial U}{\partial S} = \frac{\frac{\partial U}{\partial T}}{\frac{\partial S}{\partial T}} \,. \tag{12.22}$$

For the derivative in the numerator we get

$$\frac{\partial U}{\partial T} = \sum_i \frac{\partial W_i}{\partial T} E_i \; . \tag{12.23}$$

Computing the derivate in the denominator yields

$$\frac{\partial S}{\partial T} = -k_{\mathrm{B}} \sum_i \left(\frac{\partial W_i}{\partial T} \ln W_i + \frac{\partial W_i}{\partial T} \right) \; . \tag{12.24}$$

Because the order for the summation and the derivative can be exchanged on the right hand side of (12.24) and as $\sum_i W_i = 1$, the last term vanishes. Together with (12.21) we thus get

$$\frac{\partial S}{\partial T} = -k_{\mathrm{B}} \sum_i \frac{\partial W_i}{\partial T} \left(-\frac{E_i}{k_{\mathrm{B}} T} - \ln \sum_j \exp\left(-\frac{E_j}{k_{\mathrm{B}} T} \right) \right) \; . \tag{12.25}$$

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

$$\frac{\partial S}{\partial T} = \frac{1}{T} \sum_i \frac{\partial W_i}{\partial T} E_i \; . \tag{12.26}$$

Inserting (12.23) and (12.26) into (12.22) eventually yields

$$\frac{\partial U}{\partial S} = T \; , \tag{12.27}$$

which means that for this kind of process our temperature exhibits the desired behavior.

12.3.2 Energy Deposition

Now we consider a large system in isolating contact with an environment, into which energy is deposited by any kind of process. The internal energy of such a system reads:

$$U = \int W(E) E \, \mathrm{d}E \; , \tag{12.28}$$

where $W(E)$ is again the probability of finding the system at some energy, not in a single energy eigenstate. The entropy of such a system in microcanonical equilibrium is with (9.25)

$$S = -k_{\mathrm{B}} \int W(E) \, \ln \frac{W(E)}{N(E)} \, \mathrm{d}E \; . \tag{12.29}$$

According to Sect. 11.2 we can assume that the width of the energy distribution of the system is small enough so that the state density $N(E)$ is well described by some exponential within the region, where $W(E)$ takes on substantial values. As has already been explained, this region can be fairly broad, if the system is large. In this case we can replace

$$N(E) \approx N(U)\,\mathrm{e}^{\beta(E-U)} \,. \tag{12.30}$$

Doing so we find

$$S \approx -k_{\mathrm{B}} \int W(E)\Big(\ln W(E) - \ln N(U) + \beta(E-U)\Big)\,\mathrm{d}E \tag{12.31}$$

and after integrating the last two terms

$$S \approx k_{\mathrm{B}} \ln N(U) - k_{\mathrm{B}} \int W(E) \ln W(E)\,\mathrm{d}E \,. \tag{12.32}$$

As an instructive example we consider the case of energy probability $W(E)$ being uniformly distributed over an interval of length ϵ. In this case we find from (12.32)

$$S \approx k_{\mathrm{B}} \ln N(U) + k_{\mathrm{B}} \ln \epsilon \,. \tag{12.33}$$

The change of entropy δS that arises in such a situation from a change of the mean energy, δU, and a change of the width of the distribution by a factor C is

$$\delta S \approx k_{\mathrm{B}} \frac{\partial}{\partial U} \ln N(U)\,\delta U + k_{\mathrm{B}} C \,. \tag{12.34}$$

To get an idea for the orders of magnitude involved we set

$$\frac{\partial}{\partial U} \ln N(U)\,\delta U =: \frac{1}{k_{\mathrm{B}} T_{\mathrm{emp}}} \,, \tag{12.35}$$

where T_{emp} is the empirical temperature as defined in (12.1), yielding

$$\delta S \approx \frac{\delta U}{T_{\mathrm{emp}}} + k_{\mathrm{B}} C \,. \tag{12.36}$$

This may become more elucidating by plugging in numbers and dimensions

$$\delta S \approx \frac{\delta U[\mathrm{J}]}{T_{\mathrm{emp}}[\mathrm{K}]} + 1.38 \times 10^{-23}[\mathrm{J/K}] C \,. \tag{12.37}$$

From this equation it is obvious that for state changes involving macroscopic energy changes δU at reasonable temperatures the second term, corresponding to the change of the width of the energy probability distribution, becomes negligible, unless the width is increased by a factor of $C > 10^{15}$ or so. Such a change of the width, however, seems implausible from what we know about, say, mechanical energy depositing processes, even if they do not proceed as described by adiabatical following (see Chap. 13). A very similar picture will result for non-uniform energy probability distributions. Thus it is safe to drop the second term on the right hand side of (12.32), so that

$$S \approx k_{\mathrm{B}} \ln N(U) \,. \tag{12.38}$$

Thus we are eventually able to calculate the entropy change per energy change for typical processes:

$$\frac{\partial S}{\partial U} = k_{\mathrm{B}} \frac{\partial}{\partial U} \ln N(U) \,. \tag{12.39}$$

This result has now to be compared with the spectral temperature for this situation. With the definition of the inverse spectral temperature (12.5) we obtain

$$\frac{1}{T} = k_{\mathrm{B}} \int W(E) \frac{\mathrm{d}}{\mathrm{d}E} \ln N(E) \, \mathrm{d}E \tag{12.40}$$

or, consistently assuming the same situation as above (exponential growth of state density) and approximating the logarithm of the state density around the internal energy U,

$$\begin{aligned} \frac{1}{T} = k_{\mathrm{B}} \int W(E) \frac{\mathrm{d}}{\mathrm{d}E} \Bigg(& \ln N(E)|_U \\ & + \frac{\partial}{\partial E} \ln N(E)\bigg|_U (E-U) + O(E^2) \Bigg) \mathrm{d}E \,. \end{aligned} \tag{12.41}$$

The first term is constant and therefore the derivative vanishes, leading us to

$$\frac{1}{T} = k_{\mathrm{B}} \, \frac{\partial}{\partial E} \ln N(E)\bigg|_U \int W(E) \, \frac{\mathrm{d}}{\mathrm{d}E} (E-U) \, \mathrm{d}E \,. \tag{12.42}$$

After integration we find

$$\frac{1}{T} = k_{\mathrm{B}} \, \frac{\partial}{\partial E} \ln N(E)\bigg|_U = k_{\mathrm{B}} \frac{\partial}{\partial U} \ln N(U) \,, \tag{12.43}$$

which is evidently the same as the entropy change per energy change as given by (12.39). Thus, we finally conclude that the temperature according to our definition features the properties needed to guarantee agreement with the Gibbsian fundamental form.

13 Pressure

> ... *the laws of macroscopic bodies are quite different from those of mechanics or electromagnetic theory. They do not afford a complete microscopic description of a system. They provide certain macroscopic observable quantities, such as pressure or temperature. These represent averages over microscopic properties.*
>
> — F. Mandl [80]

Technically one could introduce pressure within classical statistical mechanics as 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 interpreted as pressure. And indeed, there are simple models of ideal gases which can account for some of their properties in this way [97, 104]. 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 the internal energy U as a function of entropy S and volume V. The derivative of the internal energy with respect to volume, while keeping entropy constant, is then identified with negative pressure (cf. (3.18))

$$\left(\frac{\partial U}{\partial V}\right)_{S=\mathrm{const.}} := -p \,. \tag{13.1}$$

This amounts to identifying the pertinent force with the change of energy per change of length, which appears quite convincing, but the claim is that the change appears in such a way that entropy does not change. The internal energy of the system could, in principle, change in many ways but it is assumed that a process is selected that keeps entropy constant. Without this assumption the above definition (13.1) would be meaningless.

In this way pressure is defined by an infinitesimal step of an adiabatic process. It has to be examined if, and under what conditions, adiabatic processes occur at all. In the case of temperature it was rather 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.

13.1 On the Concept of Adiabatic Processes

At first sight, isentropic processes may appear almost trivial: If the influence of the environment on the system under consideration, g, would be described by means of a time-dependent change of some parameter $a(t)$ entering the

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 133–141 (2004)
`http://www.springerlink.com/`

Hamiltonian of the system g, i.e., if the environment could be reduced to a changing "effective potential", a classical control by $\hat{H}^g(a(t))$ would result. Irrespective of $a(t)$, the von Neumann entropy of g would necessarily remain constant.

However, in the context of the present theory, such a reduction is considered "unphysical". The environment, regardless of whether or not it gives rise to a changing Hamiltonian for the considered system, will always become entangled with the considered system, thus causing the local entropy of the latter to increase (see Sect. 9.1). To understand the combined effect of an "adiabatic process inducing" environment onto the system, we divide the continuous evolution into steps alternating between two different mechanisms: during one step-type the effect of the environment is modeled only by the changing parameter in the local Hamiltonian, $a(t)$, during the other only by the inevitable relaxation into the microcanonical equilibrium as described in Sect. 9.1. Letting the step duration go to zero should result in the true combined effect. Since the relaxation to microcanonical equilibrium makes the off-diagonal elements (in energy-representation) of the local density operator, $\hat{\rho}^{\mathrm{g}}$, vanish, the remaining entropy is controlled by the energy occupation probabilities. Thus, if those change during the "parameter changing steps", entropy changes inevitably as well under the full evolution. Therefore, adiabatic processes are not trivial at all in a true physical process. The invariance of entropy, however, can be guaranteed if the occupation numbers do not change during the parameter changing steps. (They will obviously not be changed during the "relaxation steps", for we assume microcanonical conditions.) In quantum mechanics such a behaviour can be found within the scheme of adiabatic following.

Under the conditions of adiabatic following not only S, but all occupation numbers of states remain constant. Similar to the classical picture, for adiabatic following to work, the speed of change must be low enough. This is shortly explained in the following.

The *adiabatic approximation* (see [31, 113] and for the classical version remember Sect. 4.5) is a method of solving the time dependent Schrödinger equation with a time-dependent Hamiltonian. If a Hamiltonian contains a parameter $a(t)$ like length or volume that varies in time, it will have the following form

$$\hat{H}(a(t)) = \sum_i E_i(a(t)) \, |i, a(t)\rangle\langle i, a(t)| = \sum_i E_i(t) \, |i, t\rangle\langle i, t| \,. \tag{13.2}$$

At each time t a momentary Hamiltonian with a momentary set of eigenvectors and eigenvalues is defined. If the wave function is expanded in terms of this momentary basis with an adequate phase factor, i.e., with the definition

$$\psi_i := \langle i, t|\psi\rangle \exp\left(\frac{1}{\mathrm{i}\hbar} \int_0^t E_i \mathrm{d}t'\right) \tag{13.3}$$

the time dependent Schrödinger equation can be transformed to the form

$$\frac{\partial \psi_j}{\partial t} = -\sum_i \psi_i \langle j,t|\Big(\frac{\partial}{\partial t}|i,t\rangle\Big) \exp\Big(\frac{1}{\mathrm{i}\hbar}\int_0^t (E_i - E_j)\mathrm{d}t'\Big) \; . \tag{13.4}$$

The bracket term on the right hand side of (13.4) scales with the velocity of the parameter change, $\mathrm{d}a(t)/\mathrm{d}t$, this term gets multiplied by a rotating phase factor that rotates the faster the larger the energy distance $E_i - E_j$. This means that if the initial state is a momentary eigenstate of the Hamiltonian $|\psi(0)\rangle = |i, 0\rangle$, the transition rate to other eigenstates will be extremely small if the velocity of the parameter change is low, and it will fall off like $(E_i - E_j)^{-1}$ for transitions to eigenstates that are energetically further away. Thus in this case of slow parameter change we have as an approximate solution

$$|\psi(t)\rangle \approx |i, t\rangle \; . \tag{13.5}$$

Obviously, for such an evolution entropy is conserved. This is what is called the *adiabatic approximation* or the *adiabatic following*. This behavior has been discussed in detail 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 wave function 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. However, if the edges of those peaks overlap, the corresponding amplitudes ψ_i have to be added coherently (according to (13.4)), thus producing, possibly, a higher probability density in the overlap region than the simple addition of the probability densities of the corresponding peaks would suggest. 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 tried to describe macroscopic thermodynamic systems as entirely isolated, non-interacting systems, one would have to admit that these are most likely in superpositions of 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:

$$\hat{H}(t) = \hat{H}^{\mathrm{g}}(a(t)) + \hat{H}^{\mathrm{c}} + \hat{I}^{\mathrm{gc}}(t) \; . \tag{13.6}$$

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. However, any change of energy is induced by the parameter $a(t)$, thus, if $a(t)$ stopped changing at some time, energy should no longer change either. Demanding this behavior we get as a condition for the interaction $\hat{I}^{\mathrm{gc}}(t)$

$$\left[\hat{H}^{\mathrm{g}}(t), \hat{I}^{\mathrm{gc}}(t)\right] = 0 \; . \tag{13.7}$$

As described in Sect. 9.1, the effect of a suitable coupled environment system is to reduce purity within the gas system down to the limit set by the conserved quantities derived from (13.7). This amounts to making the off-diagonal elements of $\hat{\rho}^{\,\mathrm{g}}$, represented in the basis of the momentary eigenvectors of $\hat{H}^{\mathrm{g}}(t)$, vanish. In order to get a qualitative understanding of the type of evolution that a Hamiltonian like the one defined in (13.6) will typically give rise to, we refer to the same scheme as introduced at the beginning of this section, i.e., we decompose the continuous evolution into two different types of (infinitesimal) time steps. In one type of step we imagine the interaction to be turned off and the system to develop according to its local Hamiltonian $\hat{H}^{\mathrm{g}}(t)$, this evolution being described by the corresponding von Neumann equation. During the other type of step, we imagine the interaction to be turned on, but constant in time, as well as the local Hamiltonian. During this period the evolution is described by the Schrödinger equation for the full system, and will result in quenching the momentary off-diagonal elements. These two types of steps are now supposed to interchange. In the limit of the steps becoming infinitesimally short, the true, continuous evolution results.

For the first type the von Neumann equation reads

$$\mathrm{i}\hbar \, \frac{\partial \hat{\rho}^{\,\mathrm{g}}}{\partial t} = \left[\hat{H}^{\mathrm{g}}(t), \hat{\rho}^{\,\mathrm{g}}\right] \; . \tag{13.8}$$

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

$$W_i = \langle i,t|\hat{\rho}^{\,\mathrm{g}}(t)|i,t\rangle \; . \tag{13.9}$$

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

$$\frac{\partial}{\partial t}W_i = \Big(\frac{\partial}{\partial t}\langle i,t|\Big)\hat{\rho}^{\,\mathrm{g}}(t)|i,t\rangle + \langle i,t|\frac{\partial \hat{\rho}^{\,\mathrm{g}}(t)}{\partial t}|i,t\rangle + \langle i,t|\hat{\rho}^{\,\mathrm{g}}(t)\Big(\frac{\partial}{\partial t}|i,t\rangle\Big) \; . \tag{13.10}$$

Splitting up $\hat{\rho}^{\,\mathrm{g}}$ into a diagonal part and an off-diagonal part E

$$\hat{\rho}^{\,\mathrm{g}} =: \sum_i W_i |i,t\rangle\langle i,t| + \mathsf{E} \tag{13.11}$$

and inserting (13.8) and (13.11) into (13.10) yields

$$\begin{aligned}\frac{\partial}{\partial t}W_i = & \, W_i \left(\Big(\frac{\partial}{\partial t}\langle i,t|\Big)|i,t\rangle + \langle i,t|\Big(\frac{\partial}{\partial t}|i,t\rangle\Big)\right) \\ & + \Big(\frac{\partial}{\partial t}\langle i,t|\Big)\mathsf{E}|i,t\rangle + \langle i,t|\mathsf{E}\Big(\frac{\partial}{\partial t}|i,t\rangle\Big) \\ & + \langle i,t|\big[\hat{H}^{\mathrm{g}},\mathsf{E}\big]|i,t\rangle \; . \end{aligned} \tag{13.12}$$

The first part on the right hand side of (13.12) vanishes since

$$\left(\Big(\frac{\partial}{\partial t}\langle i,t|\Big)|i,t\rangle + \langle i,t|\Big(\frac{\partial}{\partial t}|i,t\rangle\Big)\right) = \frac{\partial}{\partial t}\langle i,t|i,t\rangle = 0\ . \tag{13.13}$$

Thus (13.12) reduces to

$$\frac{\partial}{\partial t}W_i = \Big(\frac{\partial}{\partial t}\langle i,t|\Big)\mathsf{E}|i,t\rangle + \langle i,t|\mathsf{E}\Big(\frac{\partial}{\partial t}|i,t\rangle\Big) + \langle i,t|\big[\hat{H}^{\mathrm{g}},\mathsf{E}\big]|i,t\rangle\ . \tag{13.14}$$

Obviously, this derivative vanishes, if E vanishes. This means that if, during the intermediate step, in which the interaction is active, the off-diagonal elements were completely suppressed, the rate of change of the probability would vanish at the beginning of each step of the von Neumann equation type. It would take on non-zero values during this step, especially if the step were long and $\hat{\rho}^{\mathrm{g}}(t)$ changed quickly. If we made 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 $\hat{\rho}^{\mathrm{g}}(t)$ tends to make the adiabatic approximation fail, while the contact with the environment that quickly reduces the off-diagonal elements 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, which models the effect of the environment the way it was found in Sect. 9.1. Such an equation reads

$$\mathrm{i}\hbar\,\frac{\partial\hat{\rho}^{\mathrm{g}}}{\partial t} = \big[\hat{H}^{\mathrm{g}}(t),\hat{\rho}^{\mathrm{g}}\big] - \mathrm{i}\hbar\sum_{i,i'}\langle i,t|\hat{\rho}^{\mathrm{g}}|i',t\rangle\, C_{ii'}\,|i,t\rangle\langle i',t| \tag{13.15}$$

with

$$C_{ii'} = 0 \quad \text{for} \quad i = i'\ , \qquad C_{ii'} \geq 0 \quad \text{for} \quad i \neq i'\ . \tag{13.16}$$

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

$$\hat{\rho}^{\mathrm{g}}(t) = \sum_i W_i|i,t\rangle\langle i,t| + r(t)\mathsf{E}(t)\ , \quad \mathrm{Tr}\left\{\mathsf{E}^2\right\} = 1\ , \quad r \geq 0\ , \tag{13.17}$$

where the W_i's are time independent and E may also contain diagonal elements. The second term on the right hand side of (13.17) thus contains all deviations from the adiabatic behavior. With this definition r is a measure for those deviations. Taking the derivative of $\hat{\rho}^{\mathrm{g}}(t)$ according to (13.17) and observing (13.13) one finds

$$\begin{aligned}\frac{\partial}{\partial t}\hat{\rho}^{\mathrm{g}}(t) &= \sum_i W_i\left(\Big(\frac{\partial}{\partial t}|i,t\rangle\Big)\langle i,t| + |i,t\rangle\Big(\frac{\partial}{\partial t}\langle i,t|\Big)\right) + \frac{\partial r}{\partial t}\mathsf{E} + r\frac{\partial\mathsf{E}}{\partial t}\\ &= \frac{\partial r}{\partial t}\mathsf{E} + r\frac{\partial\mathsf{E}}{\partial t}\ .\end{aligned} \tag{13.18}$$

Inserted into (13.15) we get

$$\mathrm{i}\hbar\left(\frac{\partial r}{\partial t}\mathsf{E}+r\frac{\partial \mathsf{E}}{\partial t}\right)=\left[\hat{H}^{\mathrm{g}}(t),\hat{\rho}^{\mathrm{g}}\right]-\mathrm{i}\hbar\sum_{i,i'}\langle i,t|\hat{\rho}^{\mathrm{g}}|i',t\rangle C_{ii'}|i,t\rangle\langle i',t|\ . \tag{13.19}$$

Multiplying (13.19) from the left by E, taking the trace and realizing that

$$\mathrm{Tr}\left\{\mathsf{E}\frac{\partial}{\partial t}\mathsf{E}\right\}=\frac{1}{2}\frac{\partial}{\partial t}\mathrm{Tr}\left\{\mathsf{E}^2\right\}=0\ ,\quad \mathrm{Tr}\left\{\mathsf{E}\left[\hat{H}^{\mathrm{g}},r\mathsf{E}\right]\right\}=0 \tag{13.20}$$

one finds, solving finally for $\partial r/\partial t$

$$\begin{aligned}\frac{\partial r}{\partial t}=&-r\sum_{i,i'}\left|\langle i,t|\mathsf{E}|i',t\rangle\right|^2 C_{ii'}\\ &-\sum_i W_i\left(\langle i,t|\mathsf{E}\Big(\frac{\partial}{\partial t}|i,t\rangle\Big)+\Big(\frac{\partial}{\partial t}\langle i,t|\Big)\mathsf{E}|i,t\rangle\right)\ .\end{aligned} \tag{13.21}$$

If the right hand side of (13.21) consisted exclusively of the first sum, r could only decrease in time, the decrease would be the faster the bigger the $C_{ii'}$. Only the second sum can cause r to deviate from its stable value $r=0$ and this sum would be large if $\hat{H}^{\mathrm{g}}(t)$ changed quickly. Thus, a fast local decoherence should stabilize the adiabatic approximation even for rapidly changing Hamiltonians. A numerical simulation of these rather abstract concepts can be found in Sect. 18.8.

In effect, we argue here that realistic adiabatic processes cannot dispose of decohering environments: During the adiabatic expansion of the gas, say, the gas certainly remains in contact with the cylinder and the piston, the coupling to which has been identified as the source of thermal behavior. For an adiabatic process to happen we must assume that energy exchange is suppressed on the pertintent time scales. The stabilizing effect of such a microcanonical environment should not really be surprising.

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 of the local Hamiltonian. The evolution of the state will then have the following form

$$\hat{\rho}^{\mathrm{g}}(t)\approx\sum_i W_i|i(a(t))\rangle\langle i(a(t))|\ . \tag{13.22}$$

If we identify the parameter a by the volume V, we eventually find for the pressure

$$\frac{\partial U}{\partial V}=\frac{\partial}{\partial V}\mathrm{Tr}\left\{\hat{\rho}^{\mathrm{g}}\hat{H}^{\mathrm{g}}\right\}=\sum_i W_i\frac{\partial E_i}{\partial V}=-p\ . \tag{13.23}$$

In this way, pressure, or any other conjugate variable (except for temperature), is defined, whenever a local Hamiltonian $\hat{H}^{\mathrm{g}}$ can be specified such that the weak coupling limit applies (see Chap. 15) and the change of the system proceeds in such a way that, with a changing local Hamiltonian, the whole system remains within the weak coupling limit. If this is guaranteed, pressure can be defined by (13.23), regardless of whether the system is changing or not and regardless of whether the system is thermally insulated or not. The infinitesimal process step is just a virtual one, it is 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.

13.2 The Equality of Pressures in Equilibrium

To examine, which pressure will be reached in equilibrium by systems in "volume exchanging contact", we consider the often discussed case of a container with a moving wall, dividing the gas into two different compartments (see Fig. 13.1). This situation is just meant as an example (cf. Sect. 18.7), the principles of our consideration can be transferred to other situations.

To show for this example the equality of pressures in equilibrium one could try to repeat the calculation from Sect. 12.2, now with respect to volume rather than energy. However, such an approach would face two problems.

The first problem is intrinsic to the situation and arises in the case of a completely heat insulating wall. (A completely frictionless sliding of the wall is assumed anyway.) Imagine such a system to be prepared in equilibrium. If the wall is moved, equilibrium is disturbed. On the other hand, if the wall is moved slowly enough for the process to be adiabatic, entropy cannot increase. If the wall, after release, moved back towards its initial equilibrium position, it would not do so due to an increase of entropy, since, as explained, entropy will not be any larger in equilibrium. Thus, in such a situation equilibrium is not maintained for entropic reasons but due to kinetic inhibition. In fact, there may even be no relaxation to equilibrium at all on its way back, if the wall does not accelerate beyond a velocity for which the process becomes

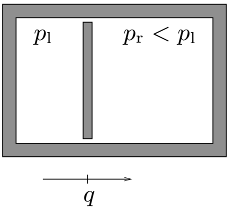


Fig. 13.1. Container with a moving wall inside, dividing the gas into two compartments.

irreversible, it would keep oscillating forever (in principle). Practically there will, of course, always be an equilibrium position of the wall. However, this may depend on details like the weight of the wall, non-equilibrium properties of the gases, etc.

In terms of our theory this difficulty results from the fact that the corresponding accessible region cannot be defined in the same simple way as discussed in Chap. 9. It is not possible to decide whether a state belongs to the accessible region or not by checking whether or not it is in accord with some constants of motion, like (9.5) and (9.25).

The second problem is of technical nature. It arises even in the case of the wall being heat conductive, a case for which the first problem does not occur.

In the case where the volume is fixed and the exchange of energy is allowed, it is possible to analyse the problem in terms of an orthogonal set of states, each of which corresponds to a given fragmentation of the full energy onto the subsystems (see Chap. 9). This is impossible in the case of volume being allowed for exchange. Unfortunately there is no set of "volume eigenstates" being orthogonal to each other. And thus the techniques from Sect. 12.2 cannot be applied.

In spite of these problems, 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 $q(t)$, 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

$$U^{\mathrm{l}}(t) + U^{\mathrm{r}}(t) + U^{\mathrm{w}}(t) = \text{const.}\,, \tag{13.24}$$

where the U's are the energy expectation values of the respective subsystems (internal energies). 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, so that for the two gas compartments the adiabatic approximation will be valid. The heavier the wall, the longer this period will last. For this period we thus get

$$U^{\mathrm{l}}(q(t)) + U^{\mathrm{r}}(-q(t)) + U^{\mathrm{w}}(t) = \text{const.}\,, \tag{13.25}$$

where q is the position of the wall and positive q values stand for the compression of the right compartment. Taking the derivative of (13.25) with respect to time yields

$$\left(\left(\frac{\partial U^{\mathrm{l}}}{\partial q}\right)_S - \left(\frac{\partial U^{\mathrm{r}}}{\partial q}\right)_S\right)\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\mathrm{d}U^{\mathrm{w}}}{\mathrm{d}t} = 0\,, \tag{13.26}$$

with the definition (13.1) we get

$$(p_l - p_r)\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\mathrm{d}U^{\mathrm{w}}}{\mathrm{d}t} \,. \tag{13.27}$$

If the pressure in the left compartment was higher than the pressure in the right compartment, the wall would move to the right while its energy rises. This process is energetically allowed, and thus likely to happen, no locally stationary situation will result. If the pressures in both compartments were the same, the wall would have to start moving without picking up internal energy. This is obviously energetically forbidden. If the wall starts moving, its internal 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.

In this way, the equality of pressures in equilibrium may be established, though not on the basis of some dominant region, like in the case of temperature.

14 Quantum Mechanical and Classical State Densities

> *Any real measurement involves some kind of coarse-grained average which will eventually obscure the quantum effects, and it is this average that obeys classical mechanics.*
>
> — L. E. Ballentine [9]

Regardless of its rather problematic foundation (see Chap. 4), Boltzmann's "recipe" to calculate thermodynamic behavior from a classical Hamilton function of a system works extremely well. This recipe essentially consists of his entropy definition, the first and second laws. Using this recipe, not only the thermodynamic behavior of gases, but also thermodynamic 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 particular approach, another (fully quantum mechanical) entropy definition is suggested, the question arises whether this other definition produces equally good if not better results. For many cases it suffices to check whether or not the classical and the quantum mechanical definitions of entropy are approximately equal

$$S^{\text{class}} \approx S^{\text{qm}} , \tag{14.1}$$

and together with the entropy definitions

$$S^{\text{class}} = k_{\text{B}} \ln G^{\text{class}}(U,V) , \qquad S^{\text{qm}} = k_{\text{B}} \ln G^{\text{qm}}(U,V) \tag{14.2}$$

it remains to investigate if

$$G^{\text{class}}(U,V) \approx G^{\text{qm}}(U,V) . \tag{14.3}$$

Here $G^{\text{class}}(U,V)$ is according to Boltzmann the number of classical micro states that is consistent with the macro state specified by U, V; stated more mathematically: the volume of the region in phase space that contains all micro states of the system that feature the energy U and are restricted to the (configuration space) volume V. This region is also referred to as the energy shell.

$G^{\text{qm}}(U,V)$ 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^{qm} is the equilibrium entropy we found for the case of microcanonical conditions and sharp energies (9.26). If the validity of (14.3) can not be established, a theory relying on S^{qm} would remain highly problematic from a practical point of view, regardless of its theoretical plausibility.

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 143–157 (2004)
http://www.springerlink.com/

From an operational point of view, the validity of (14.3) is far from obvious, because both quantities are evidently computed in entirely different ways. And of course, in general, $G^{\mathrm{qm}}(U,V)$ is discrete, while $G^{\mathrm{class}}(U,V)$ is a smooth continuous function. There are indeed cases where the recipe based on $G^{\mathrm{qm}}(U,V)$ works better than the one based on $G^{\mathrm{class}}(U,V)$. If one, e.g., changes from an ideal to a molecular gas, the deficiencies of the classical entropy definition, S^{class}, become visible at low temperatures; the heat capacity deviates significantly from the predicted behavior. This effect is referred to as the “freezing out of internal degrees of freedom”. It is due to the fact that the quantum mechanical level spacing of the spectrum arising from the internal (vibrational, rotational) degrees of freedom is much larger than that arising from the translational degrees of freedom. This behavior is described correctly by calculations based on S^{qm}. Nevertheless, if one claims S^{qm} to be the “correct” definition, the striking success of the classical entropy definition needs explanation. This can only be done by showing the validity of (14.3) 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 $G^{\mathrm{class}}(U,V)$ and $G^{\mathrm{qm}}(U,V)$: for a free particle they are the same. If the free particle is restricted to some volume, $G^{\mathrm{qm}}(U,V)$ becomes discrete, but as long as V is large, the level spacing is small and if the energy interval is chosen to contain many levels – it may still be extremely small compared to macroscopic energies – $G^{\mathrm{class}}(U,V)$ and $G^{\mathrm{qm}}(U,V)$ are still almost identical. This, eventually, is the reason why both methods lead to almost identical thermodynamic state functions for an ideal gas. A very similar situation is found for the harmonic oscillator. The quantum energy spectrum $G^{\mathrm{qm}}(U)$ of the harmonic oscillator consists of an infinite number of equidistant energy levels. The volume of the classical energy shell, $G^{\mathrm{class}}(U)$, of a harmonic oscillator is constant with respect to energy. Thus, if the level spacing is small, like 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.

In the following we want to analyse whether $G^{\mathrm{class}}(U,V)$ and $G^{\mathrm{qm}}(U,V)$ can be viewed as approximations for each other, at least for a large set of cases.

14.1 Bohr–Sommerfeld Quantization

One hint towards a possible solution in that direction comes from the Bohr–Sommerfeld Quantization [22, 127]. 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 jh

$$\oint p\,\mathrm{d}q = jh\,, \tag{14.4}$$

j being an integer. This integration over a region in phase space could be transformed into an integral over the classical state density with respect to energy

$$\int_0^{E_j} G^{\text{class}}(E)\,\mathrm{d}E = jh\ , \tag{14.5}$$

with E_j denoting the respective energy level. If this theory is right, the desired connection is established and the quantum mechanical spectrum can be calculated from $G^{\text{class}}(E)$ by (14.5), integrating only up to an energy level E_j.

A simple example for which the Bohr–Sommerfeld quantization produces good results is the harmonic oscillator in one dimension. Possible trajectories are ellipses in phase space (see Fig. 14.1). The classical phase space volume, the area of the ellipse enclosed by the trajectory, is according to (14.4)

$$\oint p\,\mathrm{d}q = \pi\sqrt{2mU}\sqrt{\frac{2U}{m\omega^2}} = \frac{U}{\nu}\ , \tag{14.6}$$

where $\nu = \omega/2\pi$ is the frequency of the oscillation. From standard quantum mechanics we know that $E_j = (j+\frac{1}{2})h\nu$. Applying (14.5) yields $U_j = jh\nu$ and is thus almost precisely correct.

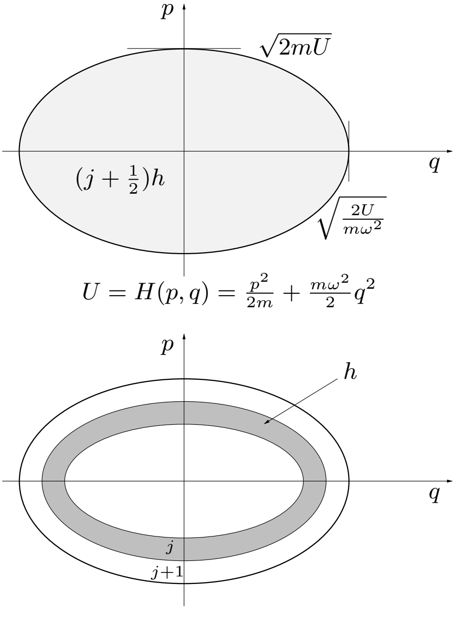


Fig. 14.1. Bohr–Sommerfeld-quantization: phase space of a one dimensional harmonic oscillator. The elliptic trajectory j includes a volume of $(j+\frac{1}{2})h$. Between two trajectories a volume of h is enclosed.

Unfortunately the Sommerfeld theory is not always applicable and the above formula holds true for some special cases only.

14.2 Partition Function Approach

Some more evidence for the similarity of G^{class} and G^{qm} can be obtained from a consideration which is usually done in the context of the partition function [108]. The partition function which, within standard classical mechanics, completely determines the thermodynamic properties of a system, reads for the quantum mechanical case

$$Z^{\mathrm{qm}} = \mathrm{Tr}\left\{\exp\left(-\frac{\hat{H}}{k_{\mathrm{B}}T}\right)\right\} = \int \mathrm{d}E\, G^{\mathrm{qm}}(E)\exp\left(-\frac{E}{k_{\mathrm{B}}T}\right) \tag{14.7}$$

and for the classical case

$$\begin{aligned} Z^{\mathrm{class}} &= \iint \exp\left(-\frac{H(\boldsymbol{q},\boldsymbol{p})}{k_{\mathrm{B}}T}\right)\prod_{\mu}\mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{p} \\ &= \int G^{\mathrm{class}}(E)\exp\left(-\frac{E}{k_{\mathrm{B}}T}\right)\,\mathrm{d}E\ . \end{aligned} \tag{14.8}$$

(If one sets $1/k_{\mathrm{B}}T = \alpha$, the partition function becomes equal to the function $R(\alpha)$, which is crucial for the spectrum of large modular systems as described in Sect. 11.2.) In the literature [108] one finds

$$\begin{aligned} &Z^{\mathrm{qm}}(\beta := \frac{1}{k_{\mathrm{B}}T}) \approx \\ &Z^{\mathrm{class}}(\beta) - \frac{\hbar\beta^3}{24m}\iint \exp(\beta H(\boldsymbol{q},\boldsymbol{p}))\sum_i\left(\frac{\partial H}{\partial q_i}\right)^2\prod_{\mu}\mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{p}\ , \end{aligned} \tag{14.9}$$

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 β, the gradient of the Hamiltonian and inverse mass $1/m$.

If $Z^{\mathrm{qm}}(\beta)$ and $Z^{\mathrm{class}}(\beta)$ were exactly the same, G^{class} and G^{qm} would have to be equal as well, since by taking derivatives with respect to β of the partition function, all moments of $\mathrm{e}^{-\beta E}G(E)$ can be produced and if all moments 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^{qm} is discrete while G^{class} is not. Thus, strictly speaking, the correction terms can never really vanish. Nevertheless (14.9) already provides a strong indication that for a large class of systems (the class of systems for which the correction term is small) at least the "rough structure" of $Z^{\mathrm{qm}}(\beta)$ and $Z^{\mathrm{class}}(\beta)$ could be the same.

Unfortunately, the smallness of the correction term does not provide a necessary criterion for the equality of G^{class} and G^{qm}. If one thinks, e.g., of a wide potential well with a saw tooth shaped bottom, it is obvious that if one makes each tooth smaller and smaller (but keeps the width the same by introducing more and more teeth), the spectrum should approach the spectrum of a flat bottom potential well for which the similarity of G^{class} and G^{qm} can be shown explicitly. The correction term for the saw tooth bottom potential well, however, does not decrease with the teeth getting smaller, it might indeed be arbitrarily big if the edges of the teeth are arbitrarily steep. Thus, there are systems for which the expansion in (14.9) does not even converge, even though G^{class} and G^{qm} of those systems may be very similar.

14.3 Minimum Uncertainty Wave Package Approach

In order to avoid the insufficiencies of the above analysis, we present here yet another treatment, which might help to clarify the relation between G^{class} and G^{qm}.

The basic idea 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 wave packages, thus each of them corresponds to a point in classical phase space. If it is then possible to show that only those wave packages contribute to $G^{\text{qm}}(U)$, which correspond to points in phase space that feature the classical energy U, i.e., if the energy spread of those packages is small, a connection between G^{class} and G^{qm} could be established.

Before we set up this complicated approximation scheme in full detail for arbitrary systems, we consider, again, as an instructive example, the one dimensional harmonic oscillator. In classical phase space the energy shells are ellipses as shown in Fig. 14.1. To each point (volume element) within this energy shell, a quantum mechanical state of minimum position-momentum uncertainty may be assigned. For the harmonic oscillator these states are known as "coherent" or "Glauber states" (see Fig. 14.2). These states are known to have an energy probability distribution centered around the energies of their classical counterparts. Furthermore the widths of these distributions decrease, relative to their mean energies, with increasing mean energies. Thus, each quantum state corresponding to a point within the energy shell may add the same "weight" within the same energy interval, to the quantum mechanical energy spectrum Fig. 14.3. In this case one will, eventually, find

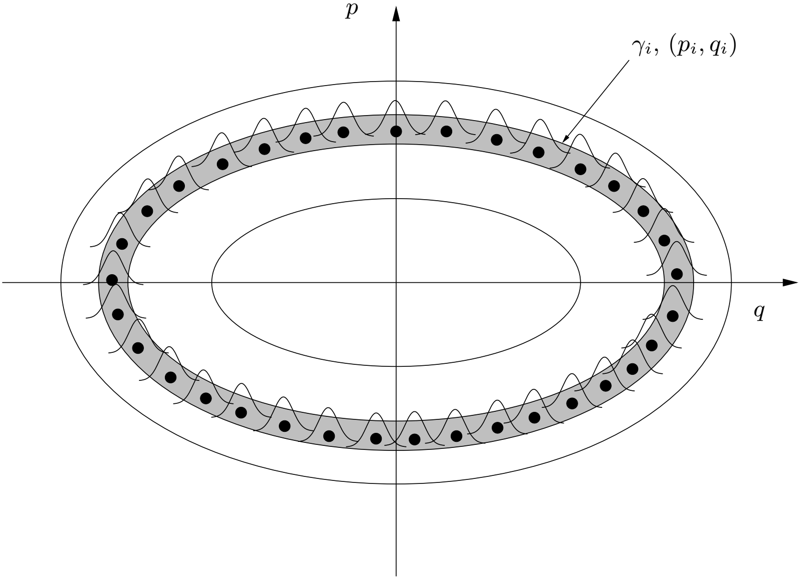


Fig. 14.2. Minimum uncertainty wave packages: phase space of the harmonic oscillator. Within the loop of volume h a minimum momentum-position uncertainty wave package γ_i is defined at every point (p_i, q_i).

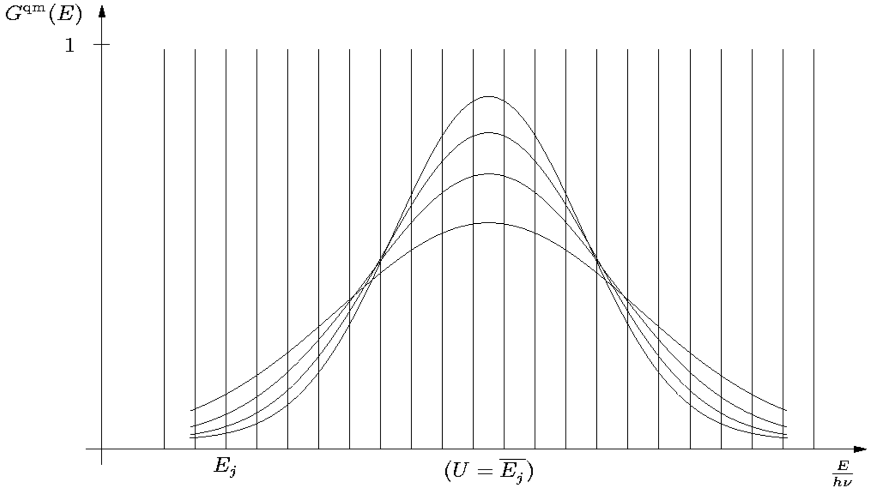


Fig. 14.3. Minimum uncertainty wave packages in energy space. Only wave packages within the phase space volume of h contribute to the state density $G^{\mathrm{qm}}(E)$ at energy E_j. All these packages add energy space to the respective level. Note that in the case of the harmonic oscillator all packages have additionally the same shape.

as many states in a certain energy interval in the quantum spectrum as there are points in the corresponding classical energy shell. This obviously establishes the similarity between G^{class} and G^{qm} that we are looking for. If, and to what extend such a scheme yields reasonable results in general, will be investigated in the following.

Now we try to apply these ideas to a more general Hamilton model. We start off by rewriting the quantum mechanical state density G^{qm}. Therefore we consider the respective Hamiltonian in energy basis which reads

$$\hat{H} = \sum_{E} E\hat{P}(E) \ , \tag{14.10}$$

where $\hat{P}(E)$ is the projector, projecting out the energy eigenspace with energy E. The quantum mechanical state density at energy E can then be written as the trace over the projector $\hat{P}(E)$

$$\mathrm{Tr}\left\{\hat{P}(E)\right\} = G^{\text{qm}}(E) \ . \tag{14.11}$$

Using a complete but not necessarily orthogonal basis $|\boldsymbol{\gamma}\rangle$, with

$$\sum_{\boldsymbol{\gamma}} |\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}| = \hat{1} \ , \qquad \langle\boldsymbol{\gamma}|\boldsymbol{\gamma}'\rangle \neq 0 \ , \tag{14.12}$$

we find for the quantum mechanical state density, carrying out the trace operation

$$G^{\text{qm}}(E) = \mathrm{Tr}\left\{\hat{P}(E)\right\} = \sum_{\boldsymbol{\gamma}} \langle\boldsymbol{\gamma}|\hat{P}(E)|\boldsymbol{\gamma}\rangle \ . \tag{14.13}$$

Furthermore, defining

$$g(\boldsymbol{\gamma}, E) := \langle\boldsymbol{\gamma}|\hat{P}(E)|\boldsymbol{\gamma}\rangle \tag{14.14}$$

we get

$$G^{\text{qm}}(E) = \sum_{\boldsymbol{\gamma}} g(\boldsymbol{\gamma}, E) \ . \tag{14.15}$$

According to its definition in (14.14) $g(\boldsymbol{\gamma}, E)$ is the energy spectrum of a single $|\boldsymbol{\gamma}\rangle$ that contributes to the spectrum of the Hamiltonian or, as mentioned before, to the energy spectrum of the $\hat{1}$-operator subject to this Hamiltonian (for a visualization see Fig. 14.4). The full spectrum $G^{\text{qm}}(E)$ of the Hamiltonian or the $\hat{1}$-operator is the sum of all individual spectra of those contributions $g(\boldsymbol{\gamma}, E)$, as stated in (14.15).

Now we choose the special complete basis as introduced in (14.12), which spans the whole Hilbert space. First of all we restrict ourselves to a two

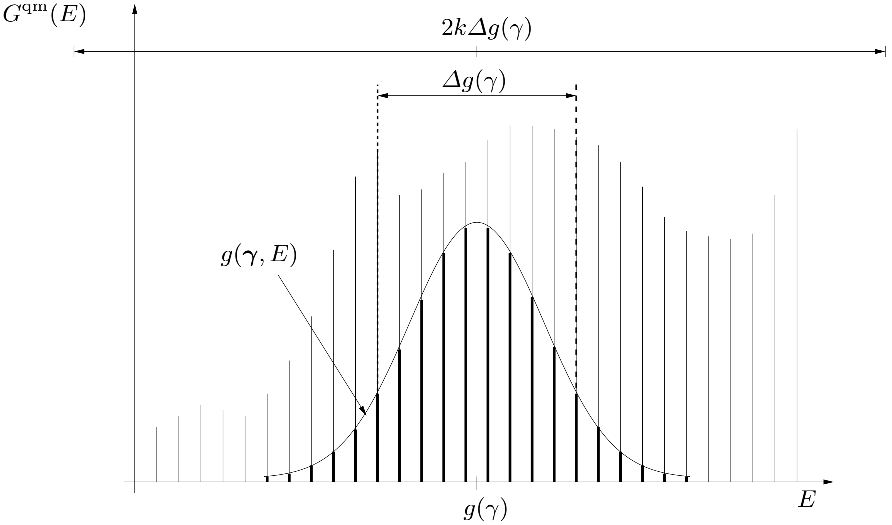


Fig. 14.4. Minimum uncertainty wave packages: contribution of a single $g(\boldsymbol{\gamma}, E)$ to the spectrum of the Hamiltonian, for an arbitrary quantum mechanical state density.

dimensional phase space (q, p), an extension to $6N$ phase space coordinates can easily be done later. In position representation this basis is defined as

$$\langle x|\boldsymbol{\gamma}\rangle := \langle x|q,p\rangle := \frac{1}{\sqrt[4]{2\pi}}\sqrt{\frac{\mathit{\Delta} q\mathit{\Delta} p}{\hbar\mathit{\Delta} x}}\exp\left(-\frac{(x-q)^2}{4(\mathit{\Delta} x)^2} - \mathrm{i}\frac{p}{\hbar}x\right) . \tag{14.16}$$

Obviously, this basis consists of Gaussian (minimum position-momentum uncertainty) wave packages with variance $\mathit{\Delta} x$, each of which corresponds to a point $\boldsymbol{\gamma} = (q, p)$ in phase space. The wave packages are defined on a lattice in phase space with distance $\mathit{\Delta} q$ and $\mathit{\Delta} p$, respectively (see Fig. 14.5), thus the coordinates q and p are integer multiples of $\mathit{\Delta} q$, $\mathit{\Delta} p$ only. In "standard" quantum mechanics one often tries to create a complete basis consisting of true quantum mechanical states by choosing the lattice grid such that $\mathit{\Delta} q\mathit{\Delta} p = h$. Thus one gets exactly one normalized state per "Planck cell" (see [89]). Note that the Gaussian wave packages defined in (14.16) are not normalized, if one does not choose this special subset. We are interested in the case $\mathit{\Delta} q \to 0$ and $\mathit{\Delta} p \to 0$, an "infinitesimal" Gaussian wave package basis. In this limit the norm of the basis states $\langle\boldsymbol{\gamma}|\boldsymbol{\gamma}\rangle$ will also vanish, but the basis remains complete. To prove this, we have to show that the following holds true

$$\begin{aligned}&\langle x|\sum_{\boldsymbol{\gamma}}|\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}|\,|x'\rangle \overset{!}{=} \langle x|\hat{1}|x'\rangle = \\ &\quad\sum_{\boldsymbol{\gamma}}\langle x|\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}|x'\rangle \overset{!}{=} \delta(x-x') .\end{aligned} \tag{14.17}$$

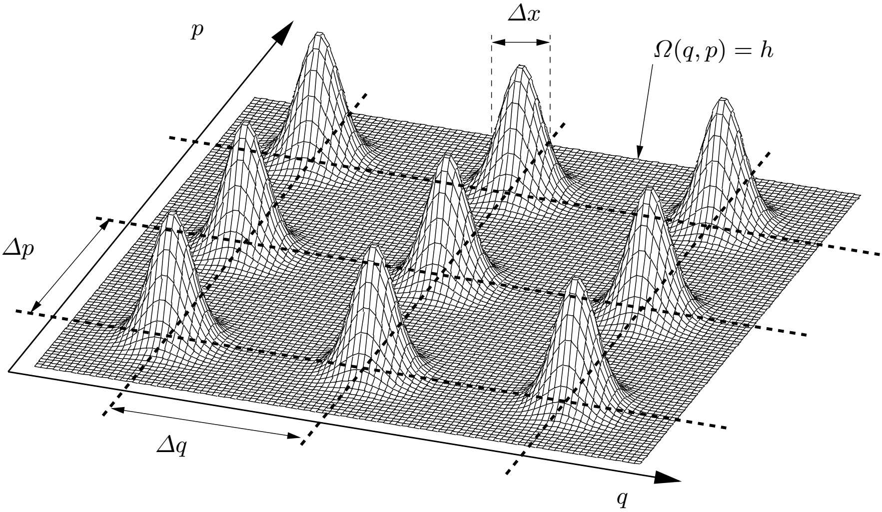


Fig. 14.5. Minimum uncertainty wave packages defined on a lattice in phase space with distance Δq and Δp, respectively.

Using the position representation of the basis states (14.16) the left hand side turns into

$$\begin{aligned}&\sum_{\gamma} \langle x|\gamma\rangle\langle\gamma|x'\rangle \\ &= \frac{1}{\sqrt{2\pi}\,\hbar\Delta x}\sum_{q,p}\exp\biggl(-\frac{(x-q)^2+(x'-q)^2}{4(\Delta x)^2}-\frac{\mathrm{i}}{\hbar}\,p\,(x-x')\biggr)\,\Delta q\Delta p\;.\end{aligned} \tag{14.18}$$

Within this equation, we perform the limit $\Delta q \to 0$, $\Delta p \to 0$ and switch from the sum to an integration

$$= \frac{1}{\sqrt{2\pi}\,\hbar\Delta x}\iint \mathrm{d}q\mathrm{d}p\,\exp\biggl(-\frac{(x-q)^2+(x'-q)^2}{4(\Delta x)^2}-\frac{\mathrm{i}}{\hbar}\,p\,(x-x')\biggr) \tag{14.19}$$

allowing for infinitesimal contributions. 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 $x = x'$, we can set $x = x'$ in the integrand and carry out the integration over the remaining Gaussian, finding

$$= \frac{1}{\sqrt{2\pi}\Delta x}\,\delta(x-x')\int \mathrm{d}q\,\exp\biggl(-\frac{(x-q)^2}{2(\Delta x)^2}\biggr) = \delta(x-x')\;, \tag{14.20}$$

which obviously shows that the chosen infinitesimal basis is complete. (This scheme is also known from the context of quantum optics, where it is usually called a "P-representation", see [118].)

The normalization of the basis states can be investigated by using again definition (14.16)

$$\langle\boldsymbol{\gamma}|\boldsymbol{\gamma}\rangle = \langle\boldsymbol{\gamma}| \int |x\rangle\langle x|\mathrm{d}x\,|\boldsymbol{\gamma}\rangle$$

$$= \frac{1}{\sqrt{2\pi}}\frac{\Delta q\Delta p}{\hbar\Delta x}\int \mathrm{d}x\exp\left(-\frac{(x-q)^2}{2(\Delta x)^2}\right) = \frac{\Delta q\Delta p}{2\pi\hbar} = \frac{\Delta q\Delta p}{h}\,. \quad (14.21)$$

From this equation it can be seen that the special complete basis consisting of normalized quantum mechanical states ($\langle\boldsymbol{\gamma}|\boldsymbol{\gamma}\rangle = 1$) can be introduced by choosing $\Delta q\Delta p = h$. This is exactly what von Neumann proposed. In the case of $\Delta q \to 0$ and $\Delta p \to 0$ the normalization vanishes. If one tried to sum the "weights" of the infinitesimal contributions in one "Planck cell", i.e., if one sums over a volume $\Omega(q,p) = h$ in phase space and does the limit later, one gets

$$\lim_{\Delta q\to 0}\lim_{\Delta p\to 0}\sum_{\substack{\boldsymbol{\gamma}\\ \Omega(q,p)=h}}\langle\boldsymbol{\gamma}|\boldsymbol{\gamma}\rangle = \lim_{\Delta q\to 0}\lim_{\Delta p\to 0}\sum_{q}\sum_{p}\frac{\Delta q\Delta p}{h}$$
$$\qquad\qquad\qquad\qquad\quad {\scriptstyle\Omega(q,p)=h}$$

$$= \frac{1}{h}\iint\limits_{\Omega(q,p)=h}\mathrm{d}q\,\mathrm{d}p = 1\,. \quad (14.22)$$

This means that all contributions coming from $\boldsymbol{\gamma}$-wave packages corresponding to a phase space volume h ("Planck cell") will add up to yield the weight one together. (For a more detailed discussion of the basis characteristics of Gaussian wave packages see [63, 115].)

Now we jump back to the considerations about the spectrum of the Hamiltonian and the $\hat{1}$-operator, respectively, to analyze the "weight" of a single $|\boldsymbol{\gamma}\rangle$ wave package to the full spectrum. As mentioned before the contribution of a single wave package corresponds to $g(\boldsymbol{\gamma},E)$. Since such a wave package might contribute to many energy levels we have to sum over all energies

$$\sum_E g(\boldsymbol{\gamma},E) = \sum_E\langle\boldsymbol{\gamma}|\hat{P}(E)|\boldsymbol{\gamma}\rangle = \langle\boldsymbol{\gamma}|\boldsymbol{\gamma}\rangle \quad \text{since} \quad \sum_E \hat{P}(E) = \hat{1}\,, \quad (14.23)$$

and therefore, together with (14.21)

$$\sum_E g(\boldsymbol{\gamma},E) = \frac{\Delta q\Delta p}{h}\,. \quad (14.24)$$

This means that while the spectrum is built up from contributions coming from $\boldsymbol{\gamma}$-wave packages, the contributions from wave packages corresponding to a phase space volume h will add exactly the weight 1 to the spectrum (see (14.22)). If this weight were now concentrated in one single energy level, one could say that the Sommerfeld quantization was exactly right and that a volume h would correspond to exactly one energy eigenstate. This, however, is typically not the case. Nevertheless, (14.24) means that, if the contributions of wave packages corresponding to classical states up to a certain energy fall entirely into some energy range of the quantum mechanical spectrum, and no other wave packages 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. This situation might very well be satisfied to good accuracy, 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.

It is straightforward to generalize the above introduced scheme to N particles, or rather $3N$ degrees of freedom. $|\gamma\rangle$ then has to be chosen as a product of the above described wave packages so that the contributions corresponding to a phase space region of volume h^{3N} will add the weight one to the spectrum. (Note that at this point it is just a generalization to a larger number of degrees of freedom, no modularity is involved yet.)

To decide now, whether or not the approximation of the number of quantum mechanical states within the respective energy interval by the volume of the classical phase space divided by h is a reasonable approximation, we analyze (14.15) more thoroughly. Since the exact form of $g(\gamma, E)$ can only be calculated with huge effort, we cannot evaluate this equation exactly. However, some properties of $g(\gamma, E)$ can be estimated with pretty good precision. The mean value, e.g., reads

$$\overline{g(\gamma)} := \frac{\sum_E E\, g(\gamma, E)}{\sum_E g(\gamma, E)} = \frac{\langle\gamma|\hat{H}|\gamma\rangle}{\langle\gamma|\gamma\rangle} \approx \frac{p^2}{2m} + \frac{1}{2m}\left(\frac{\hbar}{\Delta x}\right)^2 + V(q) \;, \quad (14.25)$$

where V is the potential energy part of the Hamiltonian and the right hand side holds if Δx is small compared to the curvature of the potential. Also the energy spread, the variance of $g(\gamma, E)$, can roughly be estimated

$$\Delta g(\gamma) := \frac{\sqrt{\langle\gamma|\hat{H}^2|\gamma\rangle - \langle\gamma|\hat{H}|\gamma\rangle^2}}{\langle\gamma|\gamma\rangle} \approx \frac{p\hbar}{m\Delta x} + \left.\frac{\partial H}{\partial x}\right|_{q,p} \Delta x \;. \quad (14.26)$$

We thus find that the mean of $g(\gamma, E)$ will be quite close to the energy of the corresponding classical state $U = H(q, p)$, and that its width will be reasonably small, at least for systems with large mass and small potential gradients. Or, stated in other words, a minimum momentum-position uncertainty wave package has a spectrum concentrated around the energy of its corresponding classical state.

Let us 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 Gaussians or Poissonians (see Fig. 14.4). Summing over the relevant energy interval only, we can approximate (14.24) by

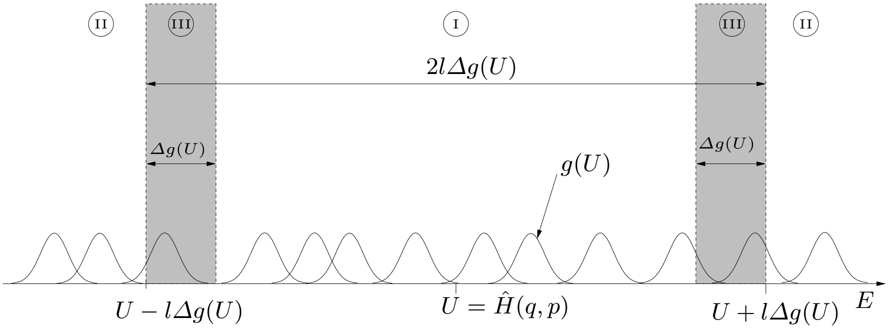


Fig. 14.6. Summation of wave packages. Large region I: all wave packages fall entirely into that part of the spectrum that is summed over. Region II: all packages are totally outside of the considered energy region. They do not contribute to the sum. Region III: problematic packages are at the edges.

$$\sum_{E=U-k\Delta g(\gamma)}^{E=U+k\Delta g(\gamma)} g(\gamma, E) \approx \frac{\Delta q \Delta p}{h} , \tag{14.27}$$

where k is some small integer and U the classical energy $H(q,p)$. Around the classical energy U we introduce a typical spread $\Delta g(U)$ of our wave packages in the energy region (e.g., an average over all spreads in the energy region). We can then approximately calculate the number of energy eigenstates within a small energy region around U. We sum over an energy interval $E = U - l\Delta g(U) \ldots U + l\Delta g(U)$, where l is an integer, so that we sum over a region of some typical spreads $\Delta g(U)$. According to the connection of $g(\gamma, E)$ and $G^{\mathrm{qm}}(U)$ in (14.15) it is possible to write

$$\sum_{E=U-l\Delta g(U)}^{E=U+l\Delta g(U)} G^{\mathrm{qm}}(E) = \sum_{E=U-l\Delta g(U)}^{E=U+l\Delta g(U)} \sum_{\gamma} g(\gamma, E) . \tag{14.28}$$

Now we want to rearrange this double sum, to use the approximation (14.27). Because of the dependencies on the different parameters, this is not so easy. In Fig. 14.6 we try to illustrate the situation. The idea is that, if l 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 to which (14.27) applies (area I in Fig. 14.6). No packages that are totally outside of the considered energy region contribute to the sum, therefore we can forget about them (area II in Fig. 14.6). Indeed there could be some packages at the edges of the considered region, for which we do not know their contribution to the sum, because they are inside as well as outside of the region (area III in Fig. 14.6). Within this interval of $2k\Delta g(\gamma)$ at the edges of the considered energy region, (14.27) can obviously not be applied, therefore we also neglect it. As a reasonable approximation for (14.28) one may accept to use

$$\sum_{E=U-l\Delta g(U)}^{E=U+l\Delta g(U)} \sum_{\gamma} g(\gamma, E) \approx \sum_{\substack{\gamma \\ U-l\Delta g(U)\leq H(\gamma)\leq U+l\Delta g(U)}} \frac{\Delta q \Delta p}{h} . \tag{14.29}$$

This is a sum, for which only those γ's contribute with $\Delta q \Delta p/h$ that are totally inside the considered energy interval around U. We need to estimate the error induced by the edges of the region. If the gradient of the classical Hamilton function in (14.26) does not vary much within the energy region (no extremely different spreads of the packages within and at the edges of the considered region), the relative error, $\tilde{\epsilon}$ can be estimated as the ratio of the energy interval at the edges, for which the $g(\gamma, E)$ do not fall entirely into the summed over part of the spectrum to the total summed energy interval

$$\tilde{\epsilon} \approx \frac{2\Delta g(U)}{2l\Delta g(U)} = \frac{1}{l} . \tag{14.30}$$

Keeping this in mind and with $\Delta q \to 0$ and $\Delta p \to 0$ in (14.29) we get

$$\begin{aligned} &\lim_{\Delta q \to 0} \lim_{\Delta p \to 0} \sum_{\substack{q,p \\ U-l\Delta g(U)\leq H(q,p)\leq U+l\Delta g(U)}} \frac{\Delta q \Delta p}{h} \\ &= \iint_{U-l\Delta g(U)\leq H(q,p)\leq U+l\Delta g(U)} \frac{\mathrm{d}q \mathrm{d}p}{h} , \end{aligned} \tag{14.31}$$

where the integral in phase space is restricted to an energy shell with the spread $2l\Delta g(U)$. In classical mechanics such an integral over an energy shell in phase space can be transformed into an integral of the classical state density $G^{\mathrm{class}}(E)$ over the energy itself. We conclude that

$$\sum_{E=U-l\Delta g(U)}^{E=U+l\Delta g(U)} G^{\mathrm{qm}}(E) \approx \frac{1}{h} \int_{U-l\Delta g(U)}^{U+l\Delta g(U)} \mathrm{d}E \, G^{\mathrm{class}}(E) , \tag{14.32}$$

with a relative error given by (14.30).

We have thus found that the number of states within a certain energy region is equal to the phase space volume corresponding to that energy region, divided by h. The energy region that has to be summed over is proportional to $l\Delta g(U)$. The relative error occurring is inversely proportional to l. Thus, this method is definitely not adequate to identify individual energy eigenstates with satisfactory precision. However, the number of states within an interval of, say 100 times the energy spread of a minimum uncertainty wave package, can be predicted with pretty good accuracy.

14.4 Implications of the Method for Minimum Uncertainty Wave Packages

The method introduced in the last section is useful for systems for which the energy spread (14.26) of a minimum uncertainty wave package is much less than its mean energy. This essentially holds in two cases (or combinations of both). One is the case of large systems (small potential gradients) with big particle masses. The other is the case of systems being made up from many identical, approximately interaction free subsystems, as will be explained below. In those cases the quantum mechanical state density in a certain finite energy interval can be approximated by the classical state density with a relative error $1/l$ (see (14.30)). Thus, for those classes of systems, *the classical and the quantum mechanical entropies are approximately the same* (see (14.10)-(14.12)). Since thermodynamic systems typically belong to the latter class, it will be very difficult to decide from thermodynamic experiments whether the classical or the quantum theory of thermodynamics holds true, since the predictions will be very similar in both theories.

Routinely, in thermodynamics we are dealing with systems consisting of a large number of identical, almost interaction free subsystems, for which the method introduced in Sect. 14.3 should yield good results. If they are assumed to be truly interaction free, the mean energy of a minimum uncertainty wave package of the full system is given by a sum over the mean values of the individual systems

$$\overline{g_{\text{tot}}(\gamma_{\text{tot}})} = \sum_{\mu=1}^{N} \overline{g_\mu(\gamma_\mu)} \,. \tag{14.33}$$

The energy spread of the full system in this case reads

$$\Delta g_{\text{tot}}(\gamma_{\text{tot}}) = \sqrt{\sum_{\mu=1}^{N} \left(\Delta g_\mu(\gamma_\mu)\right)^2} \,. \tag{14.34}$$

Because the whole system consists here of identical subsystems, we can assume that the individual mean value and the individual energy spread are of comparable size for most γ_{tot}. Therefore we can approximate the total mean value by $\overline{g_{\text{tot}}(\gamma_{\text{tot}})} = NE$ and the energy spread $\Delta g_{\text{tot}}(\gamma_{\text{tot}}) = \sqrt{N}\Delta E$. Now we can roughly estimate the ratio of energy spread over the mean value as

$$\frac{\Delta g_{\text{tot}}(\gamma_{\text{tot}})}{\overline{g_{\text{tot}}(\gamma_{\text{tot}})}} \approx \frac{1}{\sqrt{N}} \,, \tag{14.35}$$

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. Obviously this energy interval is very small for a classical number of subsystems (say some 10^{23} particles) and therefore classical and quantum mechanical considerations would not give rise to different predictions. For gases with only a small number of particles or near absolute zero temperature, where the above assumptions are definitely not true, a classical consideration would lead to totally different expectations than a quantum mechanical one. Remember, e.g., the freeze out of degrees of freedom in low temperature physics.

We conclude that there is a wide range of systems for which $G^{\mathrm{qm}}(U)$ is indeed proportional to $G^{\mathrm{class}}(U)$, in particular, large systems consisting of heavy mass particles and systems consisting of very many almost interaction free subsystems.

14.5 Correspondence Principle

It is quite comforting to see that the quantum approach to thermodynamics is compatible with classical procedures – in those cases where classical models are available at all. This fact can hardly be overestimated, as there are many branches of physics, in which thermodynamics is routinely applied while a quantum foundation would neither be feasible nor of practical interest. Pertinent examples would include, e.g., macroscopic physical and biophysical systems. The situation is different, e.g., for the concept of chaos: Here the exponential increase of distance between initially nearby states has no counterpart in quantum mechanics (where the distance is invariant under unitary transformation). This “failure of the correspondence principle”, though, is hardly a weakness of quantum theory; classical systems are essentially open systems, “the Schrödinger equation is not the appropriate tool for analyzing the route from quantum to classical chaos” ([59]). However, it is presently not clear, how conventional classical chaos might emerge from the underlying quantum substrate.

15 Sufficient Conditions for a Thermodynamic Behavior

> *It is not a question of annihilating science, but of controlling it. Science is totally dependent upon philosophical opinions for all of its goals and methods, though it easily forgets this.*
>
> — F. Nietzsche

Since the present approach is entirely based on the Schrödinger equation with no additional abstract concept like ergodicity or other "a priori postulate" being involved, it is obvious that the thermodynamic behavior, as explained in the previous chapters, should result in appropriate situations or physical set-ups only: not all systems that can be described by the Schrödinger equation are thermodynamic systems. Many of the restrictions and conditions that are necessary for thermodynamic behavior to emerge, have already been mentioned, but for clarity we want to summarize and discuss them here again explicitly. It should be pointed out that, unlike the conditions in many other approaches, the conditions introduced here are well defined physical properties of the systems under consideration, rather than abstract claims or postulates.

Furthermore, thermodynamic behavior is here considered to emerge "step by step". Contrary to other approaches, systems need not be fully thermodynamic or entirely non-thermodynamic. Since we have defined thermodynamics to encompass a multitude of properties (see Chap. 5), it could well be that systems show some aspects of thermodynamics only, while lacking others. We indeed find this to be the case.

15.1 Weak Coupling Limit

The prime prerequisite for a system to exhibit any thermodynamic 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

$$\sqrt{\langle \hat{I}^2 \rangle} \ll \langle \hat{H}^{\mathrm{g}} \rangle, \langle \hat{H}^{\mathrm{c}} \rangle \tag{15.1}$$

i.e., the energies contained in the system itself and in the surrounding one have to be much larger than the energy contained in the interaction. This has to hold for any state, into which the full system can possibly evolve. This means, in particular, that the spectrum of the joint system has to be close to the convolution of the individual spectra of the two subsystems, i.e., the

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 159–163 (2004)
http://www.springerlink.com/

interaction can be considered as a perturbation that has only a weak effect on the spectrum (cf. Sect. 7.1).

So, on the one hand, the interaction is indispensable in order for the system to evolve towards higher local entropy states, but on the other, it has to be very weak. What exactly the "optimum" interaction strength should be, 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 ((10.12), (10.13), etc.) apply even for arbitrarily small interaction strengths. Relaxation times might become extremely long, though, but eventually equilibrium will be reached. If the subsystems have a high state density, or, more precisely, a small but finite level spacing, 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. Thus, to some extent a stronger interaction might even produce "better" results, since it yields higher effective degrees of degeneracy, which enhances thermodynamic behavior (see Chap. 17). If the interaction becomes too strong, the system might still reach some local equilibrium state, but this state may not be predictable anymore from the properties of the decoupled parts of the system. Except for these overall strengths and principal structures of the interaction matrix, like non-energy exchanging (microcanonical) or energy exchanging (canonical) conditions, the details of the interaction matrix are irrelevant. All numerical simulations show, indeed, independence of such details (see, e.g., Fig. 18.5).

15.2 Microcanonical Equilibrium

For a system that cannot exchange energy with its surroundings, an equilibrium state of maximum entropy (a state that is diagonal in the local energy eigenbasis) (10.12) will be reached if

$$\sum_B \frac{\left(W(E_B^{\mathrm{c}})\right)^2}{N^{\mathrm{c}}(E_B^{\mathrm{c}})} \ll \sum_A \frac{\left(W(E_A^{\mathrm{g}})\right)^2}{N^{\mathrm{g}}(E_A^{\mathrm{g}})} , \tag{15.2}$$

where $W(E_A^{\mathrm{g}})$ ($W(E_B^{\mathrm{c}})$) are the occupation probabilities of the system (environment) and $N^{\mathrm{g}}(E_A^{\mathrm{g}})$ ($N^{\mathrm{c}}(E_B^{\mathrm{c}})$) the corresponding degrees of degeneracy (cf. (9.24)). This 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 must be expected under microcanonical conditions. Thus, a behavior as derived in Sect. 5.1 2a results.

15.3 Energy Exchange Equilibrium

If a system is in contact with its environment with energy exchange allowed, an equilibrium energy probability distribution, $W^{\mathrm{d}}(E_A^{\mathrm{g}})$ (see (9.42)), will be reached, if

$$N^{\mathrm{g}}(E_A^{\mathrm{g}})N^{\mathrm{c}}(E - E_A^{\mathrm{g}}) \ll \left(\sum_A N^{\mathrm{g}}(E_A^{\mathrm{g}})N^{\mathrm{c}}(E - E_A^{\mathrm{g}}) \right) , \tag{15.3}$$

which means that either the state density of the considered system or the state density of the surrounding has to be high (cf. (9.52)). A state of maximum entropy consistent with the equilibrium energy probability distribution (a state that is diagonal in the local energy eigenbasis), is only reached, however, if additionally (cf. (9.56))

$$\sum_B \frac{\left(W^{\mathrm{d}}(E_B^{\mathrm{c}})\right)^2}{N^{\mathrm{c}}(E_B^{\mathrm{c}})} \ll \sum_A \frac{\left(W^{\mathrm{d}}(E_A^{\mathrm{g}})\right)^2}{N^{\mathrm{g}}(E_A^{\mathrm{g}})} . \tag{15.4}$$

This is essentially the same condition as for the microcanonical equilibrium, except that equilibrium probabilities enter rather than the initial probabilities. A full equilibrium that only depends on the initial energy probability distribution of the joint system, $W(E)$, is thus only reached, if the surroundings are much larger.

15.4 Canonical Equilibrium

A full canonical equilibrium state with complete independence of the initial state of the considered system and 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 of the type

$$N^{\mathrm{c}}(E_B^{\mathrm{c}}) \propto \mathrm{e}^{\alpha E_B^{\mathrm{c}}} \tag{15.5}$$

in the region where the function $W(E)/N(E)$ takes on non-negligible values (see Sect. 11.4). This is very likely to be the case if the environment system consists of very many, identical, weakly interacting subsystems (see Sect. 11.2). The more subsystems there are in this modular environment, the larger will be the range in which the state density is well described by an exponential. Thus, in this case a behavior as stated in Sect. 5.1 2b results.

15.5 Spectral Temperature

In Sect. 12.1 we have defined a so-called spectral temperature, on the basis of the spectrum and the momentary energy probability distribution of the system. It has the following properties: two weakly interacting systems having reached the equilibrium distribution $W^{\mathrm{d}}(E_A^{\mathrm{g}})$ and $W^{\mathrm{d}}(E_B^{\mathrm{c}})$ have the same spectral temperature, if either one system has a state density typical for large modular systems, like described in Sect. 11.2 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 (see Sect. 12.2). This is the property of a temperature as demanded in Sect. 5.1 4b.

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

Furthermore, the spectral temperature has yet another feature. Firstly, consider a small system coupled to some large modular environment with a specific parameter α as stated in (15.5), in equilibrium. Now we couple it to some other large modular environment with an infinitesimally smaller specific parameter α ("infinitesimally hotter"). The system starts to decay into the new equilibrium state. The amount of energy change divided by the amount of entropy change in this process is then exactly given by the spectral temperature. Additionally, consider a large modular system, the energy distribution of which takes on non-negligible values only in a region where its state density is described by an exponential. 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. Again, the energy change divided by the entropy change in this process will be approximately the spectral temperature. The approximation will be better the higher the state density is in all regions (Sect. 12.3). These properties are, as postulated in Sect. 5.1 3b, properties of the temperature.

15.6 Parametric Pressure

If the local Hamiltonian $\hat{H}^{\mathrm{g}}(V)$ for the system g is taken to depend on a continuous parameter V, a parametric pressure (13.23) can be defined on the basis of this Hamiltonian and its momentary energy probability distribution. This parametric pressure has the following properties. Let the external parameter V be changed by a small (infinitesimal) amount only and let this change be performed slowly compared to the rate at which off-diagonal elements of the density matrix of the system would vanish due to the weak coupling of the system to a larger environment (see Sect. 15.2). Then the change of the internal energy divided by the change of the parameter V is the momentary parametric pressure of the system Sect. 13.1. This is the property of a conjugate variable like pressure as demanded in Sect. 5.1 3c.

Furthermore, consider two such systems in contact in such a way that the parameter V of the Hamiltonian in one system can only increase, if the respective parameter in the other one decreases. Let both systems be weakly and microcanonically coupled to some larger environment. Then the whole system can be at rest only if both parametric pressures are equal (see Sect. 13.2). This is the property of an intensive variable like pressure as demanded in Sect. 5.1 4b.

15.7 Extensitivity of Entropy

If a system in full canonical contact with a 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 Sect. 15.2 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 (see Sect. 11.2). Thus, in this case, entropy is an extensive quantity as claimed in Sect. 5.1 4a.

So, if all systems that one deals with are large in the sense that they consist of many identical subunits, 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.

19 Heat Conduction

> *During the International Congress on Mathematical Physics held in London in the year 2000, J. L. Lebowitz expressed his opinion that one of the great challenges to mathematical physics in the twenty-first century is the theory of heat conductivity and other transport phenomena in macroscopic bodies.*
>
> — R. Alicki and M. Fanes [3]

In Chap. 17 we have considered the route of a system into its global equilibrium state due to the coupling to an environment – the way from a non-equilibrium initial state into the stationary global equilibrium state. In fact, this relaxation process may even look like an exponential decay described by a rate equation, i.e., a statistical behavior for the considered system, while system and environment follow a Schrödinger equation. This is the key-feature of the Schrödinger dynamics of bipartite systems, by which the problem of time reversal invariance of microscopic equations and irreversibility can be overcome.

In this chapter we go another step further to non-equilibrium phenomena – the local equilibrium behavior. For such a scenario, the intensive quantities of the system are no longer equal, i.e., there is no global equilibrium established. However, in spite of this fact, in a very small part of the system one finds something like an equilibrium state only differing from subregion to subregion. (Such non-equilibrium phenomena have already been introduced in Sect. 3.2, where we discussed the ideas of linear irreversible thermodynamics.)

One of the most challenging scenarios for such a local equilibrium behavior is the heat conduction. Think, e.g., of some material bar coupled at both ends to heat baths of different temperature. In this sort of experiment, one typically finds a constant temperature gradient within the material (no global equilibrium), such that in every small segment of the chain a local equilibrium state is realized with a supposedly well-defined temperature.

Before we present some preliminary results for normal heat conduction in quantum mechanical systems, we summarize some standard theories of heat conduction.

19.1 Theories of Heat Conduction

In the region of non-equilibrium phenomena it is much harder to get a general approach to the different aspects of this subject than in the area of equilibrium. Therefore we concentrate here mainly on two aspects of this subject. Firstly, on the linear response on thermal forces and secondly on the quasi-particle approach to transport behavior.

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 221–234 (2004)
`http://www.springerlink.com/`

We restrict ourselves here to insulators. In fact, we are mainly interested in the "pure" energy current through a material, where no direct particle transport channel exists (like that implemented by electrons in metals). In the classical understanding of solid states, normal heat conduction in insulators results from the transport of quasi-particles like, e.g., phonons or magnons, etc. Therefore the transport of heat through a material would be nothing but a diffusion of those quasi-particles. However, why should the quasi-particles start to diffuse through the material? This must be due to some external force, which could induce a respective current within the material.

Let us start with the interpretation of heat conduction, mainly borrowed from electrodynamics – with the "theory of linear response".

19.1.1 Linear Response Theory

We have found in Sect. 3.2 for the energy current

$$\boldsymbol{j} = -\frac{\mathsf{L}}{T^2}\,\nabla T = -\kappa\,\nabla T\ , \tag{19.1}$$

where κ is the heat conductivity. This equation, called Fourier's law, defines a linear connection between current and temperature gradient.

To obtain an explicit expression for the conductivity a similar approach as applied in standard dc (direct current) electric conductivity is often used. In the latter case this approach is based on the idea that the electric potential is an external perturbation for the system under consideration. For a periodic time-dependent perturbation, this theory makes use of time-dependent perturbation theory of first order, in which one has to introduce a Dyson-series expansion of the time evolution, truncated after the first order term. Like in the case of Fermi's Golden Rule this perturbation leads in first order to transitions in the system. For a perturbation constant in time and space we need the zero frequency limit of this theory. If the electric field is taken as such a perturbation for the system, one is able to deduce the dc electric conductivity, given by the famous Kubo formula (for a complete derivation see [66, 67, 78]).

A similar approach for the thermal case has first been undertaken by Luttinger [67, 77], who introduced a "thermal potential" to the Hamiltonian of the system. However, this is exactly the main problem of this approach – what is a "thermal potential"? The mathematical analysis, while being completely similar to the case of electric conductivity, has to draw on ill-defined analogies: the perturbation should be proportional to the current. This assumption could not be proven mathematically and remains of unclear origin. Finally, one also gets a Kubo formula, now for thermal conductivity,

$$\kappa = -T \lim_{s\to 0}\int_0^\infty \mathrm{d}t\,\mathrm{e}^{-st}\int_0^\beta \mathrm{d}\beta'\,\langle \boldsymbol{j}(-t-\mathrm{i}\beta')\boldsymbol{j}\rangle\ . \tag{19.2}$$

Basically, this Kubo formula is a current-current correlation function, where $\boldsymbol{j}(-t - \mathrm{i}\beta')$ is a time-dependent current operator in the interaction picture.

Despite its problematic foundation the thermal Kubo formula is widely accepted. Kubo writes in his book "Statistical Physics II" (p. 183 in [67]):

> "It is generally accepted, however, that such formulas exist and are of the same form as those for responses to mechanical disturbances."

These formulas are being used in a large variety of cases to compute the thermal conductivity of experimental systems [51, 52].

19.1.2 Quasi-Particle Approach

As already mentioned, in the classical picture transport is interpreted to result from quasi-particles like, e.g., phonons or magnons diffusing through the solid. In this picture a stationary heat current results from scattering processes of the quasi-particles, in analogy to the particles proper experiencing collisions. Here, scattering processes derive from anharmonicity and the presence of defects. There are two different types of scattering: the "normal" or N-processes and so called U-processes ("Umklapp"-processes). N-processes are momentum conserving and cannot affect the heat conduction of the solid state. Only the second type of scattering, the U-processes, are believed to give rise to a finite heat conductivity. In such processes momentum is conserved only modulo a reciprocal lattice vector. The quasi-particles are scattered into different modes, possibly traveling in the reverse direction after the process. Usually such a diffusion of quasi-particles (quasi-particle transport) through the material is described by a Peierls-Boltzmann equation (see Sect. 4.1 and [82, 98, 99]). For very low temperatures these U-processes rapidly die out and the heat conductivity would diverge. In this case only impurities may limit the heat conductivity [81, 83, 100, 133]. Because of this suppression of the U-processes this approach is often discussed within a low temperature approximation.

Since we are going to study heat conductivity in spin chains later on, we focus now on magnons, the respective quasi-particles in the case of magnetic excitations (spin waves). Of course, such a quasi-particle concept would only be a good description for a big enough system. This is certainly not the case for the very small systems we are going to study later on, nevertheless let us turn to a short introduction into this theory.

In the following we consider a long chain of identical systems, taken to be spins. With the aid of a Holstein-Primakoff transformation (cf. [55]) and a Fourier transformation we can introduce magnon generation and annihilation operators $\hat{b}_k^\dagger$ and $\hat{b}_k$, fulfilling the correct bosonic commutator relations. In the low temperature limit one can approximate the Hamiltonian of the chain of spins with next neighbor Heisenberg interactions by the following diagonal Hamiltonian

$$\hat{H}_0 = \sum_k \omega_k \hat{b}_k^\dagger \hat{b}_k \,, \tag{19.3}$$

with a dispersion relation given by ω_k, where k defines the respective quasi-particle mode. The full derivation of this approximation can be found, e.g., in the book by Kittel, "Quantum Theory of Solids" [62]. An important fact is that all magnon scattering processes are removed now from the Hamiltonian and therefore we get a ballistic transport behavior, i.e., a magnon passes through the chain without any resistance. This is due to the fact that the current operator is a conserved quantity in this low temperature approximation of the Heisenberg coupled chain, as can be shown by transforming the current operator also into the mode picture and using a low temperature approximation. This current operator commutes with the Hamiltonian of (19.3) and therefore we find no normal heat transport in this low temperature approximation. A current in a cyclic structure, once induced, will never decrease in such a system, it will flow for ever.

However, if we do not neglect higher oder terms of the magnon Hamiltonian we additionally get a magnon-magnon interaction. The next higher order term of the Hamiltonian is given by

$$\hat{H}_1 \propto \sum_{k_1,k_2,k_3,k_4} \delta(k_1+k_2-k_3-k_4-2\pi j)\, W(k_1,k_2,k_3,k_4)\hat{b}_{k_1}^\dagger \hat{b}_{k_2}^\dagger \hat{b}_{k_3} \hat{b}_{k_4} \,, \tag{19.4}$$

where the $W(k_1, k_2, k_3, k_4)$ can be interpreted as transition probabilities for the respective scattering process. The delta-function $\delta(k_1+k_2-k_3-k_4-2\pi j)$ guarantees momentum conservation. Only those terms contribute, for which $k_1 + k_2 - k_3 - k_4 = 2\pi j$ with $j = 0, 1, \ldots$. This refers to the fact that the momentum is conserved only modulo by a reciprocal lattice vector. Since this is a fourth order term, the effect is very small. However, principally we expect a normal heat conductivity, because the quasi-particles are scattered now.

19.2 Quantum Heat Conduction

So far we have described several approaches to non-equilibrium behavior, in particular, to heat conduction. Here we want to present yet another method to analyze local equilibrium behavior. As discussed in the second part of this book, a quantum mechanical foundation can be established for equilibrium thermodynamics even of small systems and therefore it is tempting to expect non-equilibrium behavior also to follow directly from quantum mechanics. Indeed, in chain-like quantum systems strong indications have been found for a normal heat conduction [84, 111, 112]. However, the sort of conductivity in such systems depends on the coupling of the modular system. There are systems showing a normal heat conduction and others where the conductivity diverges. Therefore we attempt to give an introduction to the heat conduction in such chains, from the view of the theories developed in the text at hand.

19.2.1 Model Hamiltonian and Environment

The most important difference between the technique that we are going to explain in the following and the theories described before derives from the fact that we incorporate two heat baths at different temperatures explicitly into our model. This coupling eventually gives rise to an energy flux.

Of course, in such an approach we face the question of how to model the baths. In principle, we would like to include them as "full-fledged" quantum systems, featuring their own energy levels, etc. However, the numerical integration of the corresponding Schrödinger equation tends to become exponentially hard with a growing Hilbert space dimension. Especially for non-equilibrium systems, which we are going to investigate in this chapter, the dimension of the full Hilbert space is by far too large for a direct integration of the Schrödinger equation. We would have to introduce two environment systems ("baths") with some hundreds of energy levels each, coupled via our considered system with n^N levels. All together, the complete dimension of the Hilbert space was quickly over a million of energy levels. Thus we have to settle for an approximation scheme to investigate these systems with reduced computational efforts. The method proposed by Lindblad (introduced in Sect. 4.8) is exactly such a technique to approximate the influence of the environmental system, instead of integrating the full Schrödinger equation, as done so far. The numerical investigations of Sect. 18.4.3 on a chain coupled at one end to an environment, where we have compared the solution of the full Schrödinger equation with the approximation based on the Lindblad formalism showed a very good agreement between the two approaches (see Fig. 18.15). This justifies the use of the Lindblad formalism as a mathematical technique. Furthermore, the treatment of open systems by a master equation or a Lindblad formalism, is well-established.

The system we are going to investigate is a chain of N subsystems with $n = 2$ levels each (for more information about such "quantum networks" see [79] or [93]). The different subsystems are coupled by a next neighbor interaction of different types. All together, the Hamiltonian $\hat{H}$ of the system thus reads

$$\hat{H} = \sum_{\mu=1}^{N} \hat{H}_{\text{loc}}^{(\mu)} + \frac{\lambda}{\lambda_0} \sum_{\mu=1}^{N-1} \hat{H}_{\text{int}}^{(\mu,\mu+1)} \, . \tag{19.5}$$

Here, the terms of the first sum represent N equidistant 2-level subsystems. The second term accounts for pair-interactions between two adjacent subsystems, with an overall coupling strength λ. For λ to characterize the absolute interaction strength it is necessary to normalize the different interaction types by λ_0, the mean absolute value of interaction matrix elements, i.e.,

$$\lambda_0^2 = \frac{1}{n^{2N}} \operatorname{Tr} \left\{ \Big(\sum_{\mu=1}^{N-1} \hat{H}_{\text{int}}^{(\mu,\mu+1)} \Big)^2 \right\} \, . \tag{19.6}$$

In terms of the Pauli operators the local Hamiltonian of a subsystem μ with an energy spacing $\Delta E = 1$ can be written as

$$\hat{H}_{\text{loc}}^{(\mu)} = \frac{1}{2}\,\hat{\sigma}_3^{(\mu)} \,. \tag{19.7}$$

The two level systems are coupled via three alternative next neighbor interactions. Firstly, a non-resonant diagonal interaction

$$\hat{H}_{\text{NR}}^{(\mu,\mu+1)} = C_{\text{NR}}\,\hat{\sigma}_3^{(\mu)} \otimes \hat{\sigma}_3^{(\mu+1)} \,, \tag{19.8}$$

which does not account for energy transfer between the systems. Secondly, a resonant energy transfer interaction (Förster-Coupling), which reads in terms of Pauli operators,

$$\hat{H}_{\text{F}}^{(\mu,\mu+1)} = C_{\text{F}}\left(\hat{\sigma}_1^{(\mu)} \otimes \hat{\sigma}_1^{(\mu+1)} + \hat{\sigma}_2^{(\mu)} \otimes \hat{\sigma}_2^{(\mu+1)}\right) . \tag{19.9}$$

Here, C_{NR} and C_{F} can be used to adjust the relative strength of these two couplings. And last but not least a totally random next neighbor interaction

$$\hat{H}_{\text{R}}^{(\mu,\mu+1)} = \sum_{i=1}^{3}\sum_{j=1}^{3} p_{ij}\,\hat{\sigma}_i^{(\mu)} \otimes \hat{\sigma}_j^{(\mu+1)} \tag{19.10}$$

with normal distributed random numbers p_{ij}, and variance 1. Note that p_{ij} is independent of μ, since we do not allow for disorder. The coupling should be chosen randomly but the same between the different interacting pairs of subsystems. The random interaction is supposed to model "typical interactions" without any bias.

This system is now coupled to a suitable environmental system according to the Lindblad formalism. In doing so, we couple the first and the last subsystem of our chain to a separate bath, each at a different temperature. The standard Liouville equation is, therefore, supplemented by two incoherent damping terms, and reads

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -\mathrm{i}\big[\hat{H},\hat{\rho}\big] + \hat{\mathcal{L}}^{\mathrm{c}_1}\hat{\rho} + \hat{\mathcal{L}}^{\mathrm{c}_2}\hat{\rho} \,, \tag{19.11}$$

where the operators $\hat{\mathcal{L}}^{\mathrm{c}_1}$ as well as $\hat{\mathcal{L}}^{\mathrm{c}_2}$ are super-operators acting on operators of the Liouville space, so-called Lindblad operator, modelling the influence of the respective environment c_1 and c_2. In terms of super-operators we can write now the full Lindblad-equation as

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = \hat{\mathcal{L}}\,\hat{\rho} \,. \tag{19.12}$$

For the local coupling of spin ν of the chain to the environment c_1 we expand the Lindblad-part in terms of raising and lowering operators

$$\begin{aligned}\hat{\mathcal{L}}^{\mathrm{c}_1}\hat{\rho} =& \frac{W_{1\to 0}}{2}\,(2\,\hat{\sigma}_-^{(\nu)}\,\hat{\rho}\,\hat{\sigma}_+^{(\nu)} - \hat{\rho}\,\hat{\sigma}_+^{(\nu)}\hat{\sigma}_-^{(\nu)} - \hat{\sigma}_+^{(\nu)}\hat{\sigma}_-^{(\nu)}\,\hat{\rho}) + \\ & \frac{W_{0\to 1}}{2}\,(2\,\hat{\sigma}_+^{(\nu)}\,\hat{\rho}\,\hat{\sigma}_-^{(\nu)} - \hat{\rho}\,\hat{\sigma}_-^{(\nu)}\hat{\sigma}_+^{(\nu)} - \hat{\sigma}_-^{(\nu)}\hat{\sigma}_+^{(\nu)}\,\hat{\rho})\,,\end{aligned} \tag{19.13}$$

where the first term describes a decay from $|1\rangle \to |0\rangle$ with rate $W_{1\to 0}$ and the second from $|0\rangle \to |1\rangle$ with $W_{0\to 1} < W_{1\to 0}$. Now the properties of the environment (bath) only enter via these two rates. Their absolute amounts, refer to the coupling strength of the bath to the respective system. The relative differences of the rates are related to the temperature of the bath. If such a bath was brought into contact with a single two level system, the system would relax into the stationary state:

$$\hat{\rho}_{\mathrm{stat}} = \frac{1}{W_{1\to 0} + W_{0\to 1}} \begin{pmatrix} W_{1\to 0} & 0 \\ 0 & W_{0\to 1} \end{pmatrix} . \tag{19.14}$$

This Lindblad approximation is able to reduce the computational effort drastically. Instead of the very large Hilbert space we are now working in the Liouville space of the chain system with only n^{2N} dimensions.

According to Chap. 12 we use the probability of finding one of the single two level systems in the upper level as a measure for the local temperature of the respective system. For a small local coupling between the two level systems, i.e., for negligible energy contained in the interaction, we define as a measure for the local temperature

$$E_{\mathrm{loc}}^{(\mu)} = \mathrm{Tr}\left\{\hat{\rho}^{(\mu)}\hat{H}_{\mathrm{loc}}^{(\mu)}\right\} , \tag{19.15}$$

$\hat{H}_{\mathrm{loc}}^{(\mu)}$ and thus $E_{\mathrm{loc}}^{(\mu)}$ are defined in units of ΔE, $\hat{\rho}^{(\mu)}$ is the corresponding reduced density operator of site μ. Up to a logarithm this is exactly the spectral temperature defined in Chap. 12.

Before we can characterize the introduced system class of a chain coupled to two environmental systems, we have to define some more properties.

19.2.2 Current Operator and Fourier's Law

Once (19.11) has been solved, one wants to know whether or not the resulting behavior is in accord with standard linear irreversible thermodynamics, i.e., with Fourier's law

$$\boldsymbol{j} = -\kappa\,\nabla T\,. \tag{19.16}$$

This has to be read out somehow from the resulting density operator. Due to the absence of any notion of space, there is no gradient defined in our model. However, a reasonable quantity to replace the gradient seems to be the temperature difference between two adjacent spins

$$\Delta T^{(\mu,\mu+1)} = T^{(\mu)} - T^{(\mu+1)}\,. \tag{19.17}$$

Furthermore, we need an expression for the current within in the system. In order to define this consistently, we start by considering the derivative of local energy with respect to time. According to the Heisenberg equation of motion for operators (see (2.57)) one finds for the local energy in the chain

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}_{\mathrm{loc}}^{(\mu)} = \mathrm{i}\big[\hat{H}, \hat{H}_{\mathrm{loc}}^{(\mu)}\big] + \frac{\partial}{\partial t}\hat{H}_{\mathrm{loc}}^{(\mu)} , \tag{19.18}$$

where the last term vanishes because $\hat{H}_{\mathrm{loc}}^{(\mu)}$ has no explicit time dependence. Let us concentrate now on the Heisenberg coupled spin chain with $\hat{H}_{\mathrm{int}} = \hat{H}_{\mathrm{NR}} + \hat{H}_{\mathrm{F}}$, so that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\hat{H}_{\mathrm{loc}}^{(\mu)} &= \mathrm{i}\big[\hat{H}, \hat{H}_{\mathrm{loc}}^{(\mu)}\big] \\ &= \mathrm{i}\, C_{\mathrm{F}}\big(\hat{\sigma}_{-}^{(\mu-1)}\,\hat{\sigma}_{+}^{(\mu)} - \hat{\sigma}_{+}^{(\mu-1)}\,\hat{\sigma}_{-}^{(\mu)} + \hat{\sigma}_{+}^{(\mu)}\,\hat{\sigma}_{-}^{(\mu+1)} - \hat{\sigma}_{-}^{(\mu)}\,\hat{\sigma}_{+}^{(\mu+1)}\big) . \end{aligned} \tag{19.19}$$

Obviously, this quantity only depends on operators of the central system μ and its neighbors to the left and right. Note that the quantity does not change, if we consider only an energy exchange coupling (Förster coupling).

The four terms on the right hand side of (19.19) can be interpreted to describe a current into as well as a current out of the central system μ. Thus, we define these two current operators by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}_{\mathrm{loc}}^{(\mu)} = \hat{J}^{(\mu-1,\mu)} - \hat{J}^{(\mu,\mu+1)} , \tag{19.20}$$

i.e., the current operator between the sites μ and $\mu+1$ reads

$$\hat{J}^{(\mu,\mu+1)} = \mathrm{i}\, C_{\mathrm{F}}\big(\hat{\sigma}_{-}^{(\mu)}\,\hat{\sigma}_{+}^{(\mu+1)} - \hat{\sigma}_{+}^{(\mu)}\,\hat{\sigma}_{-}^{(\mu+1)}\big) . \tag{19.21}$$

In order to fulfill "Fourier's law", the current between two subsystems μ and $\mu+1$ should be connected to the temperature difference between these two systems by

$$\mathrm{Tr}\left\{\hat{\rho}\hat{J}^{(\mu,\mu+1)}\right\} = -\kappa\,\Delta T^{(\mu,\mu+1)} . \tag{19.22}$$

Based on linear irreversible thermodynamics one expects a quasi-stationary state, thus all quantities entering (19.22) should eventually become independent of time. Furthermore, if one thinks of κ as a material property that should be the same for different parts of the same material, one expects κ to be independent of the site μ.

In some cases the complete current operator is a conserved quantity of the considered system $[\hat{H}, \hat{J}\,] = 0$, i.e., once prepared a current in a closed loop system will never vanish. In this case there is no finite conductivity κ and Fourier's law is not fulfilled. This is a well known behavior in classical heat conduction scenarios. A famous example in this context is the harmonic chain, which does not have any resistance for the energy current. As already mentioned such a transport behavior is also called "ballistic transport". In solid state theories one needs anharmonic potentials to overcome this problem and implement a finite heat conductivity in a crystal. As long as we imagine the heat current as a quasi-particle transport problem (phonons, magnons), these potentials are interpreted as scattering processes for the respective quasi particles in the crystal.

For the models introduced above, we prove that the current operator is a conserved quantity depending on the commutators $[\hat{H}_{\text{loc}}, \hat{J}^{(\mu)}]$, $[\hat{H}_{\text{F}}, \hat{J}^{(\mu)}]$ and $[\hat{H}_{\text{NR}}, \hat{J}^{(\mu)}]$. The first commutator is zero according to the definition of the current itself. In a loop, i.e., summing over all commutators for all currents $\hat{J}^{(\mu)}$ between two subsystems, we find

$$\sum_{\mu} \left[\hat{H}_{\text{F}}, \hat{J}^{(\mu)}\right] = 0 \,, \quad \sum_{\mu} \left[\hat{H}_{\text{NR}}, \hat{J}^{(\mu)}\right] \neq 0 \,. \tag{19.23}$$

This means that in a chain, which is only coupled by a Förster type interaction, there is no finite conductivity, because the current is a conserved quantity in the system. However, if the system is coupled, additionally, with a non-resonant interaction, a finite heat conductivity is possible. For random couplings, the interaction does not commute in most cases with the current operator, so we also find a normal heat conduction in such coupling models. According to the classical theories of Sect. 19.1.2 we expect deviations from normal conductivity also in the case of very low temperatures, since in this limit the Hamiltonian of the system can be approximated by (19.3).

19.2.3 Linear Stability Analysis

For a numerical solution of the Liouville equation with incoherent damping terms according to the Lindblad formalism, we have to integrate the linear differential equation (19.12). This can be done by numerical diagonalization of the super-operator $\hat{\mathcal{L}}$, solving the equation in the diagonal basis and transforming back to get the whole time dependent solution of the problem. Since we are only interested in the stationary state of the problem, or in the properties of this state like local temperatures, heat currents, etc., we do not need the full time dependent solution.

The complete super-operator $\hat{\mathcal{L}}$ is a time independent operator. Therefore the linear system of differential equations is asymptotically stable, if all real parts of the eigenvalues of $\hat{\mathcal{L}}$ are less than zero. Indeed, we find that this condition is fulfilled: all real parts of the eigenvalues are less than zero except one eigenvalue, which is exactly zero. Thus, independent of the initial state, the system reaches the same stationary state, the state given by the eigenstate belonging to eigenvalue zero. We use this insight to compute directly the stationary state by looking for this special eigenstate of the system. After an appropriate normalization the stationary state $\hat{\rho}_{\text{stat}}$ is reached for all initial states after relaxation.

The different properties of the stationary state can be computed directly: the monotonous measure of the local temperature of the single two-level subsystems in the stationary state (cf. (19.15))

$$E_{\text{loc}}^{(\mu)} = T^{(\mu)} = \text{Tr}\left\{\hat{\rho}_{\text{stat}}^{(\mu)}\,\hat{H}_{\text{loc}}^{(\mu)}\right\} , \quad 0 \leq T^{(\mu)} \leq 0.5 , \tag{19.24}$$

as well as the expectation value of the heat current operator in the stationary state

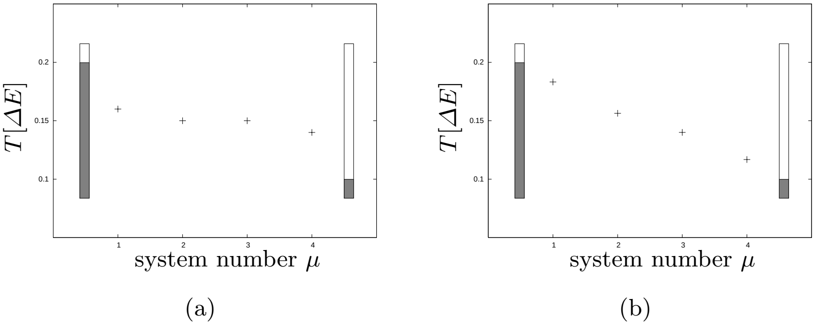
$$J^{(\mu,\mu+1)} = \text{Tr}\left\{\hat{\rho}_{\text{stat}}\,\hat{J}^{(\mu,\mu+1)}\right\} . \tag{19.25}$$

We now turn to study the non-equilibrium properties of the small quantum systems introduced in Sect. 19.2.1.

19.2.4 Heat Conduction in Low Dimensional Systems

We consider a chain of N spins coupled at both ends with the same coupling strength to separate baths of different temperature. For the higher temperature bath c_1 we set a temperature in our definition of $T^{c_1} = 0.2$, coupled to the first spin ($\mu = 1$) in the chain. The lower temperature bath c_2 ($T^{c_2} = 0.1$) is coupled to the last spin in the chain. We assume a weak coupling between the spins and their respective baths as well as within the system.

Here we investigate all three different types of internal couplings between the spins to test the dependence on the interaction. Firstly we consider the coupling due to an energy transfer only (Förster interaction). The two edge spin systems are drawn to the temperatures of their respective bath ($T^{(1)} =$



(a) (b)

Fig. 19.1. Temperature profile for a spin chain ($N = 4$) coupled to heat baths at both ends (bars in the boxes mark the appropriate temperature of the bath $T^{c_1} = 0.2, T^{c_2} = 0.1$). (**a**) Internal next neighbor Förster coupling, (**b**) internal next neighbor Förster and non-resonant coupling (Heisenberg interaction). Temperature defined as in (19.24).

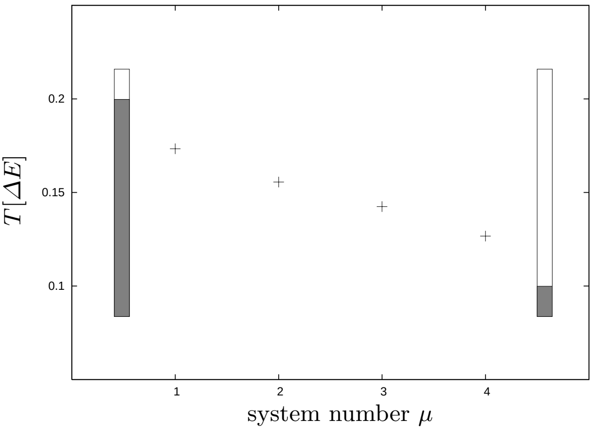


Fig. 19.2. Temperature profile: open system with weak internal random next neighbor coupling; bath temperatures $T^{(\mathrm{B1})} = 0.2$, $T^{(\mathrm{B2})} = 0.1$.

0.16 and $T^{(4)} = 0.14$, respectively, cf. Fig. 19.1(a)), whereas the two systems in the middle are exactly on the same average temperature, $T^{(2)} = T^{(3)} = 0.15$. As expected, we do not find a normal heat conduction in the middle of the system. Of course, the current does not diverge, because it is limited by the resistance of the bath contact, but, nevertheless, Fourier's law seems not to be fulfilled.

Now we couple the subsystems in the chain additionally by means of a non-resonant interaction, which by itself would not give rise to any energy exchange. This special coupling is also known as Heisenberg interaction. In this special case we find the non-vanishing temperature gradient, independent of the initial state, shown in Fig. 19.1(b). Here, Fourier's law is fulfilled showing up as a finite temperature difference in the middle of the system. This special coupling scenario will be analyzed in more detail in the next section.

Before we consider the Heisenberg chain in more detail, let us investigate a chain with random next neighbor couplings. This coupling is chosen to be weak again and we take the same coupling between different pairs of adjacent subsystems (i.e., no disorder). The random coupling is supposed to model "typical interactions" without any bias. Also in this case we find a linear temperature profile as shown in Fig. 19.2. Therefore the normal heat conduction seems to be a general result of this special class of models.

19.2.5 Fourier's Law for a Heisenberg Chain

For the normal heat conduction in the Heisenberg chain, i.e., a chain with an energy transfer coupling as well as with a non-resonant coupling, we will calculate some further interesting quantities. In Fig. 19.1(b) we have found a finite temperature gradient between the two central systems. For very low temperatures we argued in Sect. 19.2.2 that the gradient should vanish and we should find a flat profile like in the chains coupled by a Förster interaction only. To investigate this feature, we set the temperatures of the baths to $T^{c_1} = T_{\text{mean}} + \Delta T$ and $T^{c_2} = T_{\text{mean}} - \Delta T$. We choose a very small temperature difference $\Delta T = 0.0001$ and change the mean temperature within the range $T_{\text{mean}} = 0.00025 \ldots 0.49905$. In Fig. 19.3(a) we show the temperature difference between the two systems in the middle of the chain as a function of the mean temperature T_{mean}. As expected, the temperature difference vanishes for very low temperatures.

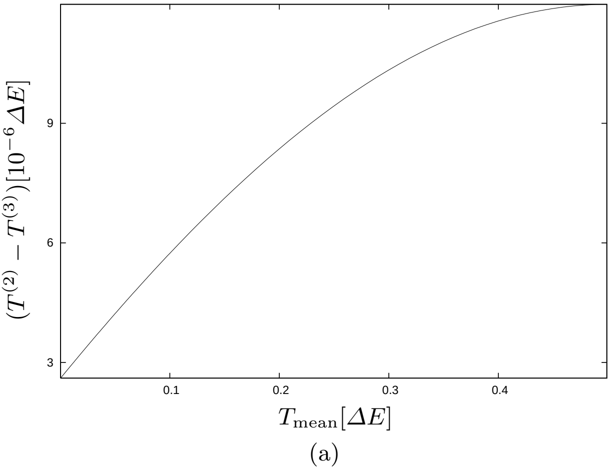
Another important quantity is the expectation value of the current operator. From the stationary state of the system we can evaluate this current according to (19.25). For the temperature difference between the two baths $\Delta T = 0.0001$ as in the above investigation and for different mean temperatures T_{mean}, we find that the current increases for lower mean temperatures (see Fig. 19.3(b)). This is a very small effect, as can be seen from the absolute amount of the increase of the current in comparison to the decrease of the temperature difference. The reason here is that the current is mainly determined by the contacts of the system to the heat baths. The resistance of these contacts is mainly responsible for the total amount of the current.

Finally, we are interested in the validity of Fourier's law in this small model system. Thus, we work at three different fixed mean temperatures $T_{\text{mean}} = 0.00015, 0.2015, 0.4515$, for which we change the temperature difference of the two heat baths $\Delta T = 0.001 \ldots 0.002$ and compute the current as well as the temperature difference of the two central systems. In Fig. 19.4(a) we show the current of the local equilibrium state $J^{(2,3)}$ as a function of the temperature difference between the two central systems. As can be seen, we get a very good linear behavior for each mean temperature and Fourier's law seems to be fulfilled. The gradient of the straight line defines the conductivity κ, which obviously depends on the mean temperature chosen.

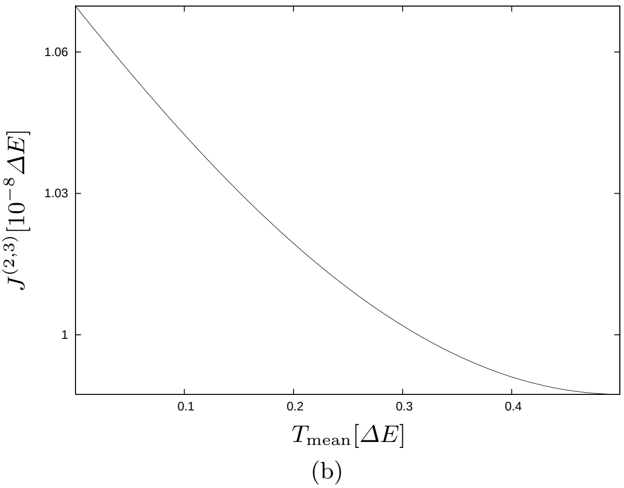
In Fig. 19.4(b) the dependence of the conductivity κ on the mean temperature is depicted over the whole temperature range. For low temperatures the conductivity increases, as expected from the considerations according to Sect. 19.2.2.

19.2.6 Implications of These Investigations

For the overwhelming majority of coupling scenarios we have found a normal heat conduction within our quantum mechanical approach. Even in the case of an unbiased random coupling within the considered chains, a local equilibrium behavior and therefore a normal heat conduction shows up. Also in the special case of the Heisenberg coupling we are able to show the validity of Fourier's law. Only for a very special interaction type – pure Förster interaction – does the conductivity diverge and we do not get a normal heat conduction.



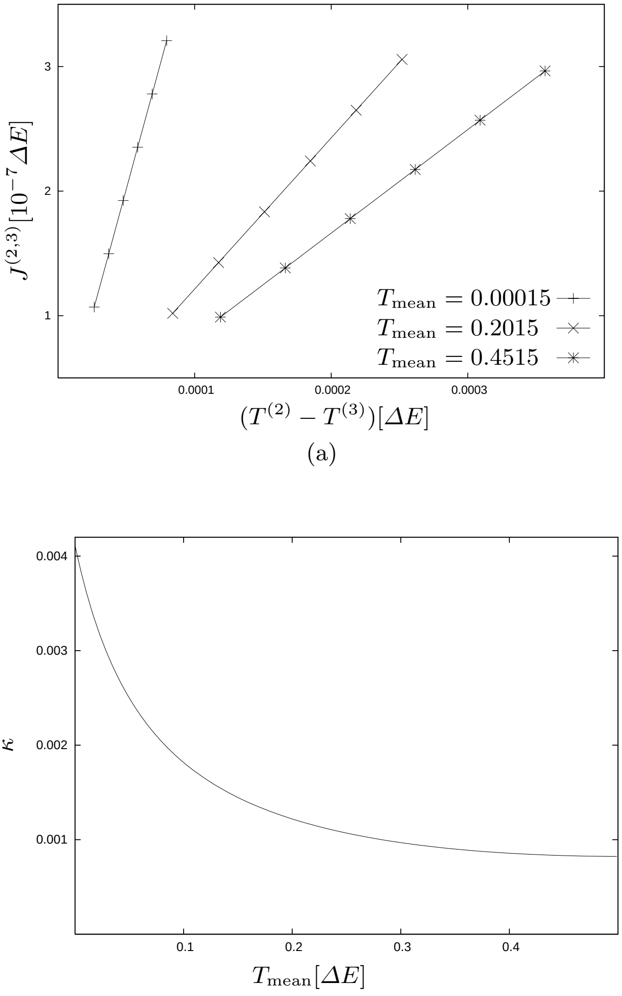
(a)



(b)

Fig. 19.3. (**a**) Temperature difference between the two central systems (2,3) of a four spin chain. (**b**) Current between the two central systems, both as a function of the mean temperature $T_{\text{mean}}(0.00025 \leq T_{\text{mean}} \leq 0.49905)$. Temperature defined as in (19.24).

To deduce these results no further assumptions are needed, only the principle of open quantum systems. Furthermore, this approach is valid even for very small systems, like the considered chain models with only four spins, where a quasi-particle approach seems more than problematic. Therefore we think that also in the region of non-equilibrium thermodynamics a quan-



(a)

(b)

Fig. 19.4. (**a**) Validity of Fourier's law. Current over temperature difference between the two central systems (2,3). The temperature differences of the heat baths are taken at mean temperatures $T_{\text{mean}} = 0.00015, 0.2015, 0.4515$, respectively. (**b**) Dependence of the conductivity κ on the mean temperature.

tum approach could be very helpful. However, at the moment a full quantum mechanical investigation of large-scale heat transport phenomena is not available.

18 Equilibrium Properties of Model Systems

> *Statistical mechanics is not the mechanics of large, complicated systems; rather it is the mechanics of limited, not completely isolated systems.*
>
> — J. M. Blatt [15]

To illustrate the principles that have been analyzed analytically and in a rather abstract way mainly in Chap. 9, we now turn to some numerical data, based on a certain type of models. These models are still rather abstract and may thus be viewed as models for a whole class of concrete physical situations, which are always taken to consist of two subsystems (confer Fig. 7.1), the considered system (gas, g) and the environment (container, c). These subsystems are specified by their respective spectra only, or, rather, those respective parts that play any role at all under the condition of energy conservation. What sort of physical structure could give rise to those spectra is not considered here, since it turns out to be irrelevant to some extent.

We want to analyze these models on the level of discrete, closed, quantum mechanical Hamilton models, for which a discrete, finite dimensional Schrödinger equation has to be solved. The spectra are translated into discrete diagonal matrix Hamiltonians that describe the decoupled bipartite system. These can be chosen to be diagonal, without any loss of generality, since any such system may be analyzed in the energy eigenbasis of its decoupled parts. The form of the interaction depends on the concrete physical subsystems and their interactions. However, since the "guess" is that for the quantities considered here (entropy, occupation probabilities, etc.) the concrete form of the interaction should not matter, the interaction is taken as some random matrix thus avoiding any bias. The interaction has to be "small" (see Sect. 7.1). In this way it is hoped that we get models that are as "typical" for general thermodynamic situations as possible. Many of the situations analyzed in this chapter are very similar to those known from the context of quantum master equations. However, note that in order to apply the theories at hand to those systems, neither a Markovian nor a Born assumption has to hold. In order to find the equilibrium state those concepts are not needed.

18.1 Entropy Under Microcanonical Conditions

All data in our first example refer to the situation depicted in Fig. 18.1 (cf. [20]). The "gas" (the system under consideration) consists of a two-level system, both levels being non-degenerate ($N_0^{\mathrm{g}} = N_1^{\mathrm{g}} = 1$), while the "container"

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 185–217 (2004)
http://www.springerlink.com/

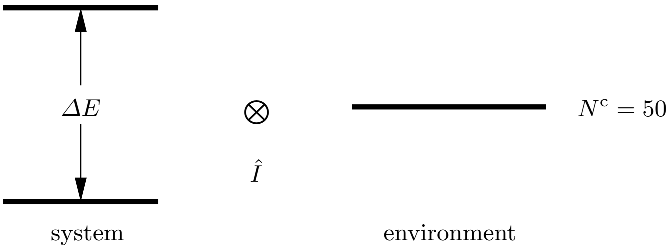


Fig. 18.1. Microcanonical scenario: a non-degenerate two-level-system (gas) is weakly coupled to a system with one energy level of degeneracy $N^{\mathrm{c}} = 50$. This is a model for a system in contact with a much larger environment such that no energy can be exchanged between the system and environment.

(the environment) consists of just one energy level with degeneracy $N^{\mathrm{c}} = 50$. This is necessarily a microcanonical situation regardless of the interaction $\hat{I}$. The container cannot absorb any energy, therefore energy cannot be exchanged between the systems. In this situation the probabilities of finding the gas system in the ground (excited) state W_0^{g} (W_1^{g}) are conserved quantities and in this example chosen as

$$W_0^{\mathrm{g}} = 0.15 \ , \quad W_1^{\mathrm{g}} = 0.85 \ . \tag{18.1}$$

As described in Sect. 9.1, the Hilbert space average of the purity of the gas system is given according to (9.24) by

$$[\![P^{\mathrm{g}}]\!] \approx \sum_{A=1}^{2} \frac{(W_A^{\mathrm{g}})^2}{N_A^{\mathrm{g}}} + \sum_{B=1}^{50} \frac{(W_B^{\mathrm{c}})^2}{N_B^{\mathrm{g}}} = 0.765 \ . \tag{18.2}$$

This should hold for large enough degeneracies of the occupied energy levels (see (9.23)), which is here the case because

$$N_0^{\mathrm{g}} N^{\mathrm{c}} + 1 = 51 \approx N_0^{\mathrm{g}} N^{\mathrm{c}} = 50 \ , \quad N_1^{\mathrm{g}} N^{\mathrm{c}} + 1 = 51 \approx N_1^{\mathrm{g}} N^{\mathrm{c}} = 50 \ . \tag{18.3}$$

The minimum purity can be computed by (9.9) and we find

$$P_{\mathrm{min}}^{\mathrm{g}} = \sum_{A=1}^{2} \frac{(W_A^{\mathrm{g}})^2}{N_A^{\mathrm{g}}} = 0.745 \ . \tag{18.4}$$

As explained in Sect. 9.1, we find here

$$[\![P^{\mathrm{g}}]\!] \approx P_{\mathrm{min}}^{\mathrm{g}} \ , \tag{18.5}$$

a situation in which almost the entire accessible region would be filled with the compartment containing only states of almost maximum local entropy.

To examine this expectation, a set of random states, uniformly distributed over the accessible region, has been generated. Their local entropies have been

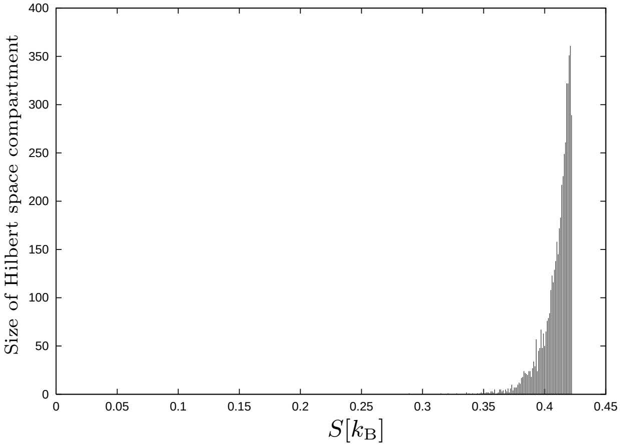


Fig. 18.2. Relative size of Hilbert space compartments: this histogram shows the relative frequency of states with a given local entropy S, among all states from the accessible region. In this case the maximum possible entropy is $S^\mathrm{g}_\mathrm{max} = 0.423\,k_\mathrm{B}$. Obviously, almost all states feature entropies close to the maximum.

calculated and sorted into a histogram. Since those states are distributed uniformly over the accessible region, the number of states in any "entropy bin" reflects the relative size of the respective Hilbert space compartment.

The histogram is shown in Fig. 18.2. The maximum local entropy in this case is $S^\mathrm{g}_\mathrm{max} = 0.423\,k_\mathrm{B}$. Obviously, almost all states have local entropies close to $S^\mathrm{g}_\mathrm{max}$. Thus compartments corresponding to entropies of, say, $S^\mathrm{g} > 0.4\,k_\mathrm{B}$ indeed fill almost the entire accessible region, just as theory predicts. Local pure states ($S^\mathrm{g} = 0$) are practically of measure zero.

In order to examine the dynamics, a coupling $\hat{I}$ is introduced. To keep the concrete example as general as 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 matrix elements of zero mean and a standard deviation of

$$\Delta I = 0.01 \Delta E \,. \tag{18.6}$$

This coupling is weak, compared to the Hamiltonian of the uncoupled system. Therefore the respective interaction cannot contain much energy. The spectrum of the system (see Fig. 18.1) does not change significantly due to the coupling, and after all the environment is not able to absorb energy.

Now the Schrödinger equation for this system, including a realization of the interaction, has been solved for initial states consistent with (18.1). Then the local entropy at each time has been calculated, thus resulting in a picture of the entropy evolution. The result is shown in Fig. 18.3. Obviously the entropy approaches $S^\mathrm{g}_\mathrm{max}$ within a reasonable time, regardless of the concrete initial state. Thus the tendency towards equilibrium is obvious. The concrete

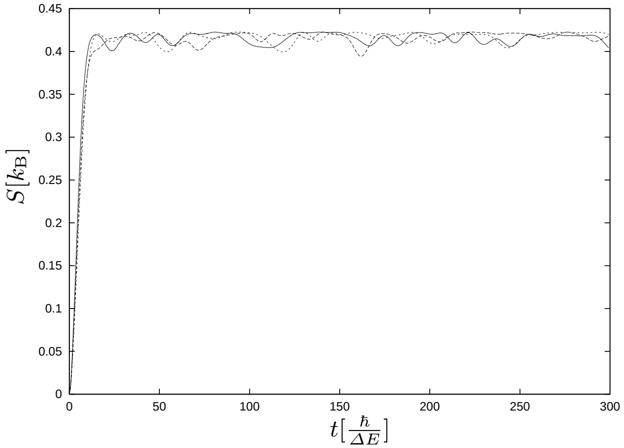


Fig. 18.3. Evolution of the local entropy for different initial states. A universal state of maximum entropy (equilibrium) is reached, independent of the initial state.

form of the interaction $\hat{I}$ only influences the details of this evolution, the equilibrium value is always the same. If the interaction is chosen to be weaker, the time scale on which equilibrium is reached gets longer, but, eventually the same maximum entropy will be reached in any case.

18.2 Occupation Probabilities Under Canonical Conditions

Just like in Sect. 18.1 we present some numerical data (cf. [20]) to support the principles derived in Sect. 9.2. The first model, which has been analyzed numerically to illustrate the above mentioned principles, is depicted in Fig. 18.4. The considered (gas) system, again, consists only of a non-degenerate two-level system. The environment (container) system in this case is a three-level system with an exponential "state density": $N_B^{\mathrm{c}} = 50 \cdot 2^B$ with $B = 0, 1, 2$. This has been chosen since for such a degeneracy scheme of the environment theory predicts an equilibrium state of the gas system, which should be independent of its initial state (see (9.47)). 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 those given could be reached, even if they were present. This is due to energy conservation and holds for the limit of weak interactions $\hat{I}$.

In this case the model can also be seen as a model for a situation with a much larger environment and we find from (9.48)

$$W^{\mathrm{d}}(E_0^{\mathrm{g}}) = \frac{2}{3}\,, \quad W^{\mathrm{d}}(E_1^{\mathrm{g}}) = \frac{1}{3}\,. \tag{18.7}$$

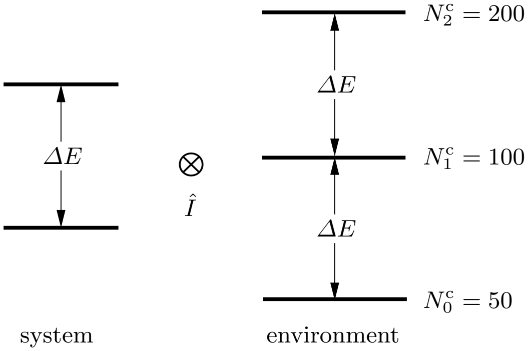


Fig. 18.4. Canonical scenario: a two-level gas system is weakly coupled to a three level environment, such that energy can be exchanged. The exponential degeneracy scheme of the container system guarantees a full independence of the equilibrium state from the initial state.

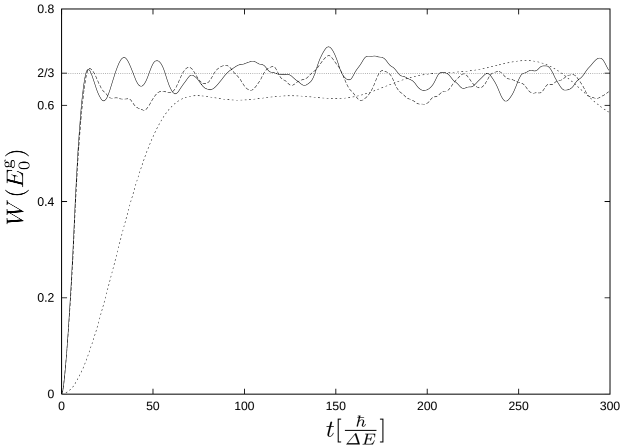


Fig. 18.5. Evolution of the ground level occupation probability for three different random interactions. The dotted line corresponds to a weaker interaction. Even in this case the same equilibrium value, $W_A^{\mathrm{d}} = \frac{2}{3}$, is approached, only on a longer timescale.

To keep the situation as general as possible, $\hat{I}$ was, like in Sect. 18.1, chosen to be a matrix with random Gaussian distributed entries in the basis of the eigenstates of the uncoupled system, but now with energy transfer allowed between the subsystems.

For this system the Schrödinger equation has been solved and the evolution of the probability of finding the gas system in its ground state, $W(E_0^{\mathrm{g}})$ is displayed in Fig. 18.5. The different curves correspond to different interaction strengths, given by the standard deviation of the distribution of the matrix elements of $\hat{I}$, ΔI:

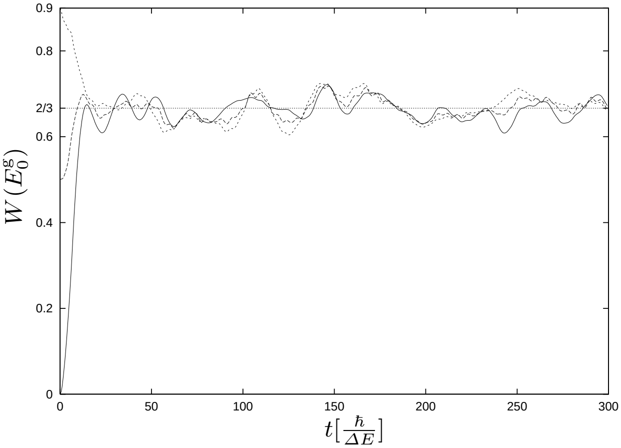


Fig. 18.6. Evolution of the ground level occupation probability for different initial states. The theoretically predicted equilibrium value is reached, independent of the initial states, as expected for canonical conditions.

$$\Delta I_{\text{solid, dashed}} = 0.0075 \Delta E \ , \quad \Delta I_{\text{dotted}} = 0.002 \Delta E \ . \tag{18.8}$$

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

Figure 18.6 displays the evolution of the same probability, $W(E_0^{\mathrm{g}})$, but now for different initial states, featuring different probabilities for the ground-state, as can be seen in the figure at $t = 0$. The equilibrium value is reached for any such evolution, regardless of the special initial state, thus we confirm the effective attractor behavior typical for thermodynamics.

Figure 18.7 displays the evolution of the local entropy of the gas system for the same three initial states as used for Fig. 18.6.

The maximum entropy, consistent with the equilibrium value of the energy probabilities, is $S_{\max}^{\mathrm{g}} = 0.637 \, k_{\mathrm{B}}$. This is also the value one finds, if one maximizes entropy for fixed mean energy (Jaynes' principle). Obviously, this value is reached for any initial state during the concrete dynamics of this model. This supports the validity of (9.45), 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 have finally investigated the system depicted in Fig. 18.8. Here the "gas" system is a non-degenerate equidistant five-level system and the container system a five-level system with degeneracies $N_B^{\mathrm{c}} = 6 \cdot 2^B$ $(B = 0, \ldots, 4)$, which should lead to a Boltzmann distribution. We restrict ourselves to initial states, where for both subsystems only the intermediate energy level is occupied (symbolized by the black dots in Fig. 18.8). Due to energy conservation other states of the container system would not play any role in this case even if they were

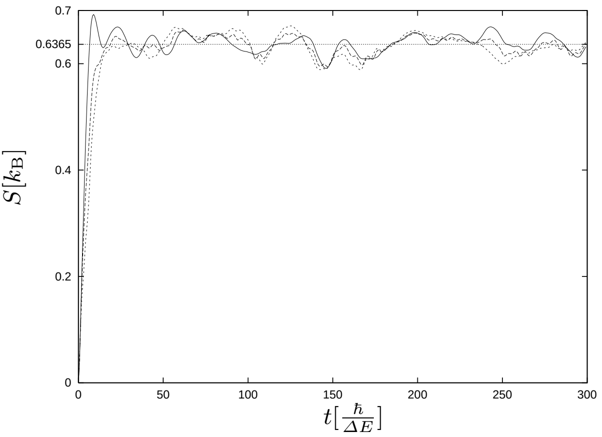


Fig. 18.7. Evolution of the local entropy for different initial states. $S = 0.637\, k_\mathrm{B}$ is the maximum entropy that is consistent with the equilibrium energy probabilities. This maximum entropy state is reached in all cases.

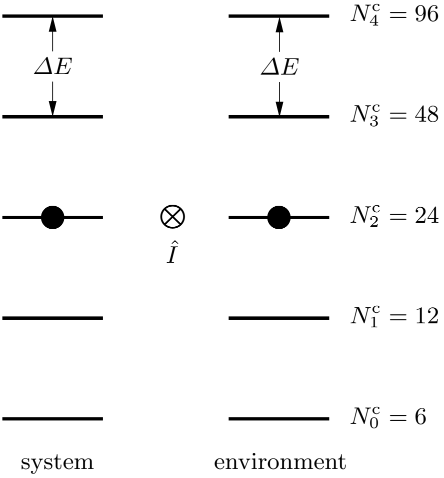


Fig. 18.8. Canonical multi-level scenario: a five-level gas system is weakly coupled to a five level container system with an exponential degeneracy scheme, such that energy may be exchanged. Black dots symbolize the initial state. This set-up should lead to a Boltzmann distribution.

present, just like in the previous model. Figure 18.9 shows the probabilities $W(E_A^\mathrm{g})$ of the different energy levels to be occupied. While the gas system starts in the intermediate (third) energy level, soon a Boltzmann distribution

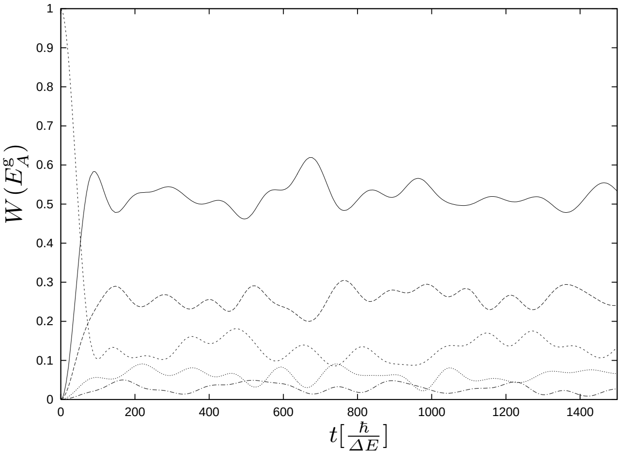


Fig. 18.9. Evolution of the energy occupation probabilities. After some relaxation time a Boltzmann distribution is reached. Each probability is twice as high as the one for the next higher energy level, as theory predicts.

develops. Obviously, each probability becomes twice as high as the one for the level above. This is exactly what theory predicts (see (9.42)) for the environment degeneracy scheme in this model.

18.3 Probability Fluctuations

To underline the results of Sect. 9.4 by some numerical data, a system almost like the one depicted in Fig. 18.4 is analyzed, but now with a degeneracy scheme given by

$$N_B^{\mathrm{c}} = \frac{N_1^{\mathrm{c}}}{2} \cdot 2^B \ . \tag{18.9}$$

The ratios between the degrees of degeneracy of the different container levels are thus the same as for the system sketched in Fig. 18.4, but the overall size of the container system is tunable by N_1^{c}. For various N_1^{c}, the Schrödinger equation has 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

$$\Delta_t^2 W_0^{\mathrm{g}} := \frac{1}{t_f - t_i}\left(\int_{t_i}^{t_f} \left(W_0^{\mathrm{g}}(t)\right)^2 \mathrm{d}t - \left(\int_{t_i}^{t_f} W_0^{\mathrm{g}}(t)\,\mathrm{d}t\right)^2\right) \ , \tag{18.10}$$

for initial states with

$$W_0^{\mathrm{g}}(0) = 0.2 \ , \quad W_1^{\mathrm{g}}(0) = 0.8 \ . \tag{18.11}$$

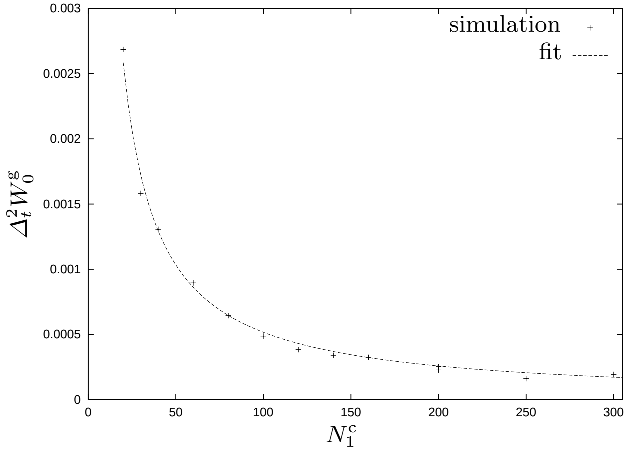


Fig. 18.10. Fluctuations of the probability for the considered system to be in the ground state $\Delta_t W_0^{\mathrm{g}}$, in dependence on the number of eigenstates of the environment system N_1^{c}. Obviously, the fluctuations go down, with increasing environment.

Figure 18.10 shows the dependence of the size of these fluctuations on the container system size N_1^{c}. The small crosses are the computed data points, the dashed line is a least square fit to a function proportional to $1/N_1^{\mathrm{c}}$. Obviously this function fits very well, confirming that fluctuations vanish like the square root of the system size (9.62). The fit reads

$$\Delta_t W_0^{\mathrm{g}} = \sqrt{\frac{0.053}{N_1^{\mathrm{c}}}} \,. \tag{18.12}$$

For the situation described above, we get, using the techniques of Sect. 9.4, especially using (9.64),

$$\Delta W_0^{\mathrm{g}} = \sqrt{\frac{0.053}{N_1^{\mathrm{c}}}} \tag{18.13}$$

and therefore the above fit is in very good agreement with the theoretical prediction, although the trajectories are not ergodic.

18.4 Spin Systems

So far we have illustrated that thermodynamic aspects can be observed in a great variety of bipartite few level systems, just as theory predicts. The bipartite system consists of a small system, the observed system as well as a larger system with some hundreds of energy levels, the environment. We have chosen the coupling between system and environment to be a random interaction, to avoid any bias.

Now we will show that based on theoretical concepts developed in this book, thermodynamic behavior can be found also in another class of systems – a class of modular systems (see Sect. 11.2) with only pair-interaction. These special interactions of subsystems considered here are far from being unbiased like a total random interaction would be. We will deal mainly with linear chains, with an identical interaction between each subsystem, e.g., a Heisenberg interaction. Nevertheless, we will find, that even in these special models, a thermodynamic behavior can be observed, without any further coupling to an environment.

The investigations here are structured in an analysis of global and local properties of such systems. Furthermore, we will observe the local behavior of these chains additionally coupled to an environment. Let us start with an investigation due to Jensen and Shankar in 1985 [58].

18.4.1 Global Properties

Jensen und Shankar considered a chain of $N = 7$ subsystems ($n = 2$ levels each), coupled by a next neighbor interaction [58]. They found hints of thermodynamic behavior in such a modular system. They investigated the Hamiltonian given by

$$\hat{H} = \sum_{\mu=1}^{N} \left(\theta_1 \, \hat{\sigma}_1^{(\mu)} + \theta_2 \, \hat{\sigma}_3^{(\mu)} \right) + \lambda \sum_{\mu=1}^{N} \hat{\sigma}_3^{(\mu)} \, \hat{\sigma}_3^{(\mu+1)} \, , \tag{18.14}$$

with the Pauli operators $\hat{\sigma}_i^{(\mu)}$ of the spin μ (see Sect. 2.2.2) and subject to cyclic boundary conditions (θ_1, θ_2 and λ are constants). After a numerical integration of the Schrödinger equation for the full system, they found an equilibrium value for some expectation values of global observables of the system. One interesting observable in this context is, e.g., the magnetization in x-direction of the whole system

$$\hat{M}_1 = \sum_{\mu=1}^{N} \hat{\sigma}_1^{(\mu)} \, . \tag{18.15}$$

The time dependence of the magnetization is shown in Fig. 18.11. Here one finds for an initial energy of $E = 4$ a mean magnetization of $\overline{M_1} = 2.46$ (solid line in Fig. 18.11). This mean value is independent of the concrete initial state of the system, but note that, of course, the equilibrium value depends on the total initial energy of this state.

To estimate the equilibrium magnetization of the system without a full solution of the Schrödinger equation we need to know the Hilbert space average of the expectation value of the magnetization over the corresponding accessible region. (Such Hilbert space averages over expectation values can be evaluated by the techniques shown in App. C.1.) In the present case the

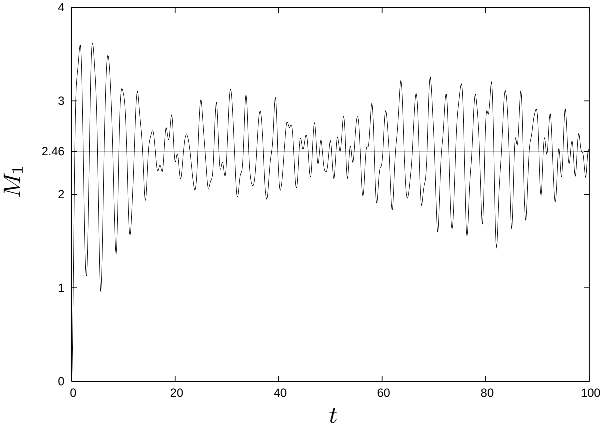


Fig. 18.11. Equilibrium behavior of the global magnetization of a spin system. Total energy of the initial state $E = 4$. The horizontal line marks the average magnetization $\overline{M_1} = 2.46$ ($\theta_1 = 1$, $\theta_2 = 0.5$ and $\lambda = 0.5$, cf. [58]).

accessible region is set by the mean energy and the energy width of the initial state. The uncertainty where, within the accessible region, to find the state at a given time is here due to not knowing the initial state in full detail rather than due to the unknown interaction with an environment. Such an environment does not even exist here, thus, e.g., an increase of entropy can certainly not be expected, nevertheless this case is similar to a "standard" microcanonical situation.

A computation of the average of the magnetization in the case on hand leads to an equilibrium value of the magnetization $M_1 = 2.35$. The comparison of the mean value of the time dependent expectation value of the magnetization with this equilibrium value shows good agreement.

18.4.2 Local Properties

Furthermore, one can observe the local behavior of such a modular system. If we imagine the system as a finite chain of subsystems (see Fig. 18.12 left hand side) and if we are only interested in the behavior of, say, the first spin, we may think of the rest of the system to be the environment for this singled out subsystem. Consider now the uncoupled system of $N-1$ subsystems together – the remaining spin should then be the considered system proper – we find N different energies for the system, each of which with a binomial distributed degeneracy, as shown in the right hand side of Fig. 18.12. For a weak coupling everywhere in the chain, the interaction provides for a small lifting of the degeneracy in the system, but leaves the spectrum qualitatively unchanged. All together the system is now in the form required for the theory. However, compared with the randomly chosen interactions, the coupling between the

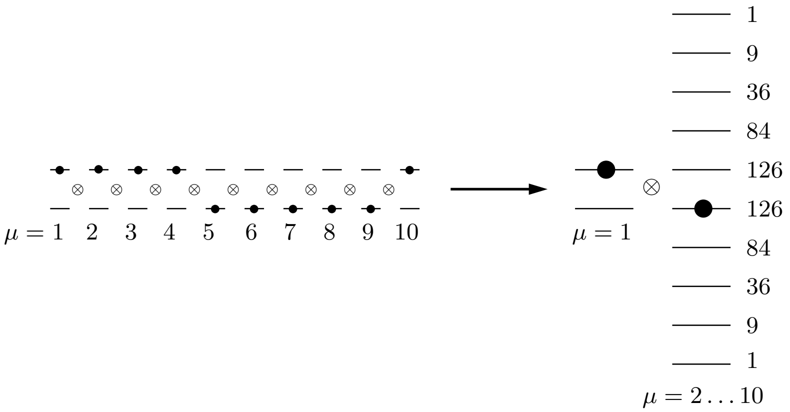


Fig. 18.12. Spectrum of an $N = 10$ spin system. On the left side, the chain of coupled spins is shown (the dots mark the initial product state). The right hand side shows the complete spectrum of $N = 9$ spins as an environment for the first spin of the chain. (Again, the dots mark the initial state)

spin of interest and the rest of the system (environment), is a very special coupling, resulting from the structure of the whole system.

In the concrete situation we are going to investigate now, the spins are coupled by a Heisenberg type of interaction with the total Hamiltonian

$$\hat{H} = \frac{1}{2} \sum_{\mu=1}^{N} \hat{\sigma}_3^{(\mu)} + \lambda \sum_{\mu=1}^{N-1} \sum_{i=1}^{3} \hat{\sigma}_i^{(\mu)} \hat{\sigma}_i^{(\mu+1)} , \tag{18.16}$$

where we choose λ to be very small in comparison to the local energy spread $\Delta E = 1$ of each spin ($\lambda = 0.002$).

In the middle of the spectrum we find the highest degeneracy and, since we need a sufficiently large subspace of Hilbert space to gain a thermodynamic behavior, we start in these highly degenerate subspaces (see Fig. 18.12). As an initial state we use product states of the uncoupled system with total energy $E = 5$. Again, the Schrödinger equation is solved for a system of $N = 10$ spins, but we observe only the first spin of the chain and a spin in the center of the chain ($\mu = 5$). The occupation probability of finding the respective subsystem in the excited energy level is shown in Fig. 18.13. In Fig. 18.13(a), we plot the probability for the first spin ($\mu = 1$) and in Fig. 18.13(b) for the spin $\mu = 5$. Both spins develop into the totally mixed state, which refers to an infinite temperature, in perfect agreement with the theoretical predictions. Since the considered system is in contact with an environment like in Fig. 18.12, where the upper band of the environment has as many levels as the lower one, we expect the system to be in this special equilibrium state.

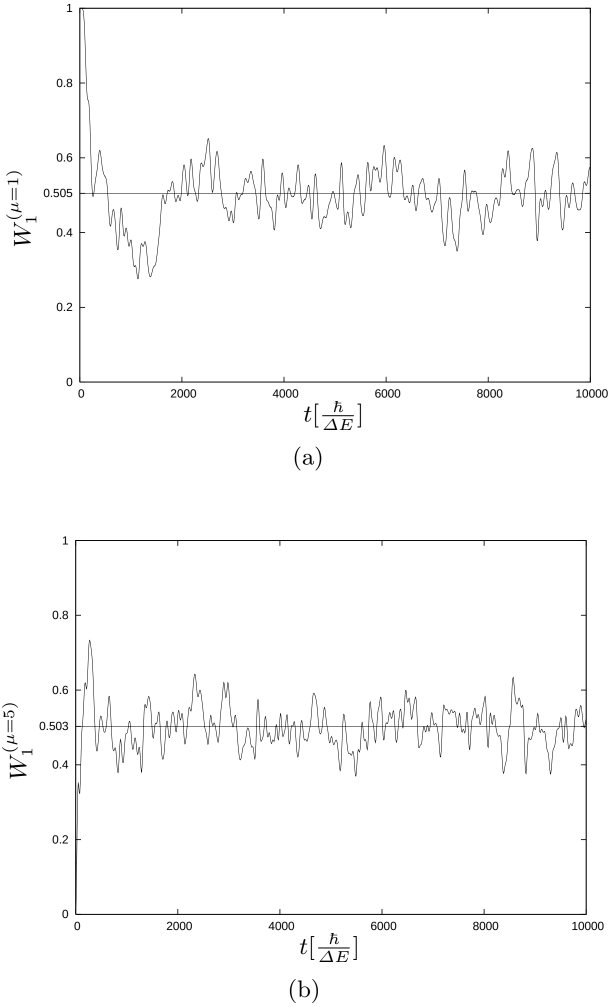


Fig. 18.13. Spin chain with Heisenberg interaction: probability of finding (**a**) the first (**b**) the fifth spin system in its upper level. Horizontal lines mark the mean value of the respective probability: (**a**) 0.504912 (**b**) 0.502989.

It is also possible to find such a local equilibrium behavior for a lower total energy of the system, but with much more fluctuations, because of the smaller parts of Hilbert space contributing to these calculations. However, if we would consider, say, some hundreds of spins, we expect to find a local thermodynamic behavior of each of the subsystems with a finite temperature and very small fluctuations.

18.4.3 Chain Coupled Locally to a Bath

After we have studied several properties of isolated chain like systems we now present yet another aspect of modular systems. Motivated by the spatial structure of our considered system it is now tempting to couple an environment not to the whole system but only at the edge of the system. Say, we will investigate systems locally coupled to an environment, as shown in Fig. 18.14. Interesting questions in this context are, whether the whole system is still to be found in the theoretically predicted equilibrium state and whether also the single subsystems are in the respective equilibrium state.

In Sect. 4.8 we have already introduced an approximate formalism to model the influence of an environment – the Lindblad formalism. Based on the full solution of the Schrödinger equation of the system with an environment, as has already been used several times in this chapter, we will demonstrate the equivalence of these two approaches.

Firstly we consider the solution of the full Schrödinger equation of the system, here two spins, and its environment. The two subsystems in the considered system together form a four-level system. Both interactions, the internal system interaction and the external environment coupling, are chosen randomly. As already done in Sect. 18.2, we use an environment of equidistant levels with a degeneracy increasing exponentially with energy. This environment is weakly coupled to the first spin only. For such an environment (c) theory would predict a canonical equilibrium state on our four-level system, with a definite temperature $T^{(c)}$ (see Sect. 9.2). The temperature of this four-level system should be the spectral temperature imposed by the degeneracy structure of the environment.

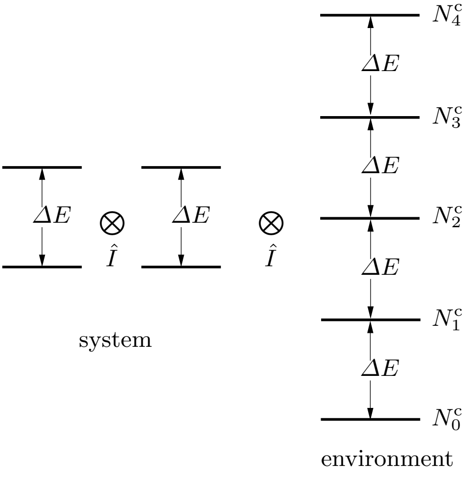


Fig. 18.14. Modular system built up of two spins weakly coupled to an environment, such that energy can be exchanged. Starting with the environment being in the third level an arbitrary initial state of the system is possible.

Let us now compare the equilibrium state of this system with the solution of the Liouville equation with a Lindblad operator as well as with the theoretical prediction according to Sect. 9.2. The rates for the Lindblad formalism are chosen such that the temperature is equivalent to the case of the solution of the Schrödinger equation and that the coupling to the environment is weak.

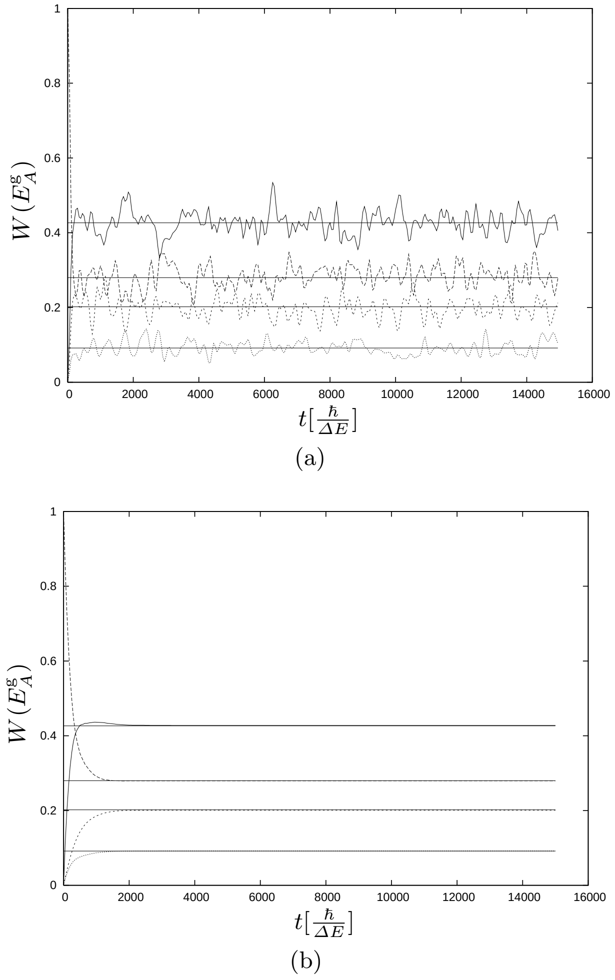


Fig. 18.15. Time dependence of the probabilities of finding the two spin system in one of its four eigenstates. (**a**) Full solution of the Schrödinger equation and (**b**) Lindblad formalism. The results of both approaches are in good accordance with the theoretically predicted values (solid lines) [54].

In Fig. 18.15 we plot the probability of finding the full four-level system in one of its eigenstates. The dynamics in Fig. 18.15(a) refers to the full Schrödinger case and in Fig. 18.15(b) one finds the dynamics according to the Lindblad formalism. Obviously, both approaches reach the same equilibrium state, which is in perfect agreement with the predictions of the theory developed in this book.

Furthermore, we are interested in the local temperatures of the single subsystems in comparison with the global temperature of the four-level system. Only if these two temperatures are equal, would temperature be an intensive and local quantity. As the pertinent temperature we use the spectral temperature of Chap. 12. Since the interaction between the two subsystems does not enter our temperature measure, we expect that we get a good agreement between global and local temperatures for internal weak couplings. Therefore we investigate the local temperature and its dependence on the internal coupling strength. In Fig. 18.16(a) we show the result for a full integration of the Schrödinger equation for given internal coupling strength λ, and in Fig. 18.16(b) the same investigation in the Lindblad formalism, together with the global fixed temperature. As can be seen in both cases, local and global temperatures are approximately the same for a small internal coupling strength λ, whereas for a larger coupling, global and local temperatures differ.

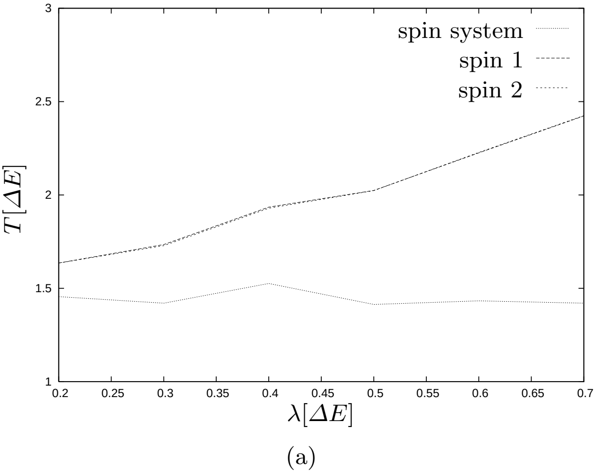
18.5 On the Existence of Local Temperatures*

As in Sect. 18.4 we consider a chain of spins with nearest neighbor interactions. However, instead of various values of the coupling strength we consider groups of different numbers of adjoining spins (see Fig. 18.17 and [46, 48]). The idea behind this approach is the following. If N adjoining spins form a group, the energy of the group is N times the average energy per spin and is thus expected to grow proportionally to N as the size of the group is increased. Since the spins only interact with their nearest neighbors, two adjacent groups only interact via the two spins at the respective boundaries. As a consequence, the effective coupling between two groups is independent of the group size and thus becomes less relevant compared to the energy contained in the groups as the group size increases.

Since we want to analyse the existence of a quantity usually assigned to equilibrium states, we should consider equilibrium scenarios or at least situations close to equilibrium. Therefore we assume that our entire spin chain is in a thermal equilibrium state. One can imagine that it may have approached this state via interactions with its surrounding as described in Chap. 9, although, for the consideration here, such details will be irrelevant.

---

[*] Based on [46, 47, 48, 49] by Hartmann et al.



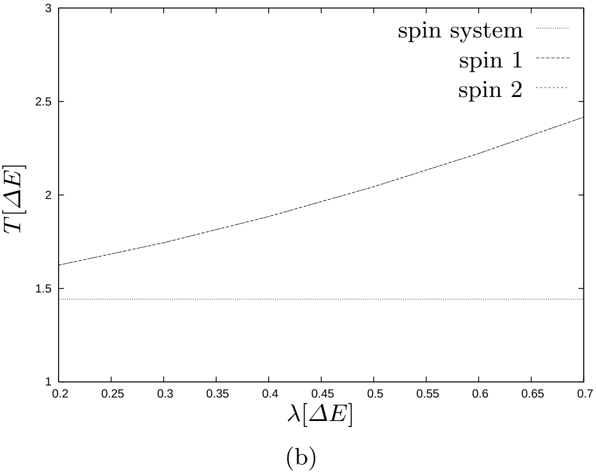


Fig. 18.16. The global (bottom solid line) and the local temperatures of the two single subsystems as a function of the internal coupling strength λ. (**a**) Solution of the full Schrödinger equation and (**b**) Lindblad formalism [54].

Before we can address the question of local temperatures, we have to clarify what we mean when we say that temperature exists or does not exist. The spectral temperature defined in Chap. 12 always exists, but it usually does not have all the properties temperature is supposed to have in thermal equilibrium.

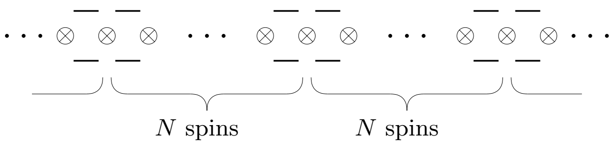


Fig. 18.17. N^{G} groups of N adjoining spins each are formed.

We adopt here the convention that local temperature exists if the respective (local) reduced density matrix is close to a canonical one. Then, the spectral temperature, which in this case coincides with the standard temperature, fully characterizes the distribution, i.e., the diagonal elements of the corresponding density matrix. If, furthermore, the local temperatures coincide with the global one, temperature is even an intensive quantity in the scenario at hand (i.e., does not change with system size).

We thus consider a very long chain of very many spins in a global thermal equilibrium state (9.49), divide the chain into N^{G} groups of N adjoining spins each (see Fig. 18.17) and test whether the local probability distribution also has the canonical form (9.49).

18.5.1 Model

We start by defining the Hamiltonian of our spin chain in the form,

$$\hat{H} = \sum_{\mu} \left(\hat{H}_{\mathrm{loc}}^{(\mu)} + \hat{H}_{\mathrm{int}}^{(\mu,\mu+1)} \right) , \tag{18.17}$$

where the index μ labels the elementary subsystems. The first term is the local Hamiltonian of subsystem μ and the second one describes the interaction between subsystem μ and $\mu + 1$. We assume periodic boundary conditions. Since this section applies to all models with the structure (18.17), we do not further specify the terms in the Hamiltonian before we apply the results to the concrete spin chain model in Sect. 18.5.4.

We now form N^{G} groups of N subsystems each with index $\nu = 1, \ldots, N^{\mathrm{G}}$ specifying the respective group, and $\mu = 1, \ldots, N$ numbers the elementary subsystems within such a group

$$\mu \mapsto (\nu - 1)N + \mu \, . \tag{18.18}$$

According to the formation of groups the total Hamiltonian splits up into two parts,

$$\hat{H} = \hat{H}_0 + \hat{I} \, , \tag{18.19}$$

where $\hat{H}_0$ is the sum of the Hamiltonians of the isolated groups,

$$\hat{H}_0 = \sum_{\nu=1}^{N^{\mathrm{G}}} \hat{H}_{\mathrm{group}}^{(\nu)} \quad \text{with} \tag{18.20}$$

$$\hat{H}_{\mathrm{group}}^{(\nu)} = \sum_{\mu=1}^{N} \hat{H}_{\mathrm{loc}}^{((\nu-1)N+\mu)} + \sum_{\mu=1}^{N-1} \hat{H}_{\mathrm{int}}^{((\nu-1)N+\mu,(\nu-1)N+\mu+1)}$$

and $\hat{I}$ contains the interaction terms of each group with its neighbor group only

$$\hat{I} = \sum_{\nu=1}^{N^{\mathrm{G}}} \hat{H}_{\mathrm{int}}^{(\nu N,\nu N+1)} \; . \tag{18.21}$$

We label the eigenstates of the total Hamiltonian $\hat{H}$ and their energies with Greek letters (φ, ψ) and eigenstates and energies of the group Hamiltonian $\hat{H}_0$ with Latin letters (a, b)

$$\hat{H}|\varphi\rangle = E_\varphi|\varphi\rangle \; , \qquad \hat{H}_0|a\rangle = E_a|a\rangle \; . \tag{18.22}$$

Here, the states $|a\rangle$ are products of group eigenstates defined as

$$\hat{H}_{\mathrm{group}}^{(\nu)}|a_\nu\rangle = E_\nu|a_\nu\rangle \; , \qquad |a\rangle = \prod_{\nu=1}^{N^{\mathrm{G}}} |a_\nu\rangle \; . \tag{18.23}$$

E_ν is the energy of one subgroup only and $E_a = \sum_{\nu=1}^{N^{\mathrm{G}}} E_\nu$.

18.5.2 Global Thermal State in the Product Basis

We assume that the total system is in a thermal state with a density matrix $\hat{\rho}$, which reads in the eigenbasis of $\hat{H}$

$$\langle\varphi|\hat{\rho}|\psi\rangle = \frac{\mathrm{e}^{-\beta E_\varphi}}{Z} \, \delta_{\varphi\psi} \; . \tag{18.24}$$

Here, Z is the partition function and $\beta = 1/(k_{\mathrm{B}}T)$ the inverse temperature. Transforming the density matrix (18.24) into the eigenbasis of $\hat{H}_0$ we obtain

$$\langle a|\hat{\rho}|a\rangle = \int_{E_0}^{E_1} \mathrm{d}E \, W_a(E) \, \frac{\mathrm{e}^{-\beta E}}{Z} \tag{18.25}$$

for the diagonal elements in the new basis. Here, the sum over all states $|a\rangle$ has been replaced by an integral over the energy. E_0 is the energy of the ground state and E_1 the upper limit of the spectrum. The density of conditional probabilities $W_a(E)$ is given by

$$W_a(E) = \frac{1}{\Delta E} \sum_{|\varphi\rangle : E \le E_\varphi < E + \Delta E} |\langle a|\varphi\rangle|^2 \; , \tag{18.26}$$

where ΔE is small and the sum runs over all states $|\varphi\rangle$ with eigenvalues E_φ in the interval $[E, E + \Delta E]$.

To compute the integral of equation (18.25) we need to know the density of the conditional probabilities $W_a(E)$. For a very large number of groups, $N^{\mathrm{G}} \gg 1$, it may be approximated by a Gaussian normal distribution (for a rigorous proof of this statement, which is a quantum analog of the central limit theorem, and further applications, see [48] and [50]),

$$\lim_{N^{\mathrm{G}} \to \infty} W_a(E) = \frac{1}{\sqrt{2\pi}\Delta_a} \exp\left(-\frac{(E - E_a - \varepsilon_a)^2}{2\,\Delta_a^2} \right) , \tag{18.27}$$

where ε_a and Δ_a are defined by

$$\varepsilon_a = \langle a|\hat{H}|a\rangle - \langle a|\hat{H}_0|a\rangle , \tag{18.28}$$

$$\Delta_a^2 = \langle a|\hat{H}^2|a\rangle - \langle a|\hat{H}|a\rangle^2 . \tag{18.29}$$

The quantity ε_a has a classical counterpart, while Δ_a^2 is purely quantum mechanical. It appears because the commutator $[\hat{H}, \hat{H}_0]$ is non-zero, and the distribution $W_a(E)$ therefore has non-zero width. Equation (18.25) can now be computed for $N^{\mathrm{G}} \gg 1$

$$\begin{aligned} \langle a|\hat{\rho}|a\rangle =& \frac{1}{Z} \exp\left(-\beta\, y_a + \frac{\beta^2 \Delta_a^2}{2} \right) \times \\ & \times \frac{1}{2} \left(\mathrm{erfc}\left(\frac{E_0 - y_a + \beta\Delta_a^2}{\sqrt{2}\Delta_a} \right) - \mathrm{erfc}\left(\frac{E_1 - y_a + \beta\Delta_a^2}{\sqrt{2}\Delta_a} \right) \right) , \end{aligned} \tag{18.30}$$

where $y_a = E_a + \varepsilon_a$ and $\mathrm{erfc}(\dots)$ is the conjugate Gaussian error function. The second term only appears if the energy is bounded and the integration extends from the energy of the ground state E_0 to the upper limit of the spectrum E_1.

The off-diagonal elements $\langle a|\hat{\rho}|b\rangle$ vanish for $|E_a - E_b| > \Delta_a + \Delta_b$ because the overlap of the two distributions of conditional probabilities becomes negligible. For $|E_a - E_b| < \Delta_a + \Delta_b$, the transformation involves an integral over frequencies and thus these terms are significantly smaller than the entries on the diagonal part.

18.5.3 Conditions for Local Thermal States

We now test under what conditions the diagonal elements of the (local) reduced density matrices are also canonically distributed with some local inverse temperature $\beta_{\mathrm{loc}}^{(\nu)}$ for each subgroup $\nu = 1, \dots, N^{\mathrm{G}}$. Since the trace of a matrix is invariant under basis transformations, it is sufficient to verify that they show the correct energy dependence. If we assume periodic boundary conditions, all reduced density matrices are equal ($\beta_{\mathrm{loc}}^{(\nu)} = \beta_{\mathrm{loc}}$ for all ν) and the products of their diagonal elements are of the form $\langle a|\hat{\rho}|a\rangle \propto \exp(-\beta_{\mathrm{loc}} E_a)$. We thus have to verify that the logarithm of the right hand side of (18.30) is a linear function of the energy E_a,

$$\ln\left(\langle a|\hat{\rho}|a\rangle\right) \approx -\beta_{\mathrm{loc}}\, E_a + c \;, \tag{18.31}$$

where β_{loc} and c are real constants. Note that (18.31) does not imply that the occupation probability of an eigenstate $|\varphi\rangle$ with energy E_φ and a product state with the same energy $E_a \approx E_\varphi$ are equal. Even if β_{loc} and β are equal with very good accuracy, but not exactly the same, occupation probabilities may differ by several orders of magnitude, provided the energy range is large enough.

Since we consider the limit $N^{\mathrm{G}} \to \infty$, we approximate the conjugate error functions of (18.30) by their asymptotic expansions (cf. [1]). This is possible because y_a and Δ_a^2 are sums of N^{G} terms and the arguments of the error functions grow proportionally to $\sqrt{N^{\mathrm{G}}}$. Inserting the asymptotic expansions into equation (18.30) shows that (18.31) can only be true if

$$\frac{E_a + \varepsilon_a - E_0}{\sqrt{N^{\mathrm{G}}}\,\Delta_a} > \beta\,\frac{\Delta_a^2}{\sqrt{N^{\mathrm{G}}}\,\Delta_a} \;, \tag{18.32}$$

(for a more detailed consideration see [47]). In this case, (18.30) may be taken to read

$$\langle a|\hat{\rho}|a\rangle = \frac{1}{Z}\exp\!\left(-\beta\left(E_a + \varepsilon_a - \frac{\beta\Delta_a^2}{2}\right)\right) \;, \tag{18.33}$$

where we have used that $y_a = E_a + \varepsilon_a$. To ensure that the criterion (18.31) is met, ε_a and Δ_a^2 have to be of the form

$$-\varepsilon_a + \frac{\beta}{2}\,\Delta_a^2 \approx c_1 E_a + c_2 \;, \tag{18.34}$$

as follows by inserting (18.33) into (18.31). In (18.34), c_1 and c_2 are arbitrary, real constants. Note that ε_a and Δ_a^2 need not be functions of E_a and therefore in general cannot be expanded in a Taylor series. In addition, the temperature becomes intensive, if the constant c_1 vanishes, in

$$\beta_{\mathrm{loc}} = \beta(1 - c_1) \;, \tag{18.35}$$

which follows by introducing the approximated diagonal matrix element (18.33) into condition (18.30) and using (18.34). If c_1 does not vanish, temperature would not be intensive, although it might exist locally.

Conditions (18.32) and (18.34) determine when temperature can exist locally. These conditions depend on the global inverse temperature β. In both equations the inverse temperature β appears together with the squared width

Δ_a^2, which is always larger than zero as long as the commutator $[\hat{H}, \hat{H}_0]$ does not vanish. The temperature dependence of the criteria for the existence of local temperatures is thus a pure quantum effect, which appears in almost all models of interest since these show $[\hat{H}, \hat{H}_0] \neq 0$. As can already be deduced from (18.32) and (18.34), the temperature dependence is most relevant at low temperatures, where minimal length scales for the existence of temperature may even become macroscopic, as we will see below (cf. [6, 60, 91, 130]).

It is sufficient to satisfy the conditions (18.32) and (18.34) for an adequate energy range $E_{\text{min}} \leq E_a \leq E_{\text{max}}$ only. For large systems with a modular structure, the density of states is a rapidly growing function of energy, as explained in Chap. 11.2. If the total system is in a thermal state, occupation probabilities decay exponentially with energy. The product of these two functions is thus sharply peaked at the expectation value of the energy $\overline{E}$ of the total system $\overline{E} + E_0 = \text{Tr}\{\hat{H}\hat{\rho}\}$, with E_0 being the ground state energy. The energy range thus needs to be centered at this peak and to be large enough. On the other hand it must not be larger than the range of values E_a can take on. Therefore, a pertinent and "safe" choice for E_{min} and E_{max} is

$$E_{\text{min}} = \max\left([E_a]_{\text{min}}\ ,\ \alpha^{-1}\,\overline{E} + E_0\right)\ , \tag{18.36}$$

$$E_{\text{max}} = \min\left([E_a]_{\text{max}}\ ,\ \alpha\,\overline{E} + E_0\right)\ , \tag{18.37}$$

where $\alpha \gg 1$ and $\overline{E}$ will in general depend on the global temperature. $[x]_{\text{min}}$ and $[x]_{\text{max}}$ denote the minimal and maximal values that x can take on.

The expectation value of the total energy, $\overline{E}$, in general depends on the global temperature. This temperature dependence, which is characteristic of the model (i.e., of the material) at hand, also enters into the criteria (18.32) and (18.34) via the definition of the relevant energy range in (18.36) and (18.37).

Figure 18.18 compares the logarithm of (18.30) for a model, where $\varepsilon_a = 0$ and $\Delta_a^2 = E_a^2/N^2$ with $N = 1000$ and the logarithm of a canonical distribution with the same β. The actual density matrix is more mixed than the canonical one. In the interval between the two vertical lines, both criteria (18.32) and (18.34) are satisfied. For $E < E_{\text{low}}$ (18.32) is violated and (18.34) for $E > E_{\text{high}}$. To allow for a description by means of a canonical density matrix, the group size N needs to be chosen such that $E_{\text{low}} < E_{\text{min}}$ and $E_{\text{high}} > E_{\text{max}}$.

For the existence of a local temperature the two conditions (18.32) and (18.34), which constitute the general result of our approach, must both be satisfied. Or, to put it differently: the two criteria determine, for a concrete situation, the minimal number of group members N (a minimal length scale) on which temperature can be defined. These fundamental criteria will now be applied to a concrete example.

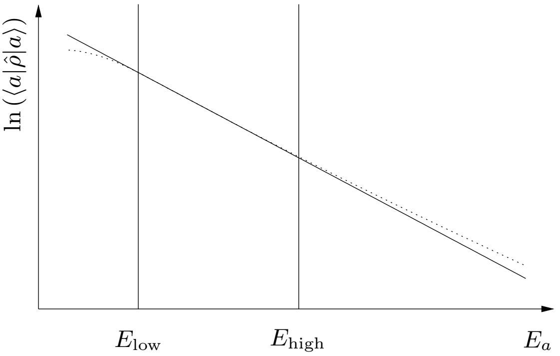


Fig. 18.18. $\ln(\langle a|\hat{\rho}|a\rangle)$ for a canonical density matrix ρ (dashed line) and for $\hat{\rho}$ as in (18.30) (solid line) with $\varepsilon_a = 0$ and $\varDelta_a^2 = E_a^2/N^2$ ($N = 1000$) [47].

18.5.4 Spin Chain in a Transverse Field

As a concrete application, we consider a spin chain in a transverse field. For this model the Hamiltonian reads

$$\begin{aligned}\hat{H}_{\mathrm{loc}}^{(\mu)} &= -\varDelta E\,\hat{\sigma}_3^{(\mu)}\\ \hat{H}_{\mathrm{int}}^{(\mu,\mu+1)} &= -\frac{\lambda}{2}\left(\hat{\sigma}_1^{(\mu)}\otimes\hat{\sigma}_1^{(\mu+1)} + \hat{\sigma}_2^{(\mu)}\otimes\hat{\sigma}_2^{(\mu+1)}\right).\end{aligned} \tag{18.38}$$

Here, $2\varDelta E$ is the Zeeman splitting (we will always assume $\varDelta E > 0$) and λ the strength of the spin-spin coupling, called Förster coupling.

If one partitions the chain into N^{G} groups of N subsystems each, the groups may be diagonalized via a Jordan-Wigner and a Fourier transformation [46]. With this procedure, the quantities E_a, ε_a and $\varDelta_a$ can be expressed in terms of fermionic occupation numbers of the group eigenstates, which in turn allows us to calculate the minimal number of spins N_{min} in each group (the minimal length scale) from conditions (18.32) and (18.34).

The technical details of this calculation, which involves a few further approximations, are not relevant for the understanding of the relevant physics. We thus refer the interested reader to App. E for a detailed discussion.

Figure 18.19 shows N_{min} as a function of temperature T in units of half the Zeeman splitting for weak coupling, $\lambda = 0.1\varDelta E$. Here, condition (18.32) is approximated by (E.7) and condition (18.34) by (E.10) (see App. E).

The strong coupling case ($\lambda = 10\varDelta E$) is shown in Fig. 18.20, where condition (18.32) is approximated by (E.9) and condition (18.34) by (E.10) (see App. E).

Apparently, criterion (18.34) and (E.10), respectively, become relevant only at high temperatures T and strong coupling $\lambda > \varDelta E$. In addition, temperature is always an intensive quantity for the spin chain considered here, irrespective of the coupling strength (see App. E for details).

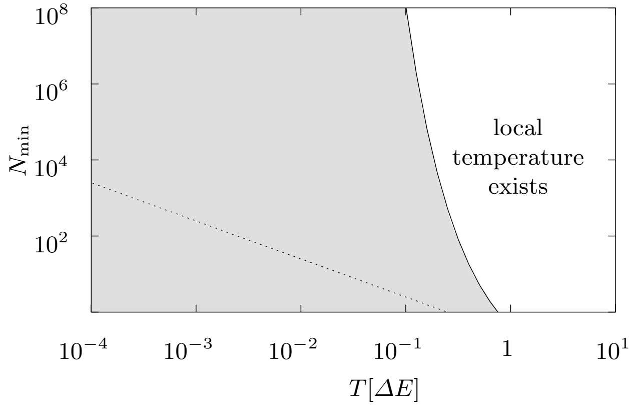


Fig. 18.19. Weak coupling ($\lambda = 0.1\Delta E$). Log-log-plot of $N_{\min}$ over T: the criteria (E.7) (solid line) and (E.10) (dashed line) define lower limits above which a local temperature is defined [47].

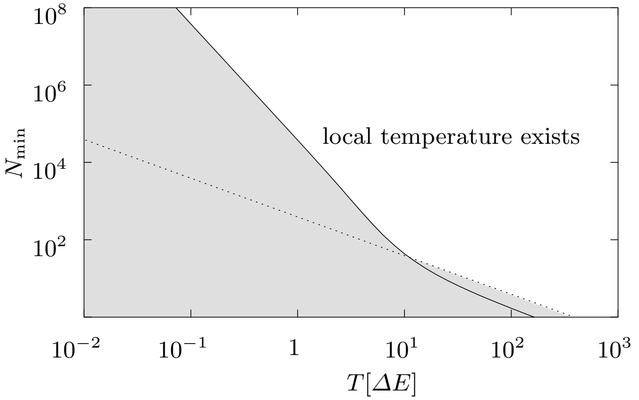


Fig. 18.20. Strong coupling ($\lambda = 10\Delta E$). Log-log-plot of $N_{\min}$ over T: the criteria (E.9) (solid line) and from (E.10) (dashed line) define a limit above which a local temperature is defined [47].

The model shows that for spin chains at high temperatures local temperatures may exist even for single spins, while the minimal group size becomes increasingly large at low temperatures. These characteristics are quite plausible if the transformation behavior of matrices under a change of basis is taken into account. At very high global temperatures, the total density matrix is almost completely mixed, i.e., proportional to the identity matrix, and thus does not change under basis transformations. There are therefore global temperatures that are high enough, so that local temperatures exist even for single spins.

On the other hand, if the global temperature is very low, the system is essentially in its ground state. This state is typically a superposition of product states and therefore strongly correlated [60, 130]. It becomes thus impossible to assign local temperatures to the ground state for any partition.

18.6 Quantum Thermometer

18.6.1 The Classical Picture

Temperature is usually associated with the "random" thermal motion, the paradigmatic model being the ideal classical gas. It thus appears meaningless to talk about the temperature of an individual particle. Temperature as a non-mechanical property is interpreted to result from some averaging over a sufficiently large sample of particles. With decreasing sample size (e.g., decreasing volume V for constant particle density) the statistics worsen and one naturally expects the temperature to be defined, if at all, with increasing uncertainty. This expectation seems to be in accord with the result for the mean variance of the temperature [69, 95]

$$\overline{(\varDelta T)^2} = k_{\mathrm{B}} \frac{T^2}{C_V} \,, \tag{18.39}$$

where C_V is the specific heat of the total system at constant volume V.

It is not clear, though, under what conditions this formula actually applies [86]. A system in contact with a thermostat should have a fixed temperature (as a boundary condition), irrespective of the size of the system. Furthermore, it is apparently possible to associate a temperature with the occupation of internal states of a single molecule or even a single spin.

18.6.2 The Szilard Picture

In his attempts to clarify confusing aspects of so-called Maxwell's demon, Szilard [123] introduced a simplified model of a heat engine, the central part of which was a single molecule in a box interacting with a heat bath. At the time, this model must have met with suspicion, as a particle in a box would seem to be an entirely mechanical problem rather than being subject to thermodynamic analysis.

In fact, this particle in a box has a simple quantum interpretation: it constitutes a system with a discrete but infinite spectrum. Its spatial extension is mesoscopic or even macroscopic and given by the box itself. However, besides this possible macroscopic size, the particle in a box is, in principle, not different from the internal (bound) spectrum of a molecule (essentially localized on a nanometer scale).

The Szilard picture fits perfectly into the approach of this book: thermodynamic properties result from the interaction of the quantum subsytem considered with its quantum environment. For weak coupling and some further requirements for the spectral density of the environment, a canonical equilibrium state of the subsystem results, even though the system as a whole had no thermodynamic properties whatsoever.

18.6.3 Temperature Fluctuations

Nevertheless, one might argue, temperature fluctuations do occur, and somehow they seem to be associated with small systems. These temperature fluctuations happen in time and for a single system, i.e., they cannot be related to the properties of an identically prepared ensemble, for which the respective temperature was taken to differ from one member to the other (statistical uncertainty). However, a closed system prepared in a canonical state cannot show temporal fluctuations, irrespective of size, as the state $\hat{\rho}$ is stationary, $[\hat{H}, \hat{\rho}] = 0$. Consequently, there are neither temperature nor entropy fluctuations.

Temporal fluctuations for an individual system in contact with an environment: this is exactly what comes out in our approach – when the relevant part of the environment (the source of thermodynamic behavior of the considered system) gets smaller and smaller. In this case we may still have a canonical state on time-average, but the occupation numbers fluctuate around their equilibrium values. These fluctuations can hardly be mapped onto a fluctuation of temperature only. A conceptually simple description was in terms of the spectral temperature (as introduced in Sect. 12.1), as this temperature can be defined for any state and any instant of time. Anyway, the fluctuations of occupation numbers will lead to fluctuations of moments of the system energy. From these moments one usually infers the temperature experimentally [86].

It is important to realize that for a molecular system (spin chain, particles in a box, etc.) partitions of the system itself (i.e., without having an external environment like a container) will already suffice to induce thermodynamic properties locally (cf. Chap. 18.4). After relaxation any local spin or particle should thus be expected to be in a thermal state (this holds for weak internal coupling only; for stronger coupling the length-scale, on which a local temperature can be defined, increases [46]). The thermal state will be subject to increased fluctuations as the total system size (pertinent state space) decreases, in qualitative agreement with (18.39). A thermostat is, by definition, a very large system making temperature fluctuations virtually absent.

18.6.4 Measurement

The concept of a thermometer is based on a tri-partite quantum system: the system proper, the large enough environment, and the (small) thermometer subsystem (cf. Fig. 18.21). They should all be weakly coupled (coupling constants λ, λ', respectively). In this case temperature becomes a local concept, i.e., the canonical state of system and thermometer generated by the envi-

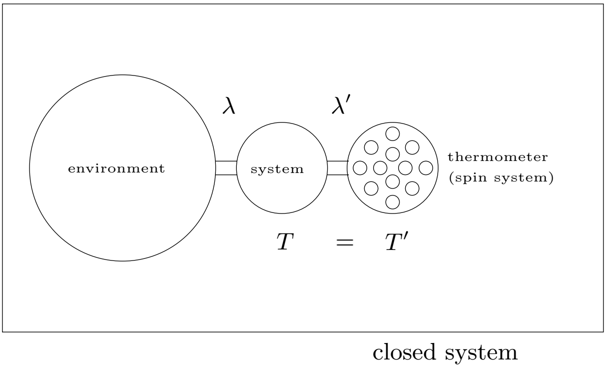


Fig. 18.21. Basic scenario for a temperature measurement, λ, λ' are the respective coupling constants.

ronment implies an occupation of energy states of system and thermometer, respectively, controlled by the same temperature ($T = T'$).

This model even takes care of the intuitive notion that we should be able to decouple the thermometer from the rest of the system before "looking at it". After having established the equilibrium state (i.e., the appropriate entanglement), interactions are no longer required to substain this state. Of course, without coupling to the environment, stability is lost; any new interaction will perturb the state of the thermometer permanently. Such an interaction, though, is needed for measurement, i.e., to implement the "looking-at" process.

We may model the thermometer by a set of non-interacting spins (two-level systems) with finite Zeemann splitting. We can thus perform an ensemble measurement of the average magnetization (mean energy). Knowing that the state is canonical, there is a unique relation between the mean energy and the temperature. This allows us to (indirectly) measure the temperature. (This final measurement step is not included in Fig. 18.21).

We note that the temperature measurement is not a quantum measurement: T is not associated with an operator. We are thus not confronted with the notorious quantum measurement problem (collapse of the wave function). The temperature measurement, when carried out on a "quantum thermometer", is rather equivalent to an ensemble measurement, which is free from those interpretation problems.

We also note that any measurement, whether "classical" or quantum, has to be based on the same aspects of non-equilibrium: a measurement is a process, and processes cease to exist in equilibrium.

18.6.5 Thermodynamic “Uncertainty Relation”

In the classical interpretation of thermodynamics, there is often some confusion about the meaning of thermodynamic control parameters. It appears that it is only for practical reasons that one prefers T over the mean energy E. It has already been conjectured by Bohr that there should be an uncertainty relation between energy and temperature [114]. Any measurement of temperature requires an energy exchange between the system and the thermometer; therefore it should make energy uncertain to some extent.

This conclusion is based on the assumption that the micro states have well-defined energy (and different temperatures) to begin with. This, in turn, is based on the classical ignorance interpretation of non-pure states. Quantum mechanically, there is no problem to accept that for any system in contact with a large environment, the energy is uncertain. Being in a stable thermal state is not a property of the considered system as such, but a consequence of its embedding in a (large) quantum environment. The assumption that temperature is unsharp if energy is sharp and vice versa ($\varDelta E \cdot \varDelta \frac{1}{T} \geq k_{\mathrm{B}}$) [114] contradicts our quantum approach. This uncertainty relation has, to the best of our knowledge, never been tested experimentally. Neither E nor T is an operator, anyway. While there is, in general, an “unexpected effectiveness” of the classical approach to thermodynamics, here is a point where the classical and the quantum approaches differ.

18.7 Quantum Manometer

18.7.1 Eigenspectrum of System and Manometer

In thermodynamics, pressure is an intensive work variable defined as (cf. Chap. 13)

$$p = -\left(\frac{\partial U}{\partial V}\right)_S , \tag{18.40}$$

where U is the internal energy of the system under consideration (not of the total system) and the volume V is the extensive work variable conjugate to p. In the one-dimensional case, p has the dimension of force:

$$F = -\left(\frac{\partial U}{\partial L}\right)_S . \tag{18.41}$$

In any case, the mechanical effect will consist of a deformation (compression) of an external (non-thermodynamic) system, working as a manometer. Here we intend to show that the conjugation between the extensive work variable and the pressure can indeed be confirmed down to the nano limit. For this purpose we first study an isolated bipartite system, i.e., without coupling

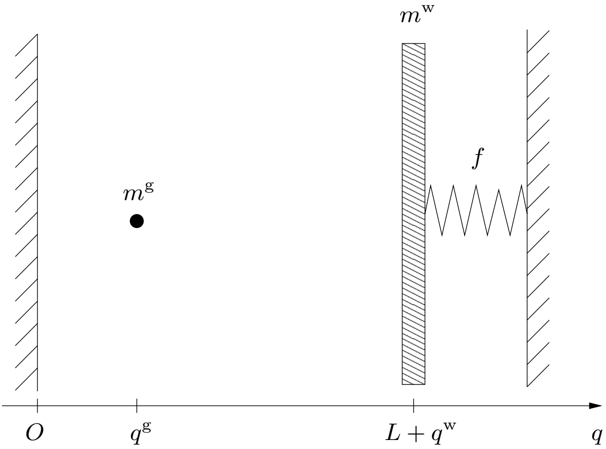


Fig. 18.22. One-dimensional model for a confined "gas particle", m^{g}, interacting with a harmonic oscillator

to an additional environment. The total model is thus supposed to remain "mechanical".

In our present one-dimensional model introduced by Borowski et al. [19] one subsystem is a single particle (g) in a box (cf. Sect. 18.6.2), where, however, one wall (w) is now replaced by a movable piston connected to a spring, the manometer (see Fig. 18.22). The respective two-particle Hamiltonian is given by

$$\hat{H}(q^{\mathrm{g}}, q^{\mathrm{w}}) = -\frac{\hbar^2}{2m^{\mathrm{g}}} \frac{\partial^2}{\partial (q^{\mathrm{g}})^2} - \frac{\hbar^2}{2m^{\mathrm{w}}} \frac{\partial^2}{\partial (q^{\mathrm{w}})^2} + \frac{f}{2} (q^{\mathrm{w}})^2 + \hat{V}(q^{\mathrm{g}}, q^{\mathrm{w}}) \ , \quad (18.42)$$

with the potential

$$\hat{V}(q^{\mathrm{g}}, q^{\mathrm{w}}) = \begin{cases} \infty & \qquad q^{\mathrm{g}} < 0 \\ \infty \quad \text{for} & \qquad q^{\mathrm{g}} < L + q^{\mathrm{w}} \\ 0 & \qquad 0 \leq q^{\mathrm{g}} \leq L + q^{\mathrm{w}} \end{cases} \ . \quad (18.43)$$

The total potential term, which is obviously not accessible from a perturbative approach

$$\hat{V}'(q^{\mathrm{g}}, q^{\mathrm{w}}) = \hat{V}(q^{\mathrm{g}}, q^{\mathrm{w}}) + \frac{f}{2} (q^{\mathrm{w}})^2 \quad (18.44)$$

can be shown to separate under the coordinate transformation

$$y^{\mathrm{g}} = \frac{q^{\mathrm{g}}}{q^{\mathrm{w}} + L} L \ , \quad y^{\mathrm{w}} = q^{\mathrm{w}} \ , \quad (18.45)$$

in the limit of $y_{\mathrm{w}} \ll L$. Of course, in order to make use of this convenient feature, one also has to transform the kinetic parts to the new coordinates.

The unperturbed part (particle of mass m^{g} in a box with infinite walls at $y^{\mathrm{g}} = 0$ and $y^{\mathrm{g}} = L$ and a particle of mass m^{w} in a harmonic potential) is then given by

$$\hat{H}_0 = -\frac{\hbar^2}{2m^{\mathrm{g}}}\frac{\partial^2}{\partial(y^{\mathrm{g}})^2} - \frac{\hbar^2}{2m^{\mathrm{w}}}\frac{\partial^2}{\partial(y^{\mathrm{w}})^2} + \hat{V}'(y^{\mathrm{g}}, y^{\mathrm{w}}) \, . \tag{18.46}$$

While the transformation of the kinetic energies produces an additional coupling term, a more careful analysis shows that this coupling term can now be dealt with by means of standard perturbation theory. The corresponding energy spectrum of $\hat{H}_0$ is

$$E_{j^{\mathrm{g}}j^{\mathrm{w}}} = E_{j^{\mathrm{g}}} + E_{j^{\mathrm{w}}} = \left(\frac{\pi\hbar j^{\mathrm{g}}}{\sqrt{2m^{\mathrm{g}}}L}\right)^2 + \hbar\sqrt{\frac{f}{m^{\mathrm{w}}}}\,(j^{\mathrm{w}} + \frac{1}{2}) \, , \tag{18.47}$$

where j^{g} as well as j^{w} are quantum numbers. Due to the particle-particle interaction one finds for the perturbed energy eigenstate with $j^{\mathrm{w}} = 0$ approximately

$$\langle q^{\mathrm{w}} \rangle_{j^{\mathrm{g}}} = \left(\frac{\pi\hbar j^{\mathrm{g}}}{L\sqrt{m^{\mathrm{g}}Lf}}\right)^2 \, . \tag{18.48}$$

For such an elongation the corresponding force is

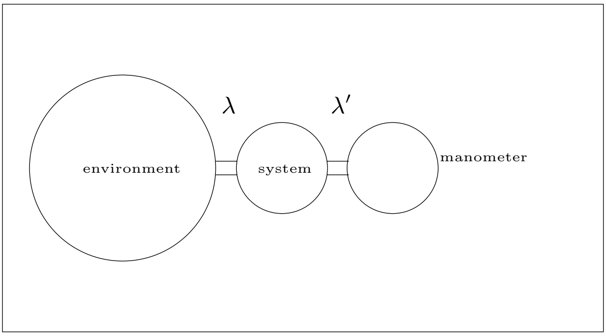
$$F_{j^{\mathrm{g}}} = f\langle q^{\mathrm{w}} \rangle_{j^{\mathrm{g}}} \, , \tag{18.49}$$

which is identical with $F = -\partial E_{j^{\mathrm{g}}}/\partial L$ as postulated. The interaction between manometer and particle in a box can be shown to lead to negligible entanglement, i.e., the local entropy remains zero.

18.7.2 The Total Scenario

The total model we envision is a tri-partite system consisting of the system and manometer part and a large environment (see Fig. 18.23). As discussed a number of times, the interaction with the environment will typically render the system and manometer thermodynamical. Here we assume that the harmonic oscillator (the manometer degree of freedom) actually remains in its ground state. Only the states j^{g} of the system become occupied according to a canonical distribution. This leads to an additional thermal averaging of $\langle q^{\mathrm{w}} \rangle$.

As the environment (its relevant part of Hilbert space) becomes smaller and smaller, the occupation probabilities for the states j^{g} start to show temporal fluctuations. These will lead to temporal fluctuations of $\langle q^{\mathrm{w}} \rangle$ as well and thus of the pressure, just like in the case of temperature (see Sect. 18.6). In so far as for a modular system (e.g., a gas) the other (weakly interacting) particles can play the role of an effective environment for any picked



closed system

Fig. 18.23. Basic scenario for a pressure measurement, λ is the coupling between system and environment, λ' describes the mechanical coupling between system and manometer.

single particle, fluctuations should increase for decreasing volume (at fixed density). This is in qualitative agreement with the standard result of classical thermodynamics [69, 95]

$$\overline{(\Delta p)^2} = -T \left(\frac{\partial p}{\partial V} \right)_S = \frac{k_{\mathrm{B}} T}{\chi_S V} \, , \tag{18.50}$$

where χ_S is the adiabatic compressibility.

For an actual measurement of $\langle q^{\mathrm{w}} \rangle$ we should be aware of the fact that $\langle q^{\mathrm{w}} \rangle$ may also suffer from quantum uncertainty. In the present case of a quantum manometer its variance can easily be of the same order as its mean value! Measurement projections will thus lead to fluctuations, in addition to the "thermal fluctuations" due to the environment.

For a macroscopic manometer the quantum uncertainty can be neglected, however, this limit also restricts the detection of fluctuations.

18.7.3 Classical Picture

The quantum mechanical "particle in a box states" are all delocalized. When weakly coupled to a quantum environment, the system will settle in a canonical state, which should not be far away from a Gaussian state [34]. This "localization" of particles in real space supports the classical view of pressure resulting from impinging classical balls on a wall. Quantum effects due to particle indistinguishability are not included, though. In this picture the measurement of pressure fluctuations can be related to the well known model of Brownian motion. The random collisions of air molecules with a suspended mirror directly allows us to visualize thermal oscillations and even to measure the Boltzmann constant [95]. The theoretical analysis is usually based on the Langevin equation. This phenomenological equation includes damping as well as a stochastic force.

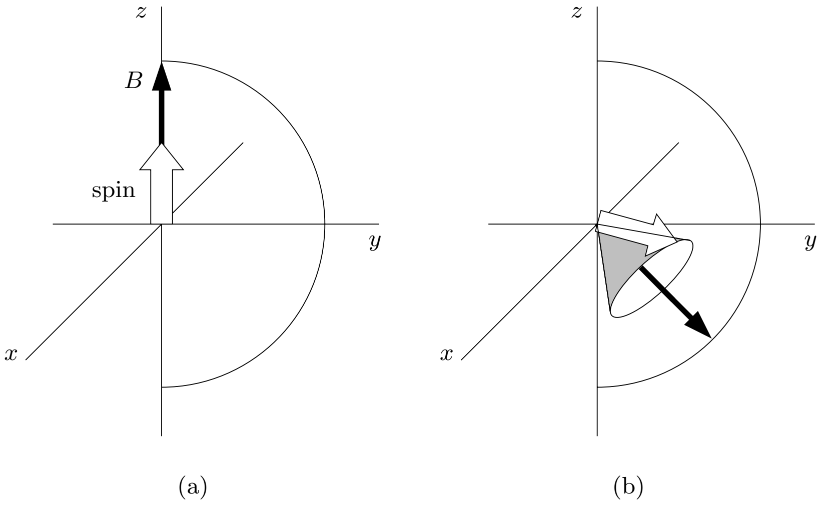
18.8 Adiabatic Following and Adiabatic Process

In Chap. 13 the adiabatic process has been discussed in the context of pressure. Its relationship with the adiabatic following, well known in quantum mechanics, has been clarified. Namely it has been pointed out that, within the theory presented here, "adiabaticity" in the thermodynamical sense is guaranteed only under adiabatic following in the quantum mechanical sense, regardless of the fact that without any coupling to an environment entropy never changes under any kind of process.

The respective adiabatic approximation is a very good technique for slow processes (quasi static), but it gets worse, if the Hamiltonian changes too fast. In the case of an additional microcanonical coupling to an environment, we have found in Chap. 13 that the adiabatic approximation should be valid even for a faster change of the Hamiltonian. Thus, this situation is characterized by a sort of antagonism. A rapidly changing Hamiltonian tends to make the adiabatic approximation fail, while contact with a microcanonical environment stabilizes it.

To investigate these properties, we analyze a model system: a spin in a constant rotating magnetic field. We start with the magnetic field as well as the spin in z-direction (see Fig. 18.24(a)). Then we rotate the magnetic field around the x axis (see Fig. 18.24(b)), whereby the Hamiltonian of the system reads

$$\hat{H} = \cos(\omega t)\,\hat{\sigma}_3 + \sin(\omega t)\,\hat{\sigma}_2 \; . \tag{18.51}$$



(a) (b)

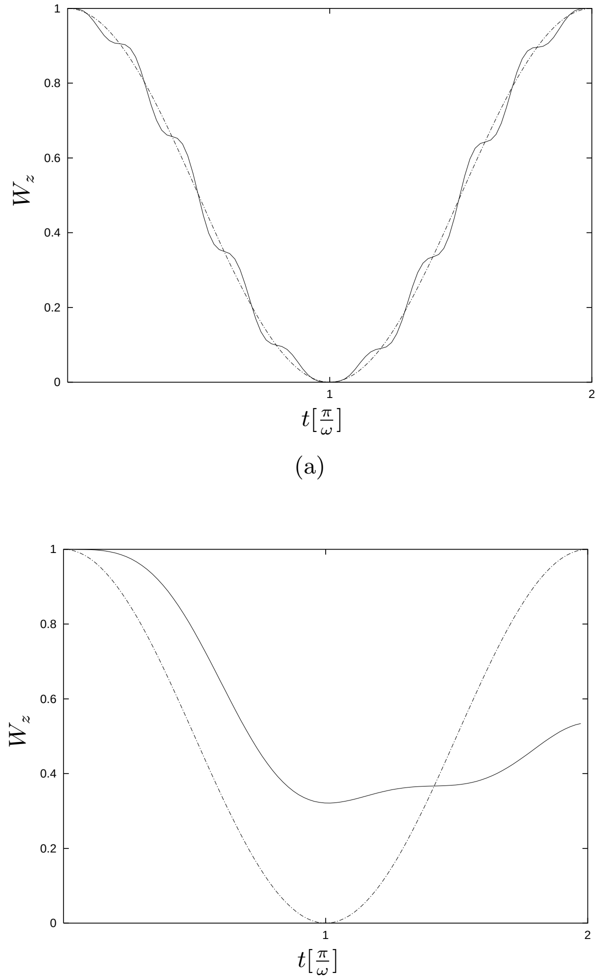
Fig. 18.24. Model system for the investigation of the adiabatic following. The magnetic field of constant modulus is rotated around the x-direction. The spin as well as the magnetic field initially point in z-direction [64].

As the field starts to rotate, the spin will follow the magnetic field for sufficiently small rotational frequencies ω. For larger ones, the spin is not able to follow the field any longer, the adiabatic approximation tends to break down.

Firstly we observe the probability W_z to find the spin in z-direction. For the strict following of the spin this probability will decrease until it is zero in case the magnetic field points in $-z$-direction and the spin should be parallel to the field in this case. This behavior can be confirmed in Fig. 18.25(a), where we rotate the field with a frequency of $\omega = 0.2$. The small oscillations are due to the Lamor frequency of the system (here $\omega_{\mathrm{L}} = 1$). In Fig. 18.25(b) one can find the result in the case of a faster rotation $\omega = 2$ and, as expected, the probability after half of a period of rotation of the field is no longer zero.

As discussed in Sect. 13.1, we are now going to additionally couple an environment to the spin system, observing the system again in the regime of the fast rotational frequency $\omega = 2$. This special microcanonical environment should only suppress the off-diagonal elements of the spin system in the momentary eigenbasis of the spin Hamiltonian, as in (13.15). We have to make sure that the environment remains microcanonical – i.e., only affects the off-diagonal part of the density operator of the system – for the whole evolution. For appropriate coupling strength of the spin system to the environment, defined by the parameters of (13.15), the spin again follows the field, as can be seen in Fig. 18.26.

So far we have used the Lindblad formalism to model the influence of an environmental system (see Sect. 13.1 and especially (13.15)). Now, we introduce a model for the environment, as done in Sect. 18.1. Here, we use a system with only one energy level but a degeneracy N^{c}, and couple our spin system randomly to this environment. We have to make sure that the environment always couples to the momentary eigenbasis of the Hamiltonian and suppresses off-diagonal elements only in this basis. After a half period of rotation we will observe the state of the spin $\hat{\rho}(\pi/\omega)$, and measure the distance of this state from the expected state according to perfect adiabatic following $\hat{\rho}$, using the fidelity $F_{\hat{\rho}(\pi/\omega),\hat{\rho}}$ introduced in Sect. 2.2.3 in (2.21). The result is to be seen in Fig. 18.27, where we show the fidelity over the rotation frequency of the field. The dashed line shows the case without an environment, whereas the solid line is the case with pure dephasing. Obviously, in the latter case one can rotate the field much faster than without an environment and this is exactly what theory predicts. The results depend on the concrete model and on the respective parameters like coupling strength and number of energy levels.



(a)

(b)

Fig. 18.25. Magnetic field rotation without an environment. Probability W_z to find the spin in z-direction. (**a**) For a small rotation frequency, $\omega = 0.2$, the spin follows the magnetic field, (**b**) for $\omega = 2$ this is no longer the case. The dashed line shows the perfect adiabatic following [64].

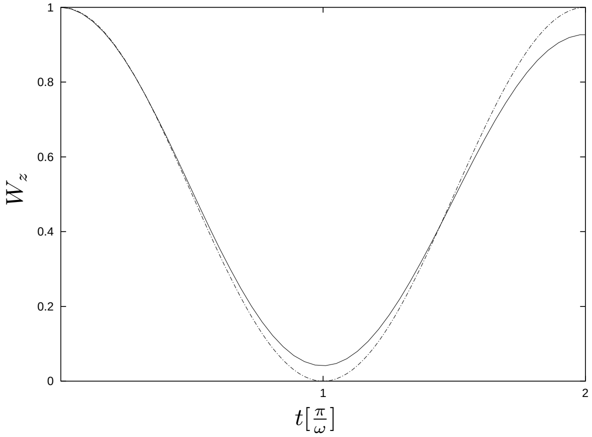


Fig. 18.26. Fast rotation of the magnetic field $\omega = 2$, but now with an additional coupling to an environment, according to (13.15). As can be seen, the spin tends to follow the field, like in the slow rotation regime [64].

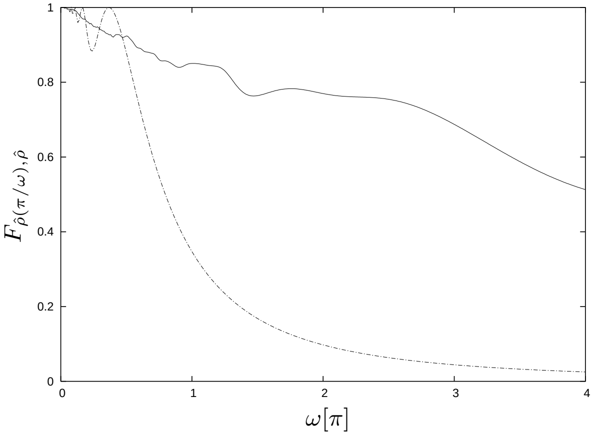


Fig. 18.27. Fidelity $F_{\hat{\rho}(\pi/\omega),\hat{\rho}}$ of the spin state after a half rotation between the state $\hat{\rho}(\pi/\omega)$ and the expected perfect following case $\hat{\rho}$, as a function of ω. Solid line: with dephasing. Dashed line: without dephasing [64].

17 The Route to Equilibrium

> *We call dynamical laws reversible if any process following this dynamics is changed into a possible process by a time reversal. In thermodynamics we know processes which do not possess this property. They are irreversible.*
>
> — F. Schlögl in [114]

The method we are going to present here also relies, like the approach based on Fermi's Golden Rule, on a step-wise solution of the Schrödinger equation, but now taking superpositions of energy eigenstates as initial states into account. Naturally, this approach, like all other approaches, relies on approximations. The crucial approximation here is the replacement of some specific quantities by their Hilbert space averages. For almost all states the Hilbert space average will be a good approximation, if and only if the Hilbert space landscape is almost flat. Therefore, the approximation is supported by showing that it is indeed possible to assume such a flat Hilbert space. Finally, it will be possible to derive a rate equation, which describes the decay qualitatively.

17.1 System and a Large Environment

The energy scheme of the situation we are going to analyze is depicted in Fig. 17.1. A two level system, g, is in contact with a "many level" environment or "container", c. Only the relevant parts of the spectrum of the environment enter the model. These are, in this case, two "bands" of width $\delta\epsilon$, containing N_1^{c} (N_0^{c}) equidistant eigenstates in the upper (lower) band. Therefore the

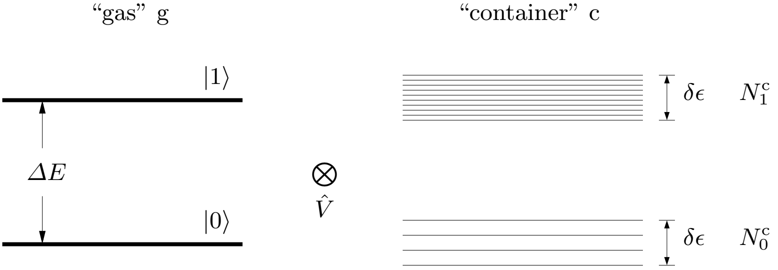


Fig. 17.1. Discrete two-level system coupled canonically to a quasi-continuous container system. This set-up should, for a sufficiently high state density in the container system, and an adequately tuned coupling, exhibit an exponential decay of an excitation in the gas system.

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 169–182 (2004)
`http://www.springerlink.com/`

level spacing within the upper (lower) energy “band” is

$$\Delta E_1^{\mathrm{c}} := \frac{\delta\epsilon}{N_1^{\mathrm{c}}} \quad \text{and} \quad \Delta E_0^{\mathrm{c}} := \frac{\delta\epsilon}{N_0^{\mathrm{c}}} \; . \tag{17.1}$$

In the following, quantities of the “upper band” of the environment get the subscript 1, where as quantities of the “lower band” get the subscript 0. We consider an evolution from an initial state, with the system in the excited state $|1\rangle$ and the environment in the “lower band”, 0. 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 $|0\rangle$ and the environment in its “upper band”, 1. The Hamiltonian within the relevant subspace of the entire Hilbert space may thus be organized as follows,

$$\hat{H} = \left(\begin{array}{ccc|ccc} \ddots & & 0 & & & \\ & \Delta E + i\Delta E_1^{\mathrm{c}} & & & \hat{V} & \\ 0 & & \ddots & & & \\ \hline & & & \ddots & & 0 \\ & \hat{V} & & & j\Delta E_0^{\mathrm{c}} & \\ & & & 0 & & \ddots \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} |\psi^{\mathrm{gr}}\rangle \\ \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} |\psi^{\mathrm{ex}}\rangle \end{array} \tag{17.2}$$

where i (j) count the levels in the upper (lower) “band” of the environment. The Hamiltonian is displayed in the eigenbasis of the uncoupled system, for simplicity we assume for the moment that the coupling $\hat{V}$ only adds terms to the Hamiltonian in the off-diagonal blocks. This corresponds to an energy transfer coupling between system and environment, say canonical conditions.

We now introduce two projectors, which project out the upper (lower) part of the state of the system

$$\hat{P}^{\mathrm{ex}} := |1\rangle\langle 1| \otimes \hat{1}^{(\mathrm{c})} \; , \quad \hat{P}^{\mathrm{gr}} := |0\rangle\langle 0| \otimes \hat{1}^{(\mathrm{c})} \; , \tag{17.3}$$

where $\hat{1}^{(\mathrm{c})}$ is the $\hat{1}$-operator in the environmental system. In the following we call that part of the wave vector that corresponds to the considered system in the excited state $|\psi^{\mathrm{ex}}\rangle$ and the part that corresponds to the system in the ground state $|\psi^{\mathrm{gr}}\rangle$, i.e.,

$$|\psi^{\mathrm{ex}}\rangle := \hat{P}^{\mathrm{ex}}|\psi\rangle \; , \quad |\psi^{\mathrm{gr}}\rangle := \hat{P}^{\mathrm{gr}}|\psi\rangle \quad \Rightarrow \quad |\psi\rangle = |\psi^{\mathrm{ex}}\rangle + |\psi^{\mathrm{gr}}\rangle \; . \tag{17.4}$$

Note that neither $|\psi^{\mathrm{ex}}\rangle$ nor $|\psi^{\mathrm{gr}}\rangle$ are normalized individually.

To analyze this model we first transform to the Dirac or interaction picture (cf. Sect. 2.5.1)

$$\hat{U}(t,0) := \hat{U}_0 := \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}_0 t} \; , \quad |\psi_{\mathrm{I}}\rangle := \hat{U}_0^{\dagger}|\psi\rangle \; , \quad \hat{V}_{\mathrm{I}} := \hat{U}_0^{\dagger}\,\hat{V}\,\hat{U}_0 \; , \tag{17.5}$$

where $\hat{H}_0$ is the Hamiltonian of the uncoupled system. The Schrödinger equation in this representation reads

$$\mathrm{i}\hbar \frac{\partial}{\partial t} |\psi_{\mathrm{I}}\rangle = \hat{V}_{\mathrm{I}} |\psi_{\mathrm{I}}\rangle \; , \tag{17.6}$$

where both states and operators are now time-dependent, i.e., also $\hat{V}_{\mathrm{I}}$ is a time dependent operator, but preserves the off-diagonal block form as before.

The crucial quantities in the context of a decay to equilibrium are the probabilities of finding the system in its exited (ground) state, W^{ex} (W^{gr}). Due to the diagonality of $\hat{H}_0$ those quantities have the same representation in the interaction as well as in the Schrödinger picture,

$$W^{\mathrm{ex}} = \langle \psi_{\mathrm{I}}^{\mathrm{ex}} | \psi_{\mathrm{I}}^{\mathrm{ex}} \rangle = \langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle \; , \qquad W^{\mathrm{gr}} = \langle \psi_{\mathrm{I}}^{\mathrm{gr}} | \psi_{\mathrm{I}}^{\mathrm{gr}} \rangle = \langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle \; . \tag{17.7}$$

For simplicity we omit in the following the interaction picture subscript "I", but all the following considerations refer to this picture.

17.2 Time Evolution

To approximate the evolution of the system for a short time step, we can truncate the corresponding Dyson series (cf. Sect. 2.5.2)

$$|\psi(\tau)\rangle \approx \left(\hat{1} - \frac{\mathrm{i}}{\hbar} \hat{U}_1(\tau) - \frac{1}{\hbar^2} \hat{U}_2(\tau) \right) |\psi(0)\rangle \; . \tag{17.8}$$

This is a truncation of second order, in which the $\hat{U}$'s are the time ordered integrals that occur in the Dyson series [113]

$$\hat{U}_1(\tau) = \int_0^{\tau} \mathrm{d}\tau' \hat{V}(\tau') \; , \quad \hat{U}_2(\tau) = \int_0^{\tau} \mathrm{d}\tau' \hat{V}(\tau') \int_0^{\tau'} \mathrm{d}\tau'' \hat{V}(\tau'') \; . \tag{17.9}$$

According to the Hermiticity of $\hat{V}(\tau)$, $\hat{U}_1(\tau)$ should be Hermitian too, which is not the case for $\hat{U}_2(\tau)$. $\hat{U}_1(\tau)$ has the same off-diagonal form as $\hat{V}(\tau)$ whereas $\hat{U}_2(\tau)$ has here a block diagonal form according to the interaction matrix. (To further simplify notation we do not write the τ dependence of the $\hat{U}$'s explicitly. Furthermore, we omit the time dependence of the wave function, if it refers to the initial state, i.e., $|\psi(0)\rangle := |\psi\rangle$.)

As mentioned above we are interested in the time evolution of the probability of finding the system in its excited state $W^{\mathrm{ex}}(\tau)$, or ground state $W^{\mathrm{gr}}(\tau)$, respectively. Initially we consider $W^{\mathrm{ex}}(\tau)$. Neglecting all terms of higher than second order (products of $\hat{U}_1$ and $\hat{U}_2$ as well as terms proportional to $\hat{U}_2^2$) we get from (17.8)

$$\begin{aligned}W^{\mathrm{ex}}(\tau) &= \langle\psi^{\mathrm{ex}}(\tau)|\psi^{\mathrm{ex}}(\tau)\rangle = \langle\psi(\tau)|\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}|\psi(\tau)\rangle \\ &= \langle\psi|\big(\hat{1}\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}\hat{1} - \frac{\mathrm{i}}{\hbar}\,\hat{1}\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}\hat{U}_1 + \frac{\mathrm{i}}{\hbar}\,\hat{U}_1\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}\hat{1} \\ &\qquad - \frac{1}{\hbar^2}\,\hat{1}\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}\hat{U}_2 + \frac{1}{\hbar^2}\,\hat{U}_1\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}\hat{U}_1 - \frac{1}{\hbar^2}\,\hat{U}_2^{\dagger}\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}}\hat{1}\big)|\psi\rangle\ .\end{aligned} \tag{17.10}$$

According to the special off-diagonal block form of the interaction, we find for the operator products

$$\hat{U}_1\hat{P}^{\mathrm{ex}} = \hat{P}^{\mathrm{gr}}\hat{U}_1 \quad \text{and} \quad \hat{U}_2\hat{P}^{\mathrm{ex}} = \hat{P}^{\mathrm{ex}}\hat{U}_2\,, \quad \hat{U}_2^{\dagger}\hat{P}^{\mathrm{ex}} = \hat{P}^{\mathrm{ex}}\hat{U}_2^{\dagger}\ , \tag{17.11}$$

and thus

$$\begin{aligned}W^{\mathrm{ex}}(\tau) = \langle\psi|\big(&\hat{P}^{\mathrm{ex}}\hat{P}^{\mathrm{ex}} - \frac{\mathrm{i}}{\hbar}\,\hat{P}^{\mathrm{ex}}\hat{U}_1\hat{P}^{\mathrm{gr}} + \frac{\mathrm{i}}{\hbar}\,\hat{P}^{\mathrm{gr}}\hat{U}_1\hat{P}^{\mathrm{ex}} + \frac{1}{\hbar^2}\,\hat{P}^{\mathrm{gr}}\hat{U}_1\hat{U}_1\hat{P}^{\mathrm{gr}} \\ &- \frac{1}{\hbar^2}\,\hat{P}^{\mathrm{ex}}\hat{U}_2\hat{P}^{\mathrm{ex}} - \frac{1}{\hbar^2}\,\hat{P}^{\mathrm{ex}}\hat{U}_2^{\dagger}\hat{P}^{\mathrm{ex}}\big)|\psi\rangle\ .\end{aligned} \tag{17.12}$$

The probability $W^{\mathrm{gr}}(\tau)$ is obtained in an analogous way. Using (17.4) we obtain for the time evolution of the probabilities $W^{\mathrm{ex}}(\tau)$ and $W^{\mathrm{gr}}(\tau)$

$$\begin{aligned}W^{\mathrm{ex}}(\tau) &= \langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle + \frac{\mathrm{i}}{\hbar}\,\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle - \frac{\mathrm{i}}{\hbar}\,\langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle \\ &\quad + \frac{1}{\hbar^2}\,\langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle - \frac{1}{\hbar^2}\,\langle\psi^{\mathrm{ex}}|(\hat{U}_2 + \hat{U}_2^{\dagger})|\psi^{\mathrm{ex}}\rangle \qquad (17.13) \\ W^{\mathrm{gr}}(\tau) &= \langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle + \frac{\mathrm{i}}{\hbar}\,\langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle - \frac{\mathrm{i}}{\hbar}\,\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle \\ &\quad + \frac{1}{\hbar^2}\,\langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle - \frac{1}{\hbar^2}\,\langle\psi^{\mathrm{gr}}|(\hat{U}_2 + \hat{U}_2^{\dagger})|\psi^{\mathrm{gr}}\rangle\ .\end{aligned} \tag{17.14}$$

The strict overall probability conservation requires

$$\begin{aligned}W^{\mathrm{ex}}(\tau) + W^{\mathrm{gr}}(\tau) &= \langle\psi^{\mathrm{ex}}(\tau)|\psi^{\mathrm{ex}}(\tau)\rangle + \langle\psi^{\mathrm{gr}}(\tau)|\psi^{\mathrm{gr}}(\tau)\rangle \\ &= \langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle + \langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle = 1\ .\end{aligned} \tag{17.15}$$

Since the normalization is already fulfilled in the zero order, all higher orders must vanish. Obviously the first order vanishes automatically. Thus, exploiting (17.15), for the second order of the sum of (17.13) and (17.14) we find

$$\langle\psi^{\mathrm{ex}}|(\hat{U}_2 + \hat{U}_2^{\dagger})|\psi^{\mathrm{ex}}\rangle = \langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle\,, \tag{17.16}$$

$$\langle\psi^{\mathrm{gr}}|(\hat{U}_2 + \hat{U}_2^{\dagger})|\psi^{\mathrm{gr}}\rangle = \langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle\ . \tag{17.17}$$

Inserting this into (17.13) and (17.14) yields

$$\begin{aligned}W^{\mathrm{ex}}(\tau) = \langle\psi^{\mathrm{ex}}(\tau)|\psi^{\mathrm{ex}}(\tau)\rangle =&\langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle - \frac{\mathrm{i}}{\hbar}\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle + \frac{\mathrm{i}}{\hbar}\langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle \\ &+ \frac{1}{\hbar^2}\langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle - \frac{1}{\hbar^2}\langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle\end{aligned} \tag{17.18}$$

$$\begin{aligned}W^{\mathrm{gr}}(\tau) = \langle\psi^{\mathrm{gr}}(\tau)|\psi^{\mathrm{gr}}(\tau)\rangle =&\langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle - \frac{\mathrm{i}}{\hbar}\langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle + \frac{\mathrm{i}}{\hbar}\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle \\ &+ \frac{1}{\hbar^2}\langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle - \frac{1}{\hbar^2}\langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle \;.\end{aligned} \tag{17.19}$$

For an exact evaluation of the right hand side one would need to know the $|\psi^{\mathrm{ex}}\rangle, |\psi^{\mathrm{gr}}\rangle$ in detail. It could, however, be the case that the right hand sides of (17.18) and (17.19) do not depend significantly on $|\psi^{\mathrm{ex}}\rangle, |\psi^{\mathrm{gr}}\rangle$, i.e., they may take on almost the same value for almost all possible $|\psi^{\mathrm{ex}}\rangle, |\psi^{\mathrm{gr}}\rangle$. This would be the case if the landscape defined by the right hand side, over the region in Hilbert space consistent with given $W^{\mathrm{ex}}(0)$, $W^{\mathrm{gr}}(0)$, were essentially flat. Whether or not this is indeed the case can only be decided by calculating the Hilbert space variances $\Delta_{\mathrm{H}}(\langle\psi^{\mathrm{ex}}(\tau)|\psi^{\mathrm{ex}}(\tau)\rangle), \Delta_{\mathrm{H}}(\langle\psi^{\mathrm{gr}}(\tau)|\psi^{\mathrm{gr}}(\tau)\rangle)$. If these are small, the right hand side of (17.18) and (17.19) could be replaced by its Hilbert space average $[\![\langle\psi^{\mathrm{ex}}(\tau)|\psi^{\mathrm{ex}}(\tau)\rangle]\!], [\![\langle\psi^{\mathrm{gr}}(\tau)|\psi^{\mathrm{gr}}(\tau)\rangle]\!]$ as a valid approximation. At the moment we will proceed to do so and come back to the Hilbert space variances later to justify this replacement.

17.3 Hilbert Space Average

We introduce the abbreviations,

$$\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle := \alpha \;, \qquad \langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle := \alpha^* \;, \tag{17.20}$$

$$\langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle := \kappa^{\mathrm{gr}} \;, \qquad \langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle := \kappa^{\mathrm{ex}} \tag{17.21}$$

and evaluate the Hilbert space average of these quantities: $[\![\kappa^{\mathrm{gr}}]\!]$, $[\![\kappa^{\mathrm{ex}}]\!]$ and $[\![\alpha]\!]$. The detailed calculation of these quantities is discussed in App. C. The general investigation of integrals in high dimensional spaces constricted to hyperspheres can be found in App. A. Here we just give the results and point out their plausibility.

From (C.23) we find

$$[\![\alpha]\!] = [\![\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle]\!] = 0 \quad \text{and} \quad [\![\alpha^*]\!] = [\![\langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle]\!] = 0 \;. \tag{17.22}$$

Since $|\psi^{\mathrm{ex}}\rangle$ and $|\psi^{\mathrm{gr}}\rangle$ lie on different, i.e., entirely independent hyperspheres, such Hilbert space averages vanish. This is due to the special structure of the respective integrals over the Hilbert space.

For the other two Hilbert space averages we find from (C.24) and (C.25)

$$[\![\kappa^{\mathrm{ex}}]\!] = [\![\langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle]\!] = \frac{\langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle}{N_0^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{ex}}\{\hat{U}_1^2\}\;, \tag{17.23}$$

$$[\![\kappa^{\mathrm{gr}}]\!] = [\![\langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle]\!] = \frac{\langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle}{N_1^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{gr}}\{\hat{U}_1^2\}\;. \tag{17.24}$$

Here $\mathrm{Tr}_{\mathrm{ex(gr)}}\{\dots\}$ is supposed to denote the trace over the upper (lower) subspace of the operator. This can be understood by thinking of the average as 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 (see App. C).

Plugging these results into (17.18) and (17.19), we get

$$W^{\mathrm{ex}}(\tau) = W^{\mathrm{ex}}(0) + \frac{W^{\mathrm{gr}}(0)}{\hbar^2 N_1^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{gr}}\{\hat{U}_1^2\} - \frac{W^{\mathrm{ex}}(0)}{\hbar^2 N_0^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{ex}}\{\hat{U}_1^2\}\;, \tag{17.25}$$

$$W^{\mathrm{gr}}(\tau) = W^{\mathrm{gr}}(0) + \frac{W^{\mathrm{ex}}(0)}{\hbar^2 N_0^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{ex}}\{\hat{U}_1^2\} - \frac{W^{\mathrm{gr}}(0)}{\hbar^2 N_1^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{gr}}\{\hat{U}_1^2\}\;. \tag{17.26}$$

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

$$\mathrm{Tr}_{\mathrm{ex}}\{\hat{U}_1^2\} = \sum_{j=1}^{N_0^{\mathrm{c}}}\langle j|\hat{U}_1^2|j\rangle = \sum_{j=1}^{N_0^{\mathrm{c}}}\left|\hat{U}_1|j\rangle\right|^2\;. \tag{17.27}$$

Here j runs over the eigenstates of $\hat{H}_0$ in the upper subspace (note this corresponds to the lower “band” of the environment). The object that is summed over here is evaluated in the literature in the context of Fermi’s Golden Rule (see Sect. 2.5),

$$\left|\hat{U}_1|j\rangle\right|^2 = \sum_{i=1}^{N_1^{\mathrm{c}}}\left|\langle i|\hat{V}|j\rangle\right|^2\frac{4\sin^2(\frac{1}{2}\,\omega_{i,j}\,\tau)}{\omega_{i,j}^2} \tag{17.28}$$

with

$$\omega_{i,j} = \frac{1}{\hbar}\left(E_j - E_i\right) = \frac{1}{\hbar}\left(j\Delta E_0^{\mathrm{c}} - i\Delta E_1^{\mathrm{c}}\right)\;, \tag{17.29}$$

see Fig. 17.3(a).

17.4 Short Time Step Equation

Our arguments, including the conditions we have to impose on the model, now follow closely the ones brought forth in the context of Fermi’s Golden Rule.

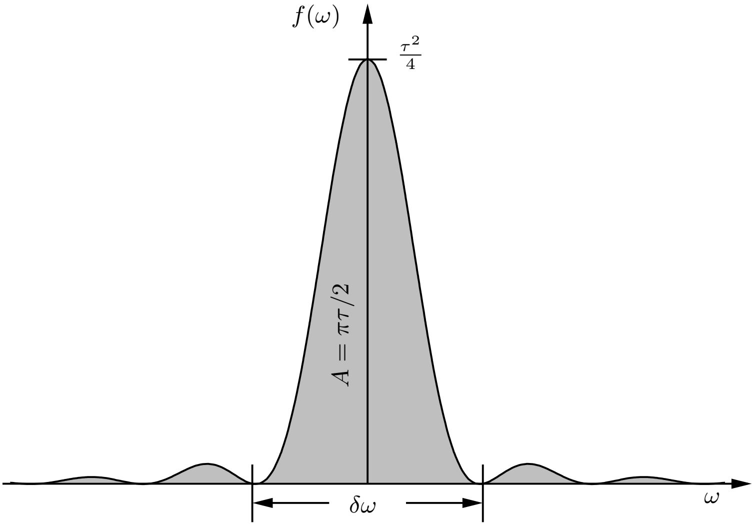


Fig. 17.2. The function $f(\omega)$ defined in (17.31).

The summation in (17.28) consists of two different terms: the transition probability or elements of the interaction matrix and a weight $f(\omega)$. The displacement of different $\omega_{i,j}$ is given by

$$\Delta\omega = \omega_{i,j} - \omega_{i+1,j} = \frac{\Delta E_1^{\mathrm{c}}}{\hbar} = \frac{\delta\epsilon}{N_1^{\mathrm{c}}\hbar} , \tag{17.30}$$

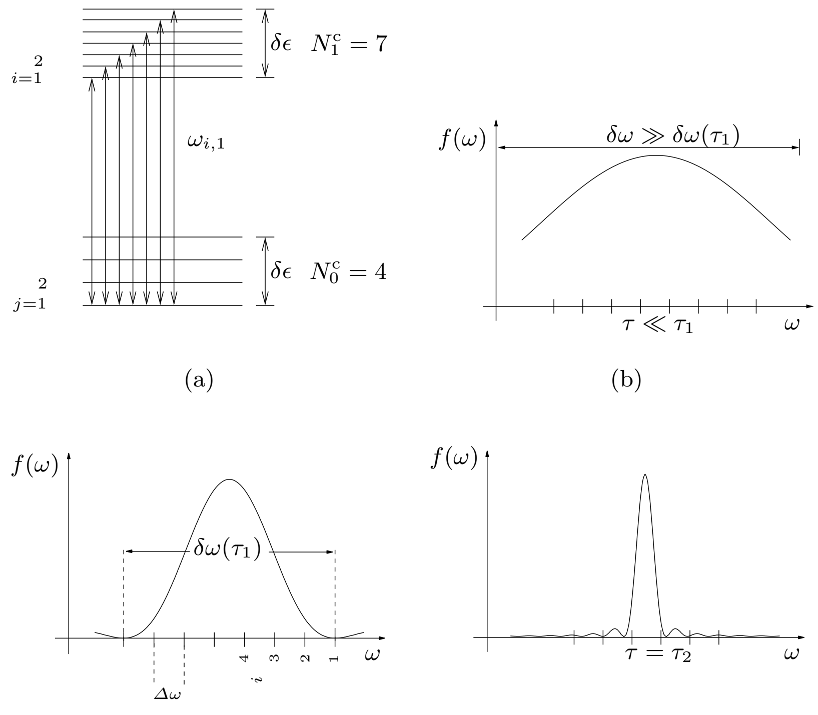
where we have used (17.1). The function

$$f(\omega) = \frac{\sin^2(\frac{1}{2}\omega\tau)}{\omega^2} \tag{17.31}$$

is basically a peak at $\omega = 0$, with the width $\delta\omega = 4\pi/\tau$ and a height of $f(0) = \tau^2/4$. The area under the function $f(\omega)$ is $A = \pi\tau/2$ (see Fig. 17.2). This means the peak gets higher and narrower as τ increases (see Fig. 17.3).

The height of the peak grows with the square of the time τ, the area under f only linearly with τ. One could thus expect two different behaviors: the square and the linear regimes. At the very beginning, the peak is very broad and therefore much broader than the "band" width $\delta\epsilon$ divided by $\hbar$. In this case, we expect that the sum grows with the square of τ, because all terms are near the maximum of the peak (see Fig. 17.3(b)). We choose some τ_1 such that the width $\delta\omega(\tau_1)$ of $f(\omega)$ has approximately the same value as the "band" width $\delta\epsilon$ divided by $\hbar$

$$\delta\omega(\tau_1) = \frac{4\pi}{\tau_1} \approx \frac{\delta\epsilon}{\hbar} \quad \Rightarrow \quad \tau_1 = \frac{4\pi\hbar}{\delta\epsilon} . \tag{17.32}$$



(a) (b)

(c) (d)

Fig. 17.3. Summation of transitions in (17.28). (**a**) Matrix elements to be summed up. (**b**) $\tau \ll \tau_1$: almost all terms are around the maximum of the peak (square regime). (**c**) $\tau \approx \tau_1$: the terms are distributed over the whole peak (linear regime). (**d**) $\tau \approx \tau_2$: only a few terms are within the peak (break-down of the approximation).

The terms are distributed over the whole width of the peak and we expect that the sum grows proportional to the area under the peak, thus linearly in τ (see Fig. 17.3(c)) . In this case and if, furthermore, the function f does not change much over many summation steps $\Delta\omega$, i.e., if

$$\Delta\omega = \frac{\delta\epsilon}{N_1^c \hbar} \ll \delta\omega(\tau_1) = \frac{\delta\epsilon}{\hbar} \quad \Rightarrow \quad N_1^c \gg 1 \,, \tag{17.33}$$

the summation averages out the different elements of the $\hat{V}$-matrix in (17.28). Therefore the sum may be approximated by the average of the interaction matrix element λ_0^2 times the integral over $f(\omega)$ according to ω. The average of the interaction matrix element is

$$\lambda_0^2 = \frac{1}{N_1^{\mathrm{c}} N_0^{\mathrm{c}}} \sum_{i=1}^{N_1^{\mathrm{c}}} \sum_{j=1}^{N_0^{\mathrm{c}}} \left| \langle i | \hat{V} | j \rangle \right|^2 = \frac{1}{2 N_1^{\mathrm{c}} N_0^{\mathrm{c}}} \mathrm{Tr}\{\hat{V}^2\} . \tag{17.34}$$

For (17.28) we then get

$$\left| \hat{U}_1 | j \rangle \right|^2 \approx \lambda_0^2 \int \frac{\mathrm{d}\omega}{\Delta\omega} 4 f(\omega) = \frac{\lambda_0^2 4 A}{\Delta\omega} = \frac{2\pi \lambda_0^2 \hbar N_1^{\mathrm{c}} \tau}{\delta\epsilon} , \tag{17.35}$$

where we have used that the area under $f(\omega)$ is $A = \pi\tau/2$, as mentioned before.

The approximation done so far breaks down later at some time τ_2, when the peak gets too narrow (see Fig. 17.3(d)), i.e., the width is smaller than the summation displacement $\Delta\omega$

$$\delta\omega(\tau_2) = \frac{4\pi}{\tau_2} = \Delta\omega = \frac{\delta\epsilon}{N_1^{\mathrm{c}} \hbar} \quad \Rightarrow \quad \tau_2 = \frac{4\pi \hbar N_1^{\mathrm{c}}}{\delta\epsilon} . \tag{17.36}$$

Thus (17.35) is a valid approximation only for $\tau_1 < \tau < \tau_2$.

Hence, plugging (17.35) into (17.27) yields

$$\mathrm{Tr}_{\mathrm{ex}} \left\{ \hat{U}_1^2 \right\} = \sum_{j=1}^{N_0^{\mathrm{c}}} \left| \hat{U}_1 | j \rangle \right|^2 \approx \frac{2\pi \lambda_0^2 \hbar N_1^{\mathrm{c}} N_0^{\mathrm{c}} \tau}{\delta\epsilon} . \tag{17.37}$$

Since this expression is symmetric under exchange of the upper and lower subspaces, the corresponding expression for the lower subspace reads

$$\mathrm{Tr}_{\mathrm{gr}} \left\{ \hat{U}_1^2 \right\} = \sum_{i=1}^{N_1^{\mathrm{c}}} \left| \hat{U}_1 | i \rangle \right|^2 \approx \frac{2\pi \lambda_0^2 \hbar N_1^{\mathrm{c}} N_0^{\mathrm{c}} \tau}{\delta\epsilon} . \tag{17.38}$$

Inserting (17.37) and (17.38) into (17.25) and (17.26) yields

$$W^{\mathrm{ex}}(\tau) = W^{\mathrm{ex}}(0) + C\tau N_0^{\mathrm{c}} W^{\mathrm{gr}}(0) - C\tau N_1^{\mathrm{c}} W^{\mathrm{ex}}(0) , \tag{17.39}$$

$$W^{\mathrm{gr}}(\tau) = W^{\mathrm{gr}}(0) + C\tau N_1^{\mathrm{c}} W^{\mathrm{ex}}(0) - C\tau N_0^{\mathrm{c}} W^{\mathrm{gr}}(0) , \tag{17.40}$$

where we have abbreviated

$$\frac{2\pi \lambda_0^2}{\delta\epsilon\, \hbar} := C . \tag{17.41}$$

Equations (17.39) and (17.40) describe, within the discussed limits, a short time step starting from any initial state, not necessarily an eigenstate of $\hat{H}$. Since they directly connect the probabilities $W^{\mathrm{ex}}(0)$, $W^{\mathrm{gr}}(0)$ of the initial state with those of the state reached after time τ, we can now iterate these equations under some specific conditions.

17.5 Derivation of a Rate Equation

Before iterating the above equations (17.39) and (17.40), one should again check the pre-conditions for the short time step equation derived so far. 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. Therefore we must check that, e.g.,

$$C\tau N_0^{\mathrm{c}}|_{\tau=\tau_1} = 8\pi^2 \frac{\lambda_0^2}{(\Delta E_0^{\mathrm{c}})^2} \frac{1}{N_0^{\mathrm{c}}} \ll 1 \, , \tag{17.42}$$

where we have used (17.1). In complete analogy we get for the other term of second order

$$C\tau N_1^{\mathrm{c}}|_{\tau=\tau_1} = 8\pi^2 \frac{\lambda_0^2}{(\Delta E_1^{\mathrm{c}})^2} \frac{1}{N_1^{\mathrm{c}}} \ll 1 \, . \tag{17.43}$$

If these two conditions are fulfilled the “linear regime” is reached while the truncation to second order is still a valid description, and we can iterate (17.39) and (17.40) after some time $\tau > \tau_1$. Obviously the linear regime is reached faster the more levels the environment contains.

However, if we want to use the above scheme (17.39) and (17.40) we should make sure that we iterate before the linear regime is left again, i.e., before τ_2. Therefore we must consider the second order terms at τ_2 (17.36) compared to one. Note that τ_2 differs for the two terms of second order, in (17.36) we only argued for one of the two energy “bands” in the environment. Thus, the case for which iterating (17.39) and (17.40) is the best description we can possibly get is

$$C\tau N_0^{\mathrm{c}}|_{\tau=\tau_2} = 8\pi^2 \frac{\lambda_0^2}{(\Delta E_0^{\mathrm{c}})^2} \geq 1 \, , \tag{17.44}$$

$$C\tau N_1^{\mathrm{c}}|_{\tau=\tau_2} = 8\pi^2 \frac{\lambda_0^2}{(\Delta E_1^{\mathrm{c}})^2} \geq 1 \, . \tag{17.45}$$

In this case we get iterating (17.39) and (17.40)

$$\frac{W^{\mathrm{ex}}((i+1)\tau) - W^{\mathrm{ex}}(i\tau)}{\tau} = CN_0^{\mathrm{c}} W^{\mathrm{gr}}(i\tau) - CN_1^{\mathrm{c}} W^{\mathrm{ex}}(i\tau) \, , \tag{17.46}$$

$$\frac{W^{\mathrm{gr}}((i+1)\tau) - W^{\mathrm{gr}}(i\tau)}{\tau} = CN_1^{\mathrm{c}} W^{\mathrm{ex}}(i\tau) - CN_0^{\mathrm{c}} W^{\mathrm{gr}}(i\tau) \, . \tag{17.47}$$

Or, in the limit of τ being extremely small

$$\frac{\mathrm{d}W^{\mathrm{ex}}}{\mathrm{d}t} = CN_0^{\mathrm{c}} W^{\mathrm{gr}} - CN_1^{\mathrm{c}} W^{\mathrm{ex}} \, , \tag{17.48}$$

$$\frac{\mathrm{d}W^{\mathrm{gr}}}{\mathrm{d}t} = CN_1^{\mathrm{c}} W^{\mathrm{ex}} - CN_0^{\mathrm{c}} W^{\mathrm{gr}} \, . \tag{17.49}$$

This evolution equation for the probabilities obviously conserves the overall probability. We have obtained a rate equation for the probabilities of finding the system in the upper and lower levels, respectively.

17.6 Solution of the Rate Equation

The solutions of the equations (17.48) and (17.49) describe 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 (see Fig. 17.4)

$$W^{\mathrm{ex}}(t) = \frac{N_0^{\mathrm{c}}}{N_1^{\mathrm{c}} + N_0^{\mathrm{c}}} + \frac{N_1^{\mathrm{c}}}{N_0^{\mathrm{c}} + N_1^{\mathrm{c}}}\, \mathrm{e}^{-C(N_0^{\mathrm{c}}+N_1^{\mathrm{c}})t} \,, \tag{17.50}$$

$$W^{\mathrm{gr}}(t) = \frac{N_1^{\mathrm{c}}}{N_0^{\mathrm{c}} + N_1^{\mathrm{c}}} \left(1 - \mathrm{e}^{-C(N_0^{\mathrm{c}}+N_1^{\mathrm{c}})t}\right) \,. \tag{17.51}$$

The equilibrium values reached after very long times are

$$W^{\mathrm{ex}}(\infty) = \frac{N_0^{\mathrm{c}}}{N_1^{\mathrm{c}} + N_0^{\mathrm{c}}} \,, \quad W^{\mathrm{gr}}(\infty) = \frac{N_1^{\mathrm{c}}}{N_0^{\mathrm{c}} + N_1^{\mathrm{c}}} \,, \tag{17.52}$$

which are obviously exactly the same as the ones derived in Sect. 9.2.4 (9.42) for the equilibrium state of a system with an energy exchange coupling to the environment.

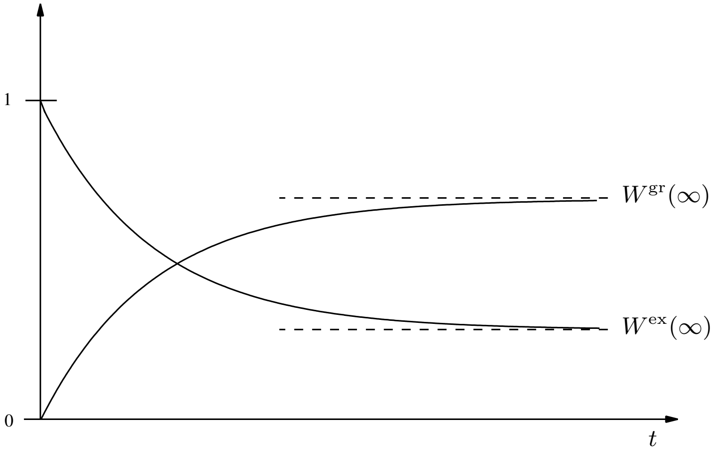


Fig. 17.4. Exponential decay into the equilibrium state, according to (17.50) and (17.51).

17.7 Hilbert Space Variances

Before we compare these results with some numerical data, we want to come back to the Hilbert space variances mentioned above. We have substituted the actual states by their Hilbert space average, an approximation which is only possible, if the landscape is sufficiently flat. This means that the Hilbert space variances must be very small. There are basically three variances to consider. The first one refers to the linear part of (17.18) and (17.19)

$$\begin{aligned}\Delta^2_{\mathrm{H}}(-\mathrm{i}(\alpha-\alpha^*)) &= [\![-(\alpha-\alpha^*)^2]\!] - [\![(\alpha-\alpha^*)]\!]^2 \\ &= [\![-\alpha^2-(\alpha^*)^2+2\alpha\alpha^*]\!] - [\![(\alpha-\alpha^*)]\!]^2 \,. \end{aligned} \tag{17.53}$$

The averages of α^2, $(\alpha^*)^2$ vanish (see (C.29)), so do, as already mentioned, the averages of α, α^* (see (C.23)). Thus it remains

$$\begin{aligned}\Delta^2_{\mathrm{H}}(\mathrm{i}(\alpha-\alpha^*)) &= 2\,[\![\alpha\alpha^*]\!] \\ &= 2\,[\![\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle\langle\psi^{\mathrm{ex}}|\hat{U}_1|\psi^{\mathrm{gr}}\rangle]\!] \\ &= 2\,\frac{\langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle\langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle}{N_0^{\mathrm{c}}N_1^{\mathrm{c}}}\,\mathrm{Tr}_{\mathrm{gr}}\{\hat{U}_1^2\} \\ &= 2\,\frac{W^{\mathrm{ex}}W^{\mathrm{gr}}}{N_0^{\mathrm{c}}N_1^{\mathrm{c}}}\mathrm{Tr}_{\mathrm{gr}}\{\hat{U}_1^2\}\,. \end{aligned} \tag{17.54}$$

The complete evaluation can be found in App. C.3, especially (C.33). Plugging in (17.38) we find in the linear regime

$$\Delta^2_{\mathrm{H}}(\mathrm{i}(\alpha-\alpha^*)) = \frac{4\pi\lambda_0^2\hbar W^{\mathrm{ex}}W^{\mathrm{gr}}\tau}{\delta\epsilon}\,. \tag{17.55}$$

If we iterate at τ_1, which is a reasonable iteration scheme, since it keeps the error caused by the truncation as small as possible, we have to evaluate the Hilbert space variance at $\tau=\tau_1$,

$$\Delta^2_{\mathrm{H}}(\mathrm{i}(\alpha-\alpha^*))\big|_{\tau=\tau_1} = \frac{\lambda_0^2\hbar^2 W^{\mathrm{ex}}W^{\mathrm{gr}}}{\delta\epsilon^2}\,. \tag{17.56}$$

Since $\delta\epsilon = N_1^{\mathrm{c}}\Delta E_1^{\mathrm{c}} = \Delta E_0^{\mathrm{c}}N_0^{\mathrm{c}}$, this Hilbert space average is small in the same limit for which the truncation scheme applies at all, (17.42) and (17.43), and gets even smaller with a growing number of levels in the environment.

The other variances that require consideration are

$$\Delta^2_{\mathrm{H}}(\kappa^{\mathrm{gr}}) = [\![(\kappa^{\mathrm{gr}})^2]\!] - [\![\kappa^{\mathrm{gr}}]\!]^2 \quad \text{and} \quad \Delta^2_{\mathrm{H}}(\kappa^{\mathrm{ex}}) = [\![(\kappa^{\mathrm{ex}})^2]\!] - [\![\kappa^{\mathrm{ex}}]\!]^2\,. \tag{17.57}$$

These are the Hilbert space averages of expectation values of Hermitian operators. For any such object, we evaluate the Hilbert space variance in App. C.1, where it is found especially in (C.34) and (C.35)

$$\Delta_{\mathrm{H}}^2(\kappa^{\mathrm{gr}}) = \frac{1}{N_1^{\mathrm{c}}+1}\left(\frac{\mathrm{Tr}\{\hat{U}_1^4\}}{N_1^{\mathrm{c}}} - \left(\frac{\mathrm{Tr}\{\hat{U}_1^2\}}{N_1^{\mathrm{c}}}\right)^2\right), \tag{17.58}$$

$$\Delta_{\mathrm{H}}^2(\kappa^{\mathrm{ex}}) = \frac{1}{N_0^{\mathrm{c}}+1}\left(\frac{\mathrm{Tr}\{\hat{U}_1^4\}}{N_0^{\mathrm{c}}} - \left(\frac{\mathrm{Tr}\{\hat{U}_1^2\}}{N_0^{\mathrm{c}}}\right)^2\right). \tag{17.59}$$

This is obviously the square of the spectral variance divided by the respective dimension of the Hilbert space compartment. Since the spectral variance of $\hat{U}_1^2$ constitutes a rather lengthy and not really instructive formula, only a rough estimate of the spectral variance shall be given in the following.

The diagonal elements of $\hat{U}_1^2$ are given by (17.28) and are all the same in the considered limit. This follows from the "regularity" of the interaction matrix elements, and holds at times for which $\delta\omega(\tau) \ll \delta\epsilon/\hbar$, except for those within a region of $\delta\omega(\tau)$ at the edges of the spectrum. From a detailed analysis one finds that the matrix elements of $\hat{U}_1^2$ involve sums of the form $\sum_j \langle i|\hat{V}|j\rangle\langle j|\hat{V}|k\rangle$. Since for off-diagonal elements ($i \neq k$) the addends of those sums typically do not have any phase correlations, those sums and thus the off-diagonal elements tend to vanish. So, in a rough approximation, $\hat{U}_1^2$ has no off-diagonal elements and the diagonal elements are all the same except for a few at the edges of the spectrum. Thus $\hat{U}_1^2$ will have a small spectral variance which is independent of N_0^{c}, N_1^{c}. Therefore, the square of the Hilbert space variance vanishes according to (17.58) with the dimension of the corresponding space $1/N_0^{\mathrm{c}}$, $1/N_1^{\mathrm{c}}$. This estimate is in accord with the results for the fluctuations in equilibrium, which are given in Sect. 9.4. It is obvious that the fluctuations during the relaxation period must eventually approach the equilibrium fluctuations.

We thus finally conclude that one can find an exponential decay into the equilibrium defined in Chap. 9, within the above mentioned boundaries. This exponential decay appears in quantum mechanics directly based on Schrödinger dynamics. Therefore, a crucial step towards non-equilibrium phenomena is done. In the next section we show that this exponential decay can, indeed, be observed.

17.8 Numerical Results for the Relaxation Period

To check the validity of the theory developed in the previous sections, a model of the type depicted in Fig. 17.1, with a Hamiltonian as described in (17.2) has been analyzed numerically by directly solving the Schrödinger equation. The interaction matrix $\hat{V}$ has been filled with random Gaussian distributed entries such that

$$\frac{\lambda_0^2}{(\Delta E_1^{\mathrm{c}})^2} \approx 1, \tag{17.60}$$

to ensure that (17.42)-(17.45) are fulfilled. Different container sizes have been analyzed, corresponding to $N_1^{\mathrm{c}} = 50$, 100, 200, 400, 800 and $N_0^{\mathrm{c}} = \frac{1}{2}N_1^{\mathrm{c}}$. For

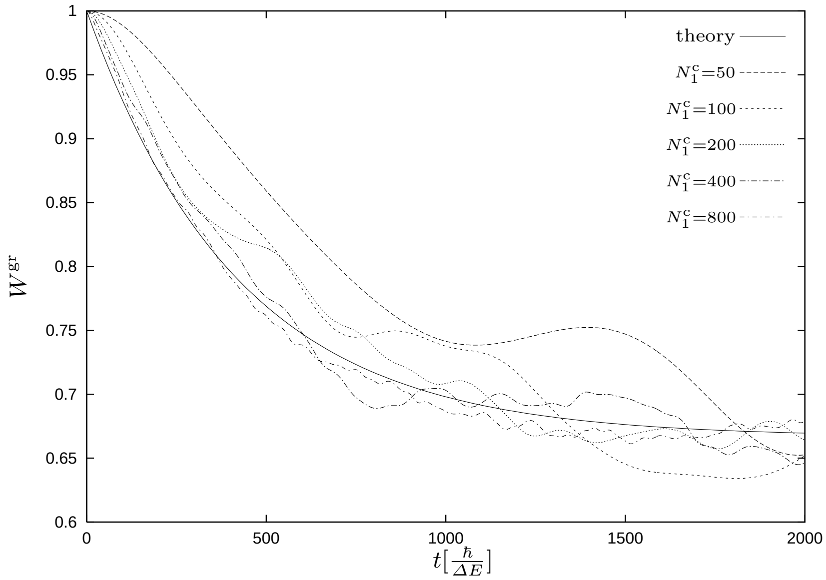


Fig. 17.5. Numerical simulation for the relaxation period. The predictions from the rate equation get better with increasing container system, N_1^{c}.

all sizes the level spacings ΔE_1^{c}, ΔE_0^{c} have been kept fixed such that for increasing container size the band widths increase.

With those parameter settings the theoretical prediction for $W^{\mathrm{gr}}(t)$ from (17.51) is the same for all container sizes. The numerical results are displayed in Fig. 17.5. The solid line is the prediction from theory.

Obviously the theoretical predictions are not accurate for "few-level" container environments. This is due to the fact that the replacement of actual quantities by their Hilbert space average is a valid approximation for high dimensional Hilbert spaces only. Furthermore, for the few-level cases the iteration step times that have to be longer than τ_1 are rather long, because already τ_1 is long. This means that the recursion cannot really be replaced by the differential equation in Sect. 17.5. 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 $N_1^{\mathrm{c}} = 800$ is obviously in good agreement with the theoretical prediction.

16 Theories of Relaxation Behavior

> *Professor Prigogine's search for "irreversibility at the microscopic level" can be seen as an attempt to extend thermodynamics all the way down to the level of quantum mechanics . . . He was hoping for a fundamental, almost deterministic partial differential equation which could describe thermodynamics at the quantum level*
>
> — K. Gustafson [44]

Almost everything stated so far has referred to situations of fully developed equilibrium. We have described 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. However, nothing has been said about how long such a relaxation time will be, or what behavior can be expected during that relaxation. This is an important aspect of non-equilibrium thermodynamics. The most common scenario for the route to equilibrium is an exponential decay scheme. One typically thinks of merely statistical processes, simply controlled by transition rates rather than of some coherent evolutions. Nevertheless such a behavior should, starting from first principles, eventually follow from the Schrödinger equation.

16.1 Fermi's Golden Rule

In the literature "Fermi's Golden Rule" is mentioned frequently to account for exponential decay scenarios, often in the context of spontaneous emission [126]. In order for exited states to decay the considered system must be coupled to a large environment. Together with this environment, the full system is assumed to feature a smooth and rather high state density. Fermi's Golden Rule (cf. Sect. 2.5.2) yields a transition rate, i.e., the system is found to disappear from its exited state with a probability P^{ex} proportional to the time passed and the probability to have been in the excited state (2.93). If such a behavior can be established at any infinitesimal time step during the evolution, an exponential decay results

$$P^{\mathrm{ex}}(t + \mathrm{d}t) = -RP^{\mathrm{ex}}(t)\,\mathrm{d}t \quad \Rightarrow \quad P^{\mathrm{ex}}(t) = P^{\mathrm{ex}}(0)\,\mathrm{e}^{-Rt} \,. \tag{16.1}$$

However, if Fermi's Golden Rule is established on the basis of a time-dependent perturbation theory, there are specific conditions, limiting the time of its applicability. On the one hand the time passed ($\mathrm{d}t$) has to be long enough to give rise to a linear growth of $P^{\mathrm{ex}}(t + \mathrm{d}t)$ with $\mathrm{d}t$ (cf. Sect. 17.4), on the other it has to be short enough to justify a truncation of the Dyson

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 165–167 (2004)
`http://www.springerlink.com/`

series at first order (2.86). Thus, the true coherent evolution during the full relaxation process can definitely not be reduced to an exponential decay behavior, controlled only by transition rates, without further considerations.

Implicitly it is often assumed that the application of Fermi's Golden Rule could somehow be iterated. The idea is to describe the respective evolution during the limited time of its applicability and then take the resulting final state for the new initial state. This, however, is not a feasible concept either, as will be explained in the following.

By means of a Dyson series, a short time evolution of the system can be described, (see (2.82))

$$|\psi(t+\mathrm{d}t)\rangle = \hat{U}_{\mathrm{I}}(t,\mathrm{d}t)|\psi(t)\rangle \; . \tag{16.2}$$

Thus, the probability for the system to be found in the energy eigenstate $|i\rangle$ at $t+\mathrm{d}t$ is given by,

$$\begin{aligned} |\langle i|\psi(t+\mathrm{d}t)\rangle|^2 &= \langle\psi(t)|\hat{U}_{\mathrm{I}}^{\dagger}(t,\mathrm{d}t)|i\rangle\langle i|\hat{U}_{\mathrm{I}}(t,\mathrm{d}t)|\psi(t)\rangle \\ &= \sum_{j,k}\langle\psi(t)|j\rangle\langle j|\hat{U}_{\mathrm{I}}^{\dagger}(t,\mathrm{d}t)|i\rangle\langle i|\hat{U}_{\mathrm{I}}(t,\mathrm{d}t)|k\rangle\langle k|\psi(t)\rangle \; . \end{aligned} \tag{16.3}$$

However, only in the case of $j=k$ this reduces to

$$|\langle i|\psi(t+\mathrm{d}t)\rangle|^2 = \sum_j |\langle i|\hat{U}_{\mathrm{I}}(t,\mathrm{d}t)|j\rangle|^2\, |\langle j|\psi(t)\rangle|^2 \; , \tag{16.4}$$

which essentially is Fermi's Golden Rule describing the transition probability from state $|j\rangle$ to state $|i\rangle$, just as (2.87) does with

$$W_{ij}(\mathrm{d}t) = |\langle i|\hat{U}_{\mathrm{I}}(t,\mathrm{d}t)|j\rangle|^2 \; . \tag{16.5}$$

Note that only for the special case of $j=k$ is it true that the new probability distribution $|\langle i|\psi(t+\mathrm{d}t)\rangle|^2$ depends only on the old probability distribution $|\langle j|\psi(t)\rangle|^2$ and that the transition rate $|\langle i|\hat{U}_{\mathrm{I}}(t,\mathrm{d}t)|j\rangle|^2$ does not depend on t. In general neither is true, but these properties are needed for an autonomous iteration. Thus, if in the beginning the initial state of the system is a simple energy eigenstate, $j=k$ holds, (16.3) may be replaced by (16.4), and Fermi's Golden Rule applies. However, if the state of the system is a superposition of energy eigenstates, as will be the case after the first time step, the evolution has to be described by (16.3). In this case, the final probability distribution depends on all details of the initial state as well as on time, and Fermi's Golden Rule does not apply. Therefore, in general, an iteration of this rule cannot be justified.

16.2 Weisskopf–Wigner Theory

An approach that is based on a (approximate) continuous solution of the Schrödinger equation for all times, rather than on an iteration scheme, is the

Weisskopf–Wigner theory. A detailed discussion of this theory is definitely beyond the scope of this text, thus we want to refer the unfamiliar reader to [118], or other standard textbooks. At this point we just want to add a few remarks.

The situation analyzed in this theory is, to some extent, the same as the one in the section above. Whether or not an approximate solution can be found depends strongly on a non-degenerate ground state of the environment and on, in the considered limit, infinitely high state density, combined with an infinitesimally weak coupling. These conditions are met, for example, by an atom coupled to the electromagnetic field in open space, but make the method difficult to apply to many other situations.

16.3 Open Quantum Systems

The pertinent ideas, although basically invented to describe the relaxation process rather than the fully developed equilibrium, have already been explained in Sect. 4.8. Here, we do not want to repeat these explanations but mention again, that those theories basically explain how a system will be driven towards equilibrium due to an environment already being in equilibrium. Furthermore, the environment (as already being in equilibrium) is assumed to be non-changing. Neither is necessarily the case for the situations analyzed in the following chapter.

20 Quantum Thermodynamic Machines

> *If an idea is really important, then it can be used to build better machines.*
>
> — R. Buckminster Fuller quoted in [109]

A machine may be defined as a physical system, which can support processes without suffering from permanent (structural) changes. In a thermodynamic machine at least part of these processes involve changes of entropy or temperature. In order to have these concepts available, it is understood that the process speed is low enough to keep the system considered close to equilibrium at all times.

Though idealized, this close-to-equilibrium condition is essential for the notorious stability and universality of thermodynamic machines; these concepts constitute the framework within which to understand even real machines.

The actual equilibrium state depends on the embedding and the control parameters entering the system Hamiltonian (see below). Reversibility implies that this state remains unchanged once those control conditions remain constant; at any instant of time the process can thus be run forward and backward.

20.1 Tri-Partite System: Sudden Changes of Embedding

Let us extend our previous considerations to a system g coupled to two mutually uncoupled environmental subsystems c_1, c_2 (cf. Fig. 20.1). The respective coupling parameters with g are λ_1, λ_2 (λ_i are the scaling factors in front of the respective coupling matrix). We assume that the effects of the two environments are in conflict in the sense that if either of them were individually coupled to g, the resulting equilibrium state for g would be different. This situation could give rise to various physical scenarios including transport. Here we restrict ourselves to a time-dependent coupling, which guarantees that only one part of the environment is coupled to g at any time. The total system is closed and subject to Schrödinger dynamics.

This change of coupling may be envisioned to happen in a periodic fashion (cf. Fig. 20.2). If $\Delta t_i, i = 1, 2$, is large compared to any relaxation times, and if the total model is such that either final state was indepenent of the initial state, the system g could switch between two canonical states. These states would then be kept also after the coupling has been switched off (period Δt_0).

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 235–240 (2004)
http://www.springerlink.com/

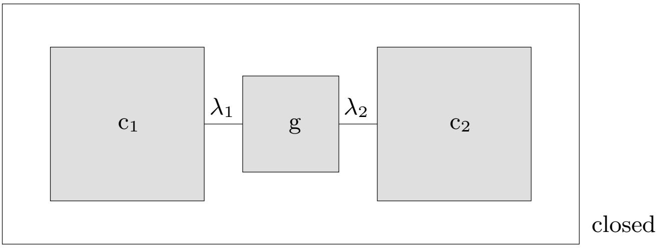


Fig. 20.1. Tri-partite system consisting of control system g and two environments c$_1$, c$_2$; λ_1, λ_2 are the respective coupling constants.

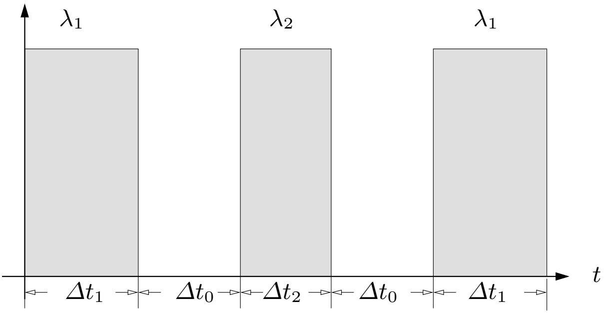


Fig. 20.2. Control pattern for a system according to Fig. 20.1.

However, the sudden changes of the coupling λ_1, λ_2 will induce relaxation processes during which the local entropy of subsystem g cannot stay constant. A reversible thermodynamic machine is based on the idea that during the periods Δt_0 the Hamiltonian of g should be changed such that the subsequent coupling to the alternate environment would not lead to any irreversible relaxation. In the ideal classical heat engines this adaption is carried out via so-called work variables. The name derives from the fact that from these variables one can calculate the work exchanged between system g and environment.

In our present approach, to be sure, thermodynamic behavior must be imposed by embedding the system under consideration into a quantum environment; changes induced via a time-dependent Hamiltonian without that embedding will not necessarily keep the system in a thermal state proper. This principal limitation of adiabatic processes also poses a principle limitation to reversible heat engines in the quantum domain [125]. When, after such an adiabatic step, the system is again brought into contact with an environment reestablishing a temperature, this adjustment will, in general, be irreversible.

20.2 Work Variables

Thermodynamic machines combine mechanical with thermal control. Mechanical control is associated with the change of (classical) parameters specifying the Hamiltonian. This change, while necessarily due to interactions with external physical subsystems and thus – in a complete quantum description – subject to entanglement, is usually treated in the classical limit. The classical limit here means that the effect of the interaction can exclusively be described in terms of "effective potentials".

Static effective potentials are a basic ingredient for defining, e.g., the constraining volume of a gas. As stressed in this book, the container does not only specify these boundary conditions, but also constitutes a source of thermodynamic behavior of the gas. As one wall of the container is replaced by a piston, the position of which is controlled from the outside, the volume V turns into an extensive work variable. Depending on the time-dependent volume, also the Hamiltonian of g becomes explicitly time-dependent (i.e., non-autonomous).

Presently it is not well understood, which (if not all) of the Hamilton parameters of a quantum system may, indeed, be made to operate as such a work variable. In classical thermodynamics we know that the list of those variables is very restricted, including, in addition to volume, extensive parameters X_i like magnetization, electrical polarization, and particle number. For these or their intensive conjugate counterparts ξ_i, practical means of control exist.

Also in the microscopic domain the type of actual control counts, not so much its mere effect; as discussed in Sect. 18.7 the spectrum of a particle in a box (delocalized eigenstates) can be changed by changing the size (volume) of the box. This control type is extensive. Alternatively, the spectrum due to bound states (e.g., of a localized atom) could be changed by a magnetic field, which would be a local (intensive) control mode. However, also new possibilities might emerge. By means of an external light field one may envision to change even the particle-particle interactions in a Bose-Einstein condensate. While the effect could not be localized in this case, the control mode was clearly intensive.

20.3 Carnot Cycle

The ideal Carnot cycle consists of alternating isothermal and adiabatic process sections, controlled by a time-dependent Hamilton parameter $a(t)$ and sudden changes of environment couplings λ_1, λ_2. The isothermal process sections allow us to change the occupation of the energy states in a robust way by slowly changing the spectrum of system g while being in contact with one of the environment parts. The low speed guarantees that the system is always in a canonical state at constant temperature (see Fig. 20.3, steps 1-2, 3-4).

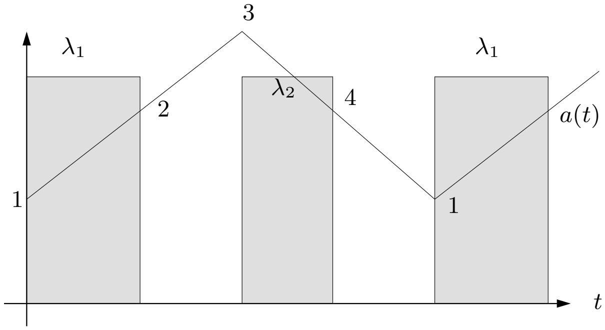


Fig. 20.3. Control pattern for system according to Fig. 20.1. $a(t)$ is the pertinent work variable.

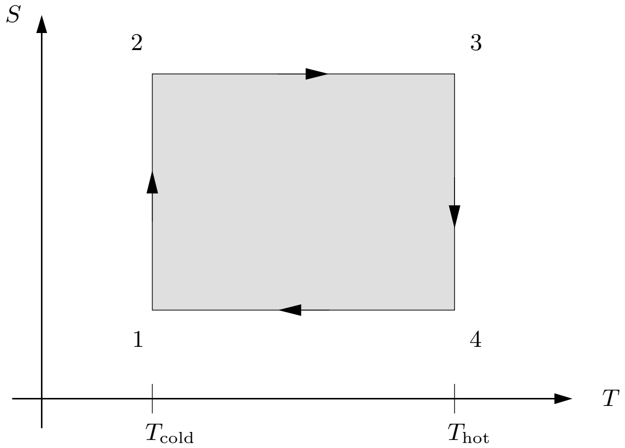


Fig. 20.4. Carnot cycle in entropy S / temperature T-space, resulting from the control pattern of Fig. 20.3.

The adiabatic process sections (Fig. 20.3, steps 2-3, 4-1) can be realized by slowly changing the spectrum of system g under isolation (no coupling to the quantum environment) or, more realistically, under microcanonical conditions (cf. Chap. 13). This will allow us to keep all occupation numbers of g constant (i.e., the entropy remains constant).

Note that the control sequence for this "four-stroke" machine lacks time-inversion symmetry. There is a physical difference in running the machine forward and backward (which translates into different applications). The whole process cycle can be reformulated in the space of thermodynamic variables as given in Fig. 20.4. This figure also shows the smoothness condition according to which the four process lines have to meet pairwise at the four points 1, 2, 3, 4, respectively. As a consequence of the energy and entropy balance, the ideal Carnot efficiency (for extracting work) is given by

$$\eta = 1 - \frac{T_{\mathrm{cold}}}{T_{\mathrm{hot}}} , \qquad T_{\mathrm{cold}} < T_{\mathrm{hot}} . \tag{20.1}$$

20.4 Generalized Control Cycles

Classical heat engines rarely comply with the ideal Carnot process; irreversibility in the form of non-equilibrium, friction, heat leakage, etc., always tends to decrease the ideal efficiency. This should not be different in the quantum case, where the notion of adiabatic processes connecting different temperatures meets with principle limitations [74]. However, the functioning of a "real machine" would still, at least in principle, be understood by referring to the ideal version. There are many different ways to perform cycles in the space of thermodynamic variables.

In the Otto cycle the two isotherms are replaced by isochores (constant extensive work variable). In a reversible fashion this could be done by connecting the working gas at constant volume with one "heat reservoir" after the other such that the temperature would gradually increase. In the practical case of a classical Otto engine, heat is supplied rather abruptly after ignition of the fuel-gas mixture (thus increasing the pressure) and removed again, when the gas is finally expanded by opening up the valve. Obviously, these two steps can physically not be run backward.

A quantum version of the Otto cycle has been studied in [35]. As the working gas, an ensemble of two-level systems has been considered, controlled by an external time-dependent magnetic field. The isochoric parts resemble the heat contacts as discussed in Chap. 20.1; this requires us to introduce here a generalized ("dynamical") temperature to account for non-equilibrium.

On the other hand, special non-equilibrium effects have even been proposed to improve efficiency. Such a case has been advocated by Scully and coworkers [119]. The pertinent machine would use a photon gas as the working substance. Work is supposed to be extracted via radiation pressure on a piston, just like for a convential gas. The photon gas can be brought into contact with two heat baths at temperature T_{cold}, T_{hot}, respectively ($T_{\mathrm{cold}} < T_{\mathrm{hot}}$), which both consist of three-level atoms. Being close to equilibrium at any time, this machine would work, as usual, as a Carnot machine.

Now, by means of an additional microwave generator the two lower levels (1,2) of the bath atoms are, before use, prepared in a coherent superposition such that the emission and the absorption of photons between levels 3 and (1,2) can be modified, due to interference (thus breaking detailed balance). The hot atoms (T_{hot}) continue to emit photons, but the cold atoms (T_{cold}) absorb less than they ordinarily would. Based on the non-equilibrium created by the microwave source, this machine would even work for $T_{\mathrm{cold}} = T_{\mathrm{hot}}$.

In fact, as the pertinent distribution function is now no longer parametrized by the temperature T alone, but by temperature and some additional parameters like the non-equilibrium polarization, it would be naive to expect that the effective machine efficiency would still depend exclusively on temperature!

While this effect clearly exploits quantum mechanics, the increased efficiency is present only as long as the cycle involves those non-equilibrium states of the respective baths (“phaseonium”). In the strict sense, the machine is not a “realistic” Carnot machine, which happens to deviate somewhat from the ideal one, but – with respect to the effect – no thermodynamic machine at all: The phaseonium is not in a thermodynamically stable state, which would maximize entropy under some specific conjugate control parameter.

Compared to the “simple” bipartite systems supporting thermal equilibrium, quantum thermodynamic machines already require more design efforts. Still being based on “incomplete control”, they, nevertheless, show robust behaviour. This class of systems can be scaled down to the quantum and nano scale, giving rise to many kinds of deviations from the classical limit. So far, such deviations have been discussed on a rudimentary level only. However, it is rather misleading to call these deviations “breaking” of the laws of thermodynamics [25, 120]. It is fairly clear that as the embedding and the working subsystems are made smaller and smaller, thermodynamics should eventually disappear: The question is, how.

Non-thermodynamic behavior is of prime interest for quantum computing [90]. The embedding of these machines is not only an unavoidable nuisance, but even an essential part of the control. While small toy systems have been demonstrated to work, it is unclear yet whether an implementation can ever be found, which may sufficiently be scalable to meet the needs of a practically useful machine (e.g., for factoring large numbers).

21 Summary and Conclusion

> *Physics in the twentieth century will probably be remembered for quantum mechanics, relativity and the standard model of particle physics. Yet the conceptual framework within which most physicists operate is not necessarily defined by the first of these and makes only reference rarely to the second two.*
>
> — P. Ball [8]

This book essentially had two intentions: to define what precisely thermodynamic behavior should be and to show that a certain class of bipartite quantum mechanical systems show this thermodynamic behavior, even if treated entirely quantum mechanically, i.e., if its behavior is exclusively described by the Schrödinger equation. Since it seems plausible that this class of systems may contain almost all systems typically considered thermodynamical (and even some more), this book may be viewed as a quantum approach to thermodynamics.

Contrary to widespread expectation, the total system is reversible, its parts, typically, show thermodynamic behavior (and irreversibility)! Thermodynamics might thus be called “non-local” (like its source, the entanglement) in the sense that it is not just a property of the system considered but of the system including its embedding. Without this embedding the stability of the thermal state is lost.

In Chap. 4, a brief overview has been given on the historical attempts to “derive” thermodynamics from an underlying theory, for the most part from classical mechanics, with an emphasis on their shortcomings. (It may be noted here that the present authors, despite their critical attitude in that chapter, do not think that those attempts are all futile, in fact they feel that some belong to the most brilliant work that has ever been done in theoretical physics). There is, in fact, a kind of “unreasonable effectiveness of classical methods in thermodynamics” (cf. Wigner's statement about mathematics [132]). 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 respective behavior without requiring additional assumptions like ergodicity.

In Chap. 5, some sort of axiomatic structure, a set of properties of thermodynamic quantities (entropy, pressure, etc.) and their respective dynamics has been given, which was meant to specify thermodynamic behavior. This set essentially corresponds to the different statements that are contained in the first and second law of thermodynamics. It is not necessarily unique, irreducible or entirely new, but organized in such a way that every rule or

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 241–244 (2004)
http://www.springerlink.com/

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 book (Part II) consists of the definitions of thermodynamic quantities on the basis of the momentary, local, quantum mechanical micro state (density operator) of the considered system, and the demonstration that these quantities, indeed, exhibit, under certain conditions, the claimed behavior (those conditions then define the thermodynamic limit). The general set-up always consists of the 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., of the considered system and the environment, is assumed to be initially given by a pure state, i.e., a wave function, and, due to the Schrödinger equation, continues to be so. Any micro state of the considered system is obtained 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 thermodynamic quantities have been defined as functions of the micro state of the full system, which can be represented by a point in Hilbert space. The whole Hilbert space may be viewed as being divided into precisely defined cells such that all points of a given cell yield the same thermodynamic quantity like entropy, energy, etc. Those cells that correspond to conserved quantities, like energy under microcanonical conditions, represent restrictions to the possible evolutions of the full system in Hilbert space. It now turns out that the biggest part of all different micro states within some “accessible cell” yield the same equilibrium values of the thermodynamic quantities, like maximum entropy or identical temperatures for subsystems, etc. This means the biggest part of the accessible Hilbert space is filled with states that correspond to thermodynamic equilibrium. The assumption now is that a system that starts in a region in Hilbert space that does not correspond to equilibrium, will typically end up wandering through a region that corresponds 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 have analyzed are definitely not ergodic, nevertheless they show thermodynamic behavior, as may be seen from the numerical simulations.)

In the present 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. In this picture, the increase of entropy is due to the build-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, though, that some features of thermodynamics, e.g., the full independence of the equilibrium state of some system from its initial state, require a certain structure of the spectrum of their environment. Since any larger macroscopic system can possibly play the role of an envinroment, this structure should be somehow generic. It has been shown in Chap. 11, that this feature results, whenever very many weakly interacting identical units form the entire system, i.e., whenever the system shows some modularity. This modularity is also linked to the extensitivity of entropy, namely that the logarithm of the state density is a homogeneous function of the first order. (This consideration is not necessarily restricted to quantum mechanics.)

Chapters 12 and 13 have dealt with the intensive quantities and their property as derivatives of energy with respect to a corresponding extensive quantity. It has been 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 has essentially been done by enlarging the concept of adiabatic following known from quantum optics.

Since entropy as a thermodynamic potential entirely determines the concrete thermodynamic behavior of a system, and since the state density 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 state density is computed based on a classical or a quantum mechanical model. (In the classical case the "state density" is just the volume of some energy shell in phase space.) Thus the connection between classical and quantum mechanical state density had to be investigated (see Chap. 14), with the result that whenever the energy spread of some minimum position-momentum uncertainty wave package is small compared to its mean energy, the classical state density divided by $\hbar$ may be taken as a good approximation to the quantum mechanical state density. The above condition is typically fulfilled either by systems consisting of heavy objects or systems made of many identical weakly interacting subsystems, which is again the condition of "modularity".

Trying to summarize the different claims and assumptions that have been exploited to establish thermodynamic behavior (Chap. 15), 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 thermodynamic behavior. Full thermodynamic behavior, however, will emerge only if all systems in quest have spectra that are typical for modular systems.

In Chap. 17 the route to equilibrium has been examined. Usually it is stated that a system in some exited 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 do so based on additional influences that are not described by the Schrödinger equation. Here we have investigated if and under what conditions such an exponential decay may result, if the respective system is, again, weakly coupled to some other system, and the whole system is described by the Schrödinger equation only. It has been 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 support and illustrate the results from this theory.

Part III deals with some more specific problems that require the use of concrete models. Chapter 18 has been devoted to numerical results pertaining to finite-size bipartite systems. These data confirm our theory, but also demonstrate "finite-size effects" in the form of time-dependent fluctuations. Equilibrium properties of spin systems and aspects of the measurement of intensive thermodynamic variables have been discussed. The existence of local temperature has been examined in some detail with special emphasis on quantum features. Chapter 19 has addressed the problem of heat conduction, notoriously a confusing field of research, especially from a classical point of view. Quantum thermodynamic machines (Chap. 20) might be seen as a direction for future research, a potential alternative to the quantum control by means of fast alternations between unitary and dissipative error correcting steps as proposed for quantum computation.

A concise conclusion from this book is hard to draw. Certainly it cannot be that all approaches to a foundation of thermodynamics so far are entirely wrong, nor that this work establishes the only access that is beyond any doubt. Nevertheless, the idea of thermodynamics being an emergent behavior pattern of a broad class of 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 more clearly. And those are not necessarily bound to some minimum number of particles of the system in quest. In the context of the present theory, thermodynamic behavior of microscopic single objects seems possible. Furthermore, violations of the second law do not appear to be entirely impossible, but just unlikely. Nevertheless, the theory might provide hints on how systems that allow to see such a violation could possibly be made. It is our hope that future experiments might help to clarify the applicability and significance of this theory, rather than offer mere theoretical reasoning.

A Hyperspheres

In the following we consider an n_{tot}-dimensional Cartesian space with the coordinates $\{x_1, \ldots, x_{n_{\mathrm{tot}}}\}$. Within this space a d dimensional hypersphere with radius R is introduced by

$$\mathrm{sph}(d) : \sum_{i=1}^{d} x_i^2 = R^2 \ , \tag{A.1}$$

where we have sorted the coordinates according to the coordinates of the hypersphere.

A.1 Surface of a Hypersphere

We restrict ourselves to a hypersphere with radius $R = 1$. To evaluate its surface $\mathcal{O}(R = 1, d)$, we consider the special integral

$$I(d) = \int_{\mathbb{R}^d} \mathrm{e}^{-\sum_{i=1}^{d} x_i^2} \prod_{i=1}^{d} \mathrm{d}x_i = \prod_{i=1}^{d} \int_{-\infty}^{\infty} \mathrm{e}^{-x_i^2} \mathrm{d}x_i = \pi^{\frac{d}{2}} \ . \tag{A.2}$$

Generalized spherical coordinates in the d dimensional subspace are defined by the transformation $\{x_1, \ldots, x_d\} \mapsto \{r, \phi_1, \ldots, \phi_{d-1}\}$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_{d-1} \\ x_d \end{pmatrix} = \begin{pmatrix} r \cos \phi_1 \\ r \sin(\phi_1) \cos(\phi_2) \\ \vdots \\ r \prod_{i=1}^{j-1} \sin(\phi_i) \cos(\phi_j) \\ \vdots \\ r \prod_{i=1}^{d-2} \sin(\phi_i) \cos(\phi_{d-1}) \\ r \prod_{i=1}^{d-2} \sin(\phi_i) \sin(\phi_{d-1}) \end{pmatrix} , \tag{A.3}$$

and the Jacobian matrix (functional matrix)

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 247–252 (2004)
`http://www.springerlink.com/`

$$\mathsf{F} = \begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \phi_1} & \cdots & \frac{\partial x_1}{\partial \phi_{d-1}} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \phi_1} & \cdots & \frac{\partial x_2}{\partial \phi_{d-1}} \\ \vdots & & \ddots & \vdots \\ \frac{\partial x_d}{\partial r} & \frac{\partial x_d}{\partial \phi_1} & \cdots & \frac{\partial x_d}{\partial \phi_{d-1}} \end{pmatrix} . \tag{A.4}$$

The determinant of this matrix can be evaluated and reads

$$\det \mathsf{F} = r^{d-1} \sin^{d-2}(\phi_1) \sin^{d-3}(\phi_2) \cdots \sin(\phi_{d-2}) . \tag{A.5}$$

The volume element transforms according to

$$\begin{aligned} \prod_{i=1}^{d} \mathrm{d}x_i &= |\det \mathsf{F}| \; \mathrm{d}r \prod_{i=1}^{d-1} \mathrm{d}\phi_i \\ &= r^{d-1} \mathrm{d}r \sin^{d-2}(\phi_1) \mathrm{d}\phi_1 \cdots \sin(\phi_{d-2})\, \mathrm{d}\phi_{d-2}\, \mathrm{d}\phi_{d-1} \\ &= r^{d-1} \mathrm{d}r\, \mathrm{d}\Omega . \end{aligned} \tag{A.6}$$

Now we evaluate the integral in spherical coordinates, too,

$$I(d) = \int_0^\infty \mathrm{e}^{-r^2} r^{d-1} \mathrm{d}r \int \mathrm{d}\Omega . \tag{A.7}$$

The last integration is just the surface $\mathcal{O}(1, d)$ of the hypersphere, and we find by evaluating the first one

$$I(d) = \frac{\Gamma(\frac{d}{2})}{2}\, \mathcal{O}(1, d) . \tag{A.8}$$

The surface of the hypersphere with radius $R = 1$ is

$$\mathcal{O}(1, d) = \frac{2I(d)}{\Gamma(\frac{d}{2})} = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} . \tag{A.9}$$

If we evaluate the surface of a hypersphere with radius R, we have to renormalize the exponent in (A.2) by $\frac{1}{R^2}$, finding $I(d) = \pi^{d/2} R^d$. Additionally, replacing the radius variable of the generalized spherical coordinate $r \mapsto \frac{r}{R}$ and integrating (A.7), we find $I(d) = \frac{1}{2} R \Gamma(\frac{d}{2}) \mathcal{O}(R, d)$, and thus

$$\mathcal{O}(R, d) = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} R^{d-1} . \tag{A.10}$$

A.2 Integration of a Function on a Hypersphere

To integrate a polynomial function over the hypersphere surface one can use the same approach as in the last section. Consider the function

$$f(x_i, x_j) = (x_i)^l (x_j)^k \quad \text{with} \quad l, k \geq 0 \;, \tag{A.11}$$

and the integration

$$\begin{aligned}
I(d) &= \int_{\mathbb{R}^d} f(x_i, x_j) \exp\left(-\frac{1}{R^2} \sum_{i'=1}^{d} x_{i'}^2\right) \prod_{j'} \mathrm{d}x_{j'} \\
&= \int_{\mathbb{R}^d} (x_i)^l (x_j)^k \exp\left(-\frac{1}{R^2} \sum_{i'=1}^{d} x_{i'}^2\right) \prod_{j'} \mathrm{d}x_{j'} \\
&= \pi^{\frac{d-2}{2}} R^{d-2} \int_{-\infty}^{\infty} (x_i)^l \exp\left(-\frac{x_i^2}{R^2}\right) \mathrm{d}x_i \int_{-\infty}^{\infty} (x_j)^k \exp\left(-\frac{x_j^2}{R^2}\right) \mathrm{d}x_j \\
&= \pi^{\frac{d-2}{2}} \frac{1}{4} (1 + (-1)^l) \Gamma(\frac{l+1}{2})(1 + (-1)^k) \Gamma(\frac{k+1}{2}) R^{d+l+k} \;.
\end{aligned} \tag{A.12}$$

We switch to generalized spherical coordinates, as in the previous section, which transforms the function (A.11) into

$$\begin{aligned}
f(x_i, x_j) \mapsto & f(\{\frac{r}{R}, \phi_1, \dots, \phi_{d-1}\}) \\
&= \left(\frac{r}{R} \prod_{i'=1}^{i-1} \sin\phi_{i'} \cos\phi_i\right)^l \left(\frac{r}{R} \prod_{i'=1}^{j-1} \sin\phi_{i'} \cos\phi_j\right)^k \\
&= \left(\frac{r}{R}\right)^l \left(\frac{r}{R}\right)^k (z_i(\{\phi_{i'}\}))^l (z_j(\{\phi_{i'}\}))^k \;,
\end{aligned} \tag{A.13}$$

and the integral into

$$I(d) = \int_0^{\infty} \mathrm{e}^{-\frac{r^2}{R^2}} \left(\frac{r}{R}\right)^{d+l+k-1} \mathrm{d}r \int (z_i(\{\phi_{i'}\}))^l (z_j(\{\phi_{i'}\}))^k \mathrm{d}\Omega \;. \tag{A.14}$$

Here we are interested in the last integral, $Z(R, d, l, k)$. From

$$I(d) = \frac{R\Gamma(\frac{d+l+k}{2})}{2} Z(1, d, l, k) \tag{A.15}$$

we find

$$\begin{aligned}
Z(R, d, l, k) &= \frac{2I(d)}{R\Gamma(\frac{d+l+k}{2})} \\
&= \frac{\pi^{\frac{d-2}{2}}}{2} (1 + (-1)^l)(1 + (-1)^k) \frac{\Gamma(\frac{l+1}{2})\Gamma(\frac{k+1}{2})}{\Gamma(\frac{d+l+k}{2})} R^{d+l+k-1} \;.
\end{aligned} \tag{A.16}$$

Note that $Z(R, d, 0, 0) = \mathcal{O}(R, d)$ and that all integrals are zero, if either l or k is odd. Here we need some special and normalized integrals only

$$\mathcal{O}(R,d,l,k) := \frac{Z(R,d,l,k)}{Z(R,d,0,0)} , \tag{A.17}$$

$$\mathcal{O}(R,d,0,1) = \mathcal{O}(R,d,1,1) = 0 , \tag{A.18}$$

$$\mathcal{O}(R,d,0,2) = \frac{R^2}{d} , \tag{A.19}$$

$$\mathcal{O}(R,d,2,2) = \frac{R^4}{d^2+2d} , \tag{A.20}$$

$$\mathcal{O}(R,d,0,4) = \frac{3R^4}{d^2+2d} . \tag{A.21}$$

A.3 Sizes of Zones on a Hypersphere

We now evaluate the area on a hypersphere with radius 1, wherein a special subsidiary condition as a function of parameter W is fulfilled. We thus consider a hyperspace $\mathrm{sph}(n_{\mathrm{tot}})$ in the space of dimension $n_{\mathrm{tot}} = d_1 + d_2$ with the Cartesian coordinates $\{x_1, \ldots, x_{d_1}, y_1, \ldots, y_{d_2}\}$ and the overall normalization

$$\mathrm{sph}(n_{\mathrm{tot}}) : \sum_{i=1}^{d_1} x_i^2 + \sum_{i=1}^{d_2} y_i^2 = 1 . \tag{A.22}$$

The above mentioned subsidiary condition for the coordinates reads

$$\mathrm{sph}(d_1) : \sum_{i=1}^{d_1} x_i^2 = W , \tag{A.23}$$

obviously also a hypersphere, now with radius $\sqrt{W}$ ($W = [0,1]$). From the normalization and from (A.23) we find for the y coordinates

$$\mathrm{sph}(d_2) : \sum_{i=1}^{d_2} y_i^2 = 1 - W , \tag{A.24}$$

again a hypersphere $\mathrm{sph}(d_2)$ with radius $\sqrt{1-W}$.

To evaluate the area on $\mathrm{sph}(n_{\mathrm{tot}})$, in which the additional subsidiary condition (A.23) is fulfilled, we integrate over the whole space, accounting for the constraints by two δ-functions

$$\mathcal{O}(\mathrm{sph}(d_1)) = \int \delta\left(\sum_{i=1}^{d_1} x_i^2 - W\right) \delta\left(\sum_{i=1}^{d_1} x_i^2 + \sum_{i=1}^{d_2} y_i^2 - 1\right) \prod_i \mathrm{d}x_i \prod_j \mathrm{d}y_j . \tag{A.25}$$

Because of the special structure of the problem (two hyperspheres $\mathrm{sph}(d_1)$ and $\mathrm{sph}(d_2)$, respectively) we switch to generalized spherical coordinates in the x subspace as well as in the y subspace

$$\begin{aligned} \{x_1,\dots,x_{d_1}\} &\rightarrow \{r_1,\phi_1,\dots,\phi_{d_1-1}\} \ , \\ \{y_1,\dots,y_{d_2}\} &\rightarrow \{r_2,\vartheta_1,\dots,\vartheta_{d_2-1}\} \ . \end{aligned} \tag{A.26}$$

The two subspaces are entirely independent and therefore the whole volume element separates into a product of two single generalized spherical volume elements of the subspaces (cf. (A.6)),

$$\begin{aligned} \prod_i \mathrm{d}x_i &= r_1^{d_1-1}\mathrm{d}r_1\mathrm{d}\boldsymbol{\phi} \ , \\ \prod_j \mathrm{d}y_j &= r_2^{d_2-1}\mathrm{d}r_2\mathrm{d}\boldsymbol{\vartheta} \ . \end{aligned} \tag{A.27}$$

Now, the integral transforms into

$$\mathcal{O}(\mathrm{sph}(d_1)) = \iint \delta\big(r_1^2 - W\big)\,\delta\big(r_1^2 + r_2^2 - 1\big)\, r_1^{d_1-1} r_2^{d_2-1}\mathrm{d}r_1\mathrm{d}r_2 \int \mathrm{d}\boldsymbol{\phi} \int \mathrm{d}\boldsymbol{\vartheta} \ . \tag{A.28}$$

According to Sect. A.1 the two integrals over $\mathrm{d}\boldsymbol{\phi}$ and $\mathrm{d}\boldsymbol{\vartheta}$ are just the surface area of the respective hyperspheres with radii 1, $\mathcal{O}(1,d_1)$ and $\mathcal{O}(1,d_2)$, respectively. For a further integration, we perform another coordinate transformation

$$\{r_1,r_2\} \mapsto \{W',r\} : \quad r_1 = \sqrt{W'} \ , \quad r_2 = \sqrt{r - W'} \ , \tag{A.29}$$

with the functional matrix and determinant

$$\mathsf{F} = \begin{pmatrix} \frac{1}{2\sqrt{W'}} & 0 \\ -\frac{1}{2\sqrt{r-W'}} & \frac{1}{2\sqrt{r-W'}} \end{pmatrix} \ , \quad \det \mathsf{F} = \frac{1}{4\sqrt{W'}}\frac{1}{\sqrt{r-W'}} \ , \tag{A.30}$$

so that we find for the integral (A.28)

$$\begin{aligned} \mathcal{O}(\mathrm{sph}(d_1)) = \frac{1}{4}\iint \delta(W'-W)\,\delta(r-1)(\sqrt{W'})^{d_1-2}(\sqrt{r-W'})^{d_2-2}\mathrm{d}r\mathrm{d}W' \\ \cdot\, \mathcal{O}(1,d_1)\,\mathcal{O}(1,d_2) \ . \end{aligned} \tag{A.31}$$

Final integration over r and W' leads to

$$\mathcal{O}(\mathrm{sph}(d_1)) = \frac{1}{4}(\sqrt{W})^{d_1-2}(\sqrt{1-W})^{d_2-2}\mathcal{O}(1,d_1)\,\mathcal{O}(1,d_2) \ . \tag{A.32}$$

This function describes the size of the zone in which the subsidiary condition (A.23) is fulfilled, as a function of the parameter W. The relative size of this zone compared to the whole area of the hypersphere $\mathrm{sph}(n_{\mathrm{tot}})$ is given by

$$\mathcal{V}(W) = \frac{\mathcal{O}(\mathrm{sph}(d_1))}{\mathcal{O}(1,n_{\mathrm{tot}})} = \frac{\Gamma\big(\frac{d_1+d_2}{2}\big)}{\Gamma\big(\frac{d_1}{2}\big)\,\Gamma\big(\frac{d_2}{2}\big)}(\sqrt{W})^{d_1-2}(\sqrt{1-W})^{d_2-2} \ . \tag{A.33}$$

This function represents the distribution of relative sizes on such a hypersphere or the relative frequency of states featuring the subsidiary condition (A.23). Obviously, the function is zero at $W = 0$ and $W = 1$ and peaked at

$$W_{\max} = \frac{d_1 - 2}{d_1 + d_2 - 4} \,. \tag{A.34}$$

Its mean value is

$$\overline{W} := \int_0^1 W \mathcal{V}(W) \, \mathrm{d}W = \frac{d_1}{d_1 + d_2} \,. \tag{A.35}$$

Thus, the maximum converges against the mean value for large d_1, d_2. The standard deviation of the function $\mathcal{V}(W)$ is

$$\Delta W := \sqrt{\int_0^1 W \mathcal{V}(W)^2 \, \mathrm{d}W - \overline{W}^2} = \sqrt{\frac{2 d_1 d_2}{(d_1 + d_2)^2 (d_1 + d_2 + 2)}} \,, \tag{A.36}$$

which means that the peak will be the sharper, the bigger d_1, d_2. Especially, if d_1, d_2, both scale with the factor α, the variance will scale as:

$$\Delta W \propto \sqrt{\frac{1}{\alpha}} \,. \tag{A.37}$$

B Hilbert Space Average under Microcanonical Conditions

We consider a space with the Cartesian coordinates $\{\eta_{ab}^{AB}, \xi_{ab}^{AB}\}$, wherein we have to evaluate the integral

$$[\![P^{\mathrm{g}}]\!] = \frac{\int_{\mathrm{AR}} P^{\mathrm{g}}(\{\eta_{ab}^{AB}, \xi_{ab}^{AB}\}) \prod_{abAB} \mathrm{d}\eta_{ab}^{AB} \mathrm{d}\xi_{ab}^{AB}}{\int_{\mathrm{AR}} \prod_{abAB} \mathrm{d}\eta_{ab}^{AB} \mathrm{d}\xi_{ab}^{AB}} \ . \tag{B.1}$$

Since this integral refers only to the accessible region (AR), the coordinates have to fulfill some subsidiary conditions defined in subspaces with definite A, B

$$\mathrm{sph}(2N_{AB}) : \sum_{ab} \left((\eta_{ab}^{AB})^2 + (\xi_{ab}^{AB})^2 \right) = W_{AB} \ . \tag{B.2}$$

Obviously, these subsidiary conditions are hyperspheres in each subspace AB with radius $\sqrt{W_{AB}}$. In the subspace AB there are $d_{AB} = 2N_{AB}$ coordinates $\{\eta_{ab}^{AB}, \xi_{ab}^{AB}\}$ (N_{AB} is the number of amplitudes ψ_{ab}^{AB} in the respective subspace), numbered by a and b. Therefore we find

$$\sum_{ab} 1 = N_{AB} = N_A N_B \ . \tag{B.3}$$

Additionally, we have to ensure that the normalization is conserved

$$\sum_{AB} \sum_{ab} \left((\eta_{ab}^{AB})^2 + (\xi_{ab}^{AB})^2 \right) = \sum_{AB} W_{AB} = 1 \ . \tag{B.4}$$

Adapting to these special subsidiary conditions we switch again to generalized spherical coordinates severally in each subspace AB: $\{r^{AB}, \phi_1^{AB}, \ldots, \phi_{d_{AB}-1}^{AB}\}$ (cf. App. A). Because of the independence of subspaces the functional matrix F then has block diagonal form, and reads

$$\mathsf{F} = \begin{pmatrix} \tilde{\mathsf{F}}^1 & & & & 0 \\ & \tilde{\mathsf{F}}^2 & & & \\ & & \ddots & & \\ & & & \tilde{\mathsf{F}}^{AB} & \\ 0 & & & & \ddots \end{pmatrix} = \prod_{AB} \mathsf{F}^{AB} , \tag{B.5}$$

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 253–258 (2004)
`http://www.springerlink.com/`

where F^{AB} has the block $\tilde{\mathsf{F}}^{AB}$ at the position AB and else the $\hat{1}$-operator. Because the determinant of a product is the product of the single determinants, we find for the functional determinant

$$\det \mathsf{F} = \det \left(\prod_{AB} \mathsf{F}^{AB} \right) = \prod_{AB} \det \mathsf{F}^{AB} \, . \tag{B.6}$$

Note that we do not integrate over the whole Hilbert space but over AR only. Thus the integral over the whole space must be restricted to AR via the product of δ-functions $\delta(r^{AB} - \sqrt{W_{AB}})$, one for each subspace AB. Because of these δ-functions we are immediately able to evaluate the integral over r^{AB}

$$\begin{aligned} &[\![P^{\mathrm{g}}]\!] \\ &= \frac{\int \int P^{\mathrm{g}}(\{r^{AB}, \phi_n^{AB}\}) \prod_{AB} \delta(r^{AB} - \sqrt{W_{AB}}) \det \mathsf{F}^{AB} \mathrm{d}r^{AB} \prod_{n=1}^{d_{AB}-1} \mathrm{d}\phi_n^{AB}}{\prod_{AB} \int \delta(r^{AB} - \sqrt{W_{AB}}) \det \mathsf{F}^{AB} \mathrm{d}r^{AB} \prod_{n=1}^{d_{AB}-1} \mathrm{d}\phi_n^{AB}} \\ &= \frac{\int P^{\mathrm{g}}(\{\sqrt{W_{AB}}, \phi_n^{AB}\}) \prod_{AB} \mathrm{d}\Omega^{AB}}{\prod_{AB} \int \mathrm{d}\Omega^{AB}} \, , \end{aligned} \tag{B.7}$$

where we have used the considerations on generalized spherical coordinates according to App. A.2, especially (A.5) and (A.6). First of all we consider the remaining integral in the denominator, finding the surface area of each subspace hypersphere sph$(2N_{AB})$

$$\prod_{AB} \int \mathrm{d}\Omega^{AB} = \prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0) \, . \tag{B.8}$$

The remaining integral over P^{g} reads with the definition of P^{g} in Cartesian coordinates according to (7.10),

$$\begin{aligned} &\frac{\int P^{\mathrm{g}}(\{\sqrt{W_{AB}}, \phi_n^{AB}\}) \prod_{AB} \mathrm{d}\Omega^{AB}}{\prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)} \\ &= \sum_{A'B'CD} \sum_{abcd} \frac{\int \psi_{ab}^{A'B'} (\psi_{cb}^{CB'})^* \psi_{cd}^{CD} (\psi_{ad}^{A'D})^* \prod_{AB} \mathrm{d}\Omega^{AB}}{\prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)} \, . \end{aligned} \tag{B.9}$$

Mind you that $\psi_{ab}^{A'B'}$ is a linear function of the coordinates $\psi_{ab}^{A'B'} = \eta_{ab}^{A'B'} + \mathrm{i}\, \xi_{ab}^{A'B'}$ and therefore P^{g} is a polynomial function of fourth order in the coordinates. According to App. A.2 we are able to evaluate all integrals over polynomial functions on hypersphere surfaces. However, all integrals over an odd power of a single coordinate vanish, as mentioned in App. A.2. Therefore we have to discriminate three different index combinations, for which the integrals are non zero: 1. $C = A'$, $D = B'$ and $c = a$, $d = b$ 2. $C = A'$ and $c = a$ 3. $D = B'$ and $d = b$, thus at least two of the amplitudes from P^{g} are equal. In all other cases we have terms with odd powers of coordinates and the integral vanishes. Let us start to consider these terms separately:

1. $C = A'$, $D = B'$ and $c = a$, $d = b$:

$$T^{(I)} = \sum_{A'B'} \sum_{ab} \frac{\int \left|\psi_{ab}^{A'B'}\right|^4 \prod_{AB} \mathrm{d}\Omega^{AB}}{\prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)} \,. \tag{B.10}$$

This is a product of integrals over all subspaces AB

$$= \sum_{A'B'} \sum_{ab} \frac{\int \prod_{\substack{AB \\ AB \neq A'B'}} \mathrm{d}\Omega^{AB} \int \left|\psi_{ab}^{A'B'}\right|^4 \mathrm{d}\Omega^{A'B'}}{\prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)} \,, \tag{B.11}$$

where the first integrals in the nominator are just $Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)$ and cancel the integrals in the denominator except one, so that

$$\begin{aligned} &= \sum_{A'B'} \sum_{ab} \frac{\int \left|\psi_{ab}^{A'B'}\right|^4 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \\ &= \sum_{A'B'} \sum_{ab} \frac{\int \left((\eta_{ab}^{A'B'})^4 + (\xi_{ab}^{A'B'})^4 + 2(\eta_{ab}^{A'B'})^2 (\xi_{ab}^{A'B'})^2\right) \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \,. \end{aligned} \tag{B.12}$$

Now we transform to generalized spherical coordinates using (A.13) (where we have to account for $R = r = \sqrt{W_{AB}}$), thus

$$\begin{aligned} = \sum_{A'B'} \sum_{ab} \Bigg(&\frac{\int (z_{ab}(\{\phi_n^{A'B'}\}))^4 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} + \frac{\int (z'_{ab}(\{\phi_n^{A'B'}\}))^4 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \\ &+ 2 \frac{\int (z_{ab}(\{\phi_n^{A'B'}\}))^2 (z'_{ab}(\{\phi_n^{A'B'}\}))^2 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \Bigg) \,. \end{aligned} \tag{B.13}$$

Using the results of App. A.2 (the first two integrals are equal) and inserting $d_{A'B'} = 2N_{A'B'}$

$$\begin{aligned} &= \sum_{A'B'} \sum_{ab} \left(\frac{2Z(\sqrt{W_{A'B'}}, d_{A'B'}, 4, 0)}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} + \frac{2Z(\sqrt{W_{A'B'}}, d_{A'B'}, 2, 2)}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \right) \\ &= \sum_{A'B'} \sum_{ab} \left(\frac{3(W_{A'B'})^2}{N_{A'B'}(2N_{A'B'} + 2)} + \frac{(W_{A'B'})^2}{N_{A'B'}(2N_{A'B'} + 2)} \right) \\ &= \sum_{A'B'} \sum_{ab} \frac{2(W_{A'B'})^2}{N_{A'B'}(N_{A'B'} + 1)} \,. \end{aligned} \tag{B.14}$$

Since there is no dependence on a and b any longer, the sum leads just to a factor $N_{A'B'}$ (cf. (B.3))

$$T^{(I)} = \sum_{A'B'} \frac{2(W_{A'B'})^2}{(N_{A'B'} + 1)} \,. \tag{B.15}$$

2. $C = A'$ and $c = a$:

$$\sum_{A'B'D} \sum_{abd} \frac{\int \left|\psi_{ab}^{A'B'}\right|^2 \left|\psi_{ad}^{A'D}\right|^2 \prod_{AB} \mathrm{d}\Omega^{AB}}{\prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)} \tag{B.16}$$

Here we have to discriminate two further cases:

a) $D = B'$:

$$T^{(II)} = \sum_{A'B'} \sum_{abd} \frac{\int \left|\psi_{ab}^{A'B'}\right|^2 \left|\psi_{ad}^{A'B'}\right|^2 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} , \tag{B.17}$$

where we have removed immediately all factors 1. Furthermore we have to exclude the case $d = b$ (formulated by a δ-function), which would lead back to the integral of the first case (cf. (B.10)):

$$= \sum_{A'B'} \sum_{abd} (1 - \delta_{db}) \frac{\int \left|\psi_{ab}^{A'B'}\right|^2 \left|\psi_{ad}^{A'B'}\right|^2 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} . \tag{B.18}$$

Considering the integrand we find four terms with squares of two different coordinates, which lead to the same value,

$$\begin{aligned} &= \sum_{A'B'} \sum_{abd} (1 - \delta_{db}) \frac{4Z(\sqrt{W_{A'B'}}, d_{A'B'}, 2, 2)}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \\ &= \sum_{A'B'} \sum_{abd} \frac{(W_{A'B'})^2}{N_{A'B'}(N_{A'B'} + 1)} - \sum_{A'B'} \sum_{ab} \frac{(W_{A'B'})^2}{N_{A'B'}(N_{A'B'} + 1)} . \end{aligned} \tag{B.19}$$

According to (B.3) the summation over a and b leads just to the number of coordinates in the subspace $A'B'$, namely $N_{A'}$ and $N_{B'}$. The summation over d leads to the number of coordinates in the container part $D = B'$, $N_{B'}$, so that

$$\begin{aligned} T^{(II)} &= \sum_{A'B'} \sum_{d} \frac{(W_{A'B'})^2}{(N_{A'B'} + 1)} - \sum_{A'B'} \frac{(W_{A'B'})^2}{(N_{A'B'} + 1)} \\ &= \sum_{A'B'} (N_{B'} - 1) \frac{(W_{A'B'})^2}{(N_{A'B'} + 1)} . \end{aligned} \tag{B.20}$$

b) $D \neq B'$:

$$T^{(III)} = \sum_{A'B'} \sum_{\substack{D \\ D \neq B'}} \sum_{abd} \frac{\int \left|\psi_{ab}^{A'B'}\right|^2 \left|\psi_{ad}^{A'D}\right|^2 \prod_{AB} \mathrm{d}\Omega^{AB}}{\prod_{AB} Z(\sqrt{W_{AB}}, d_{AB}, 0, 0)} . \tag{B.21}$$

Since $D \neq B'$, the integral factorizes into two single integrals

$$= \sum_{A'B'} \sum_{\substack{D \\ D \neq B'}} \sum_{abd} \frac{\int \left|\psi_{ab}^{A'B'}\right|^2 \mathrm{d}\Omega^{A'B'}}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \frac{\int \left|\psi_{ad}^{A'D}\right|^2 \mathrm{d}\Omega^{A'D}}{Z(\sqrt{W_{A'D}}, d_{A'D}, 0, 0)} \,. \tag{B.22}$$

Like in the previous case, one transforms to generalized spherical coordinates (remember that the absolute value has two terms with squares of coordinates) and integrates,

$$\begin{aligned} &= \sum_{A'B'} \sum_{\substack{D \\ D \neq B'}} \sum_{abd} \frac{2Z(\sqrt{W_{A'B'}}, d_{A'B'}, 2, 0)}{Z(\sqrt{W_{A'B'}}, d_{A'B'}, 0, 0)} \frac{2Z(\sqrt{W_{A'D}}, d_{A'D}, 2, 0)}{Z(\sqrt{W_{A'D}}, d_{A'D}, 0, 0)} \\ &= \sum_{A'B'} \sum_{\substack{D \\ D \neq B'}} \sum_{abd} \frac{W_{A'B'}}{N_{A'B'}} \frac{W_{A'D}}{N_{A'D}} \,. \end{aligned} \tag{B.23}$$

From (B.3) we can replace the sums over a, b and d by the number of coordinates in the respective subspace, provided the addend does not depend on these indices

$$\begin{aligned} &= \sum_{A'B'} \sum_{\substack{D \\ D \neq B'}} N_{A'} N_{B'} N_D \frac{W_{A'B'}}{N_{A'} N_{B'}} \frac{W_{A'D}}{N_{A'} N_D} \\ &= \sum_{A'B'D} (1 - \delta_{DB'}) \frac{W_{A'B'} W_{A'D}}{N_{A'}} \\ &= \sum_{A'B'D} \frac{W_{A'B'} W_{A'D}}{N_{A'}} - \sum_{A'B'} \frac{(W_{A'B'})^2}{N_{A'}} \,. \end{aligned} \tag{B.24}$$

3. $D = B'$ and $d = b$: here the considerations are the same as in the case 2., with the indices of bath and container being exchanged. After integration one finds again two terms of the form

$$T^{(IV)} = \sum_{A'B'} \left((N_{A'} - 1) \frac{(W_{A'B'})^2}{(N_{A'B'} + 1)} \right) \tag{B.25}$$

$$T^{(V)} = \sum_{A'B'C} \frac{W_{A'B'} W_{CB'}}{N_{B'}} - \sum_{A'B'} \frac{(W_{A'B'})^2}{N_{B'}} \,. \tag{B.26}$$

Since all other integrals are zero, we finally get for the Hilbert space average of P^{g} (replacing $A' \to A$ and $B' \to B$)

$$\begin{aligned}
[\![P^{\mathrm{g}}]\!] &= T^{(I)} + T^{(II)} + T^{(III)} + T^{(IV)} + T^{(V)} \\
&= \sum_{AB} \frac{2(W_{AB})^2}{(N_{AB}+1)} \\
&+ \sum_{AB} \left((N_B - 1) \frac{(W_{AB})^2}{(N_{AB}+1)} \right) + \sum_{ABD} \frac{W_{AB} W_{AD}}{N_A} - \sum_{AB} \frac{(W_{AB})^2}{N_A} \\
&+ \sum_{AB} \left((N_A - 1) \frac{(W_{AB})^2}{(N_{AB}+1)} \right) + \sum_{ABC} \frac{W_{AB} W_{CB}}{N_B} - \sum_{AB} \frac{(W_{AB})^2}{N_B} \\
&= \sum_{AB} \frac{(W_{AB})^2}{N_{AB}+1} (N_A + N_B) \\
&+ \sum_{AB} \frac{W_{AB}}{N_A} \left(\sum_D W_{AD} - W_{AB} \right) + \sum_{AB} \frac{W_{AB}}{N_B} \left(\sum_C W_{CB} - W_{AB} \right) .
\end{aligned} \tag{B.27}$$

In the case of an initial product state, $W_{AB} = W_A W_B$, we can further simplify

$$\begin{aligned}
[\![P^{\mathrm{g}}]\!] &= \sum_{AB} W_A^2 W_B^2 \frac{N_A + N_B}{N_A N_B + 1} \\
&+ \sum_A \frac{W_A^2}{N_A} \Big(\sum_B W_B \sum_D W_D - \sum_B W_B^2 \Big) \\
&+ \sum_B \frac{W_B^2}{N_B} \Big(\sum_A W_A \sum_C W_C - \sum_A W_A^2 \Big) .
\end{aligned} \tag{B.28}$$

Because W_A, W_C, W_B and W_D are probabilities, the complete sum over such probabilities has to be 1, so that

$$\begin{aligned}
[\![P^{\mathrm{g}}]\!] &= \sum_{AB} W_A^2 W_B^2 \frac{N_A + N_B}{N_A N_B + 1} \\
&+ \sum_A \frac{W_A^2}{N_A} \Big(1 - \sum_B W_B^2 \Big) + \sum_B \frac{W_B^2}{N_B} \Big(1 - \sum_A W_A^2 \Big) .
\end{aligned} \tag{B.29}$$

C Hilbert Space Averages and Variances

In this chapter we calculate the concrete Hilbert space average of the special quantities introduced in Chap. 17, i.e. κ^{gr}, κ^{ex} and α. We will need similar techniques as introduced in App. B for calculating Hilbert space averages and especially the integration of polynomial functions over a hypersphere in a high dimensional space (see App. A.2). Additionally, we calculate the variances of these quantities, as required in Sect. 17.7.

Let us start with a general type of Hilbert space average and Hilbert space variance, which will simplify the later considerations.

C.1 General Considerations

For later use we calculate the Hilbert space average and the Hilbert space variance of some Hermitian operator $\hat{A}$ in general. Therefore we consider a n_{tot} dimensional Hilbert space and the general state $|\psi\rangle$ within this space. Remember that the overall normalization condition for the state defines a hypersphere in the $2n_{\mathrm{tot}}$ dimensional parameter space $\{\eta_i, \xi_i\}$

$$\mathrm{sph}(2n_{\mathrm{tot}}) : \sum_i \left(\eta_i^2 + \xi_i^2\right) = \langle\psi|\psi\rangle = 1 \; . \tag{C.1}$$

In order to calculate the Hilbert space average of the expectation value $\langle\psi|\hat{A}|\psi\rangle$, it is useful to consider this quantity in explicit tensor notation

$$[\![\langle\psi|\hat{A}|\psi\rangle]\!] = [\![\sum_{ij} \psi_i^* A_{ij} \psi_j]\!] = \sum_{ij} A_{ij} [\![\psi_i^* \psi_j]\!] \; . \tag{C.2}$$

The last expression is just an integral of a polynomial function in parameter space over the hypersphere with radius 1. Such integrals over coordinates are certainly zero, if we integrate over a single linear coordinate (cf. A.2). Therefore this sum has non-zero addends for $i = j$ only, implying for the Hilbert space average

$$[\![\psi_i^* \psi_i]\!] = \frac{\int_{\mathrm{AR}} |\psi_i|^2 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{\int_{\mathrm{AR}} \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i} \; . \tag{C.3}$$

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 259–264 (2004)
`http://www.springerlink.com/`

The integral in the denominator is just the surface area of the respective hypersphere $Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\text{tot}}, 0, 0)$

$$= \frac{\int_{\text{AR}}(\eta_i^2 + \xi_i^2) \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\text{tot}}, 0, 0)} . \tag{C.4}$$

With App. A.2 and especially (A.19) we find eventually

$$= \frac{2Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\text{tot}}, 2, 0)}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\text{tot}}, 0, 0)} = \frac{\langle\psi|\psi\rangle}{n_{\text{tot}}} . \tag{C.5}$$

The Hilbert space average of the expectation value now reads

$$[\![\langle\psi|\hat{A}|\psi\rangle]\!] = \sum_{ij} A_{ij} \frac{\langle\psi|\psi\rangle}{n_{\text{tot}}} \delta_{ij} = \frac{\langle\psi|\psi\rangle}{n_{\text{tot}}} \sum_i A_{ii} = \frac{\langle\psi|\psi\rangle}{n_{\text{tot}}} \mathrm{Tr}\{\hat{A}\} , \tag{C.6}$$

and, because of $\langle\psi|\psi\rangle = 1$, especially

$$[\![\langle\psi|\hat{A}|\psi\rangle]\!] = \frac{\mathrm{Tr}\{\hat{A}\}}{n_{\text{tot}}} . \tag{C.7}$$

Turning to the Hilbert space variance of the expectation value of a Hermitian operator $\hat{A}$,

$$\Delta_{\text{H}}(\langle\psi|\hat{A}|\psi\rangle) = \sqrt{[\![(\langle\psi|\hat{A}|\psi\rangle)^2]\!] - [\![\langle\psi|\hat{A}|\psi\rangle]\!]^2} , \tag{C.8}$$

we consider in tensor notation

$$[\![(\langle\psi|\hat{A}|\psi\rangle)^2]\!] = [\![\langle\psi|\hat{A}|\psi\rangle\langle\psi|\hat{A}|\psi\rangle]\!] = \sum_{ijkl} A_{ij} A_{kl} [\![\psi_i^* \psi_j \psi_k^* \psi_l]\!] . \tag{C.9}$$

In the case of $i \neq j \neq k \neq l$ all addends of the sum are zero, and also, if three of the four indices are equal (see A.2). Only if all indices are equal or if there are always pairs of two equal indices, may the addend be non-zero. Collecting all possible non-zero terms we get

$$\begin{aligned} [\![(\langle\psi|\hat{A}|\psi\rangle)^2]\!] =& \sum_{\substack{ij \\ i\neq j}} [\![|\psi_i|^2 \, |\psi_j|^2]\!] (A_{ii}A_{jj} + A_{ij}A_{ji}) \\ &+ \sum_{\substack{ij \\ i\neq j}} [\![(\psi_i^*)^2 (\psi_j)^2]\!] A_{ij} A_{ij} + \sum_i A_{ii}^2 [\![|\psi_i|^4]\!] . \end{aligned} \tag{C.10}$$

In the following we calculate the three remaining Hilbert space averages. The first one reads ($i \neq j$)

$$\begin{aligned}
[\![|\psi_i|^2\,|\psi_j|^2]\!] &= \frac{\int_{\mathrm{AR}} |\psi_i|^2\,|\psi_j|^2 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{\int_{\mathrm{AR}} \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i} = \frac{\int_{\mathrm{AR}} |\psi_i|^2\,|\psi_j|^2 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} \\
&= \frac{4Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 2, 2)}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} = \frac{1}{n_{\mathrm{tot}}(n_{\mathrm{tot}}+1)} \,. \qquad \text{(C.11)}
\end{aligned}$$

The second one is a little bit more complicated, we have to look for the expression in the coordinates $(i \neq j)$

$$\begin{aligned}
[\![(\psi_i^*)^2(\psi_j)^2]\!] &= \frac{\int_{\mathrm{AR}} (\psi_i^*)^2(\psi_j)^2 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} \\
&= \frac{\int_{\mathrm{AR}} \eta_i^2\eta_j^2 - \eta_i^2\xi_j^2 - \xi_i^2\eta_j^2 + \xi_i^2\xi_j^2 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} \\
&+ \frac{\int_{\mathrm{AR}} 2\mathrm{i}\eta_i\xi_i(\xi_j^2 - \eta_j^2) + 2\mathrm{i}\eta_j\xi_j(\eta_i^2 - \xi_i^2) \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} \,. \qquad \text{(C.12)}
\end{aligned}$$

The last two integrals are zero because of the linear coordinates. Since the first integrals all have the same value but different signs, these integrals together are also zero,

$$[\![(\psi_i^*)^2(\psi_j)^2]\!] = 0 \,. \qquad \text{(C.13)}$$

The remaining Hilbert space average reads

$$\begin{aligned}
[\![|\psi_i|^4]\!] &= \frac{\int_{\mathrm{AR}} |\psi_i|^4 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} = \frac{\int_{\mathrm{AR}} \eta_i^4 + \xi_i^4 + 2\eta_i^2\xi_i^2 \prod_n \mathrm{d}\eta_i \mathrm{d}\xi_i}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} \\
&= \frac{2Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 4, 0)}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} + \frac{2Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 2, 2)}{Z(\sqrt{\langle\psi|\psi\rangle}, 2n_{\mathrm{tot}}, 0, 0)} \\
&= \frac{2}{n_{\mathrm{tot}}(n_{\mathrm{tot}}+1)} \,. \qquad \text{(C.14)}
\end{aligned}$$

All together we find now

$$[\![(\langle\psi|\hat{A}|\psi\rangle)^2]\!] = \frac{1}{n_{\mathrm{tot}}(n_{\mathrm{tot}}+1)} \Big(\sum_{\substack{ij \\ i\neq j}} (A_{ii}A_{jj} + A_{ij}A_{ji}) + 2\sum_i A_{ii}^2 [\![|\psi_i|^4]\!] \Big) \,. \qquad \text{(C.15)}$$

The missing terms in the first two sums can be found in the last sum

$$\begin{aligned}
[\![(\langle\psi|\hat{A}|\psi\rangle)^2]\!] &= \frac{1}{n_{\mathrm{tot}}(n_{\mathrm{tot}}+1)} \sum_{ij} (A_{ii}A_{jj} + A_{ij}A_{ji}) \\
&= \frac{1}{n_{\mathrm{tot}}(n_{\mathrm{tot}}+1)} \big(\mathrm{Tr}\{\hat{A}\}^2 + \mathrm{Tr}\{\hat{A}^2\} \big) \,. \qquad \text{(C.16)}
\end{aligned}$$

With this result and the Hilbert space average of the expectation value we can now calculate the Hilbert space variance of the expectation value

$$\Delta_{\mathrm{H}}^2(\langle\psi|\hat{A}|\psi\rangle) = \frac{1}{n_{\mathrm{tot}}+1}\left(\frac{\mathrm{Tr}\{\hat{A}^2\}}{n_{\mathrm{tot}}} - \left(\frac{\mathrm{Tr}\{\hat{A}\}}{n_{\mathrm{tot}}}\right)^2\right) \tag{C.17}$$

thus, the Hilbert space variance squared is simply the spectral variance squared, divided by the dimension plus 1.

C.2 Special Hilbert Space Averages

As explained in Sect. 17.1, the total wavevector is given by $|\psi\rangle = |\psi^{\mathrm{gr}}\rangle + |\psi^{\mathrm{ex}}\rangle$ with

$$\langle\psi|\psi\rangle = \langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle + \langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle = 1\ . \tag{C.18}$$

According to this normalization condition, it is again possible to formulate a hypersphere subsidiary condition in each subspace, subspace gr (system in the ground state) with dimension N_1^{c} as well as subspace ex (system in the excited state) with the dimension N_0^{c},

$$\mathrm{sph}(2N_1^{\mathrm{c}}) : \sum_i (\eta_i^{\mathrm{gr}})^2 + (\xi_i^{\mathrm{gr}})^2 = \langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle = 1 - \langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle\ , \tag{C.19}$$

$$\mathrm{sph}(2N_0^{\mathrm{c}}) : \sum_j (\eta_j^{\mathrm{ex}})^2 + (\xi_j^{\mathrm{ex}})^2 = \langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle\ . \tag{C.20}$$

These are two hyperspheres with the radii $\sqrt{\langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle}$ and $\sqrt{\langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle}$, respectively. The index i is restricted to the subspace gr with $2N_1^{\mathrm{c}}$ coordinates $(\{\eta_i^{\mathrm{gr}}, \xi_i^{\mathrm{gr}}\})$, whereas j belongs to the other subspace ex with $2N_0^{\mathrm{c}}$ coordinates $(\{\eta_j^{\mathrm{ex}}, \xi_j^{\mathrm{ex}}\})$.

Let us start now to calculate $[\![\alpha]\!]$ by writing it out in detailed tensor notation

$$[\![\alpha]\!] = [\![\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle]\!] = [\![\sum_{ij}(\psi_i^{\mathrm{gr}})^* U_{ij}\psi_j^{\mathrm{ex}}]\!] = \sum_{ij} U_{ij}[\![(\psi_i^{\mathrm{gr}})^*\psi_j^{\mathrm{ex}}]\!]\ . \tag{C.21}$$

Since $|\psi^{\mathrm{gr}}\rangle$, $|\psi^{\mathrm{ex}}\rangle$ refer to entirely different subspaces, the ψ_i^{gr} and ψ_j^{ex} always consist of different coordinates, and we get for the Hilbert space average, as argued in App. A.2

$$[\![(\psi_i^{\mathrm{gr}})^*\psi_j^{\mathrm{ex}}]\!] = \frac{\int_{\mathrm{AR}}(\psi_i^{\mathrm{gr}})^* \prod_n \mathrm{d}\eta_n^{\mathrm{gr}}\mathrm{d}\xi_n^{\mathrm{gr}} \int_{\mathrm{AR}}\psi_j^{\mathrm{ex}} \prod_m \mathrm{d}\eta_m^{\mathrm{ex}}\mathrm{d}\xi_m^{\mathrm{ex}}}{\int_{\mathrm{AR}}\prod_n \mathrm{d}\eta_n^{\mathrm{gr}}\mathrm{d}\xi_n^{\mathrm{gr}} \prod_m \mathrm{d}\eta_m^{\mathrm{ex}}\mathrm{d}\xi_m^{\mathrm{ex}}} = 0\ . \tag{C.22}$$

(Integrals of a single coordinate over a hypersphere are always zero.) Thus we get for the complete Hilbert space average

$$[\![\alpha]\!] = 0\ . \tag{C.23}$$

Of course, the Hilbert space average of the adjoint of α is zero as well.

For $[\![\kappa^{\mathrm{gr}}]\!]$ we can use the general result (C.5) from Sect. C.1, because this is just the Hilbert space average of an expectation value of a Hermitian operator. Note that $|\psi^{\mathrm{gr}}\rangle$ is not normalized and refers to the subspace gr. It follows that the trace does not run over the whole space but over the subspace gr only, finding now

$$[\![\kappa^{\mathrm{gr}}]\!] = [\![\langle\psi^{\mathrm{gr}}|\hat{U}_1^2|\psi^{\mathrm{gr}}\rangle]\!] = \frac{\langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle}{N_1^{\mathrm{c}}}\mathrm{Tr}_{\mathrm{gr}}\{\hat{U}_1^2\}\ . \tag{C.24}$$

$\mathrm{Tr}_{\mathrm{gr}}\{\dots\}$ denotes the partial trace in the subspace gr. In the same way we can calculate the Hilbert space average of κ^{ex}

$$[\![\kappa^{\mathrm{ex}}]\!] = [\![\langle\psi^{\mathrm{ex}}|\hat{U}_1^2|\psi^{\mathrm{ex}}\rangle]\!] = \frac{\langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle}{N_0^{\mathrm{c}}}\mathrm{Tr}_{\mathrm{ex}}\{\hat{U}_1^2\}\ , \tag{C.25}$$

thus getting a totally symmetrical expression under exchange of the subspaces.

C.3 Special Hilbert Space Variances

For Sect. 17.7 we need the variance of the quantity

$$\begin{aligned}\Delta_{\mathrm{H}}^2(-\mathrm{i}(\alpha-\alpha^*)) &= [\![-(\alpha-\alpha^*)^2]\!] - [\![(\alpha-\alpha^*)]\!]^2\\ &= -[\![\alpha^2]\!] - [\![(\alpha^*)^2]\!] + 2[\![\alpha\alpha^*]\!] - ([\![\alpha]\!] - [\![\alpha^*]\!])^2\ ,\end{aligned} \tag{C.26}$$

where we have used the definition of the Hilbert space variance from (8.26). The last two terms are zero according to Sect. C.2. We consider now the first term in tensor notation,

$$\begin{aligned}[\![\alpha^2]\!] &= [\![(\langle\psi^{\mathrm{gr}}|\hat{U}_1|\psi^{\mathrm{ex}}\rangle)^2]\!] = [\![\sum_{ijkl}(\psi_i^{\mathrm{gr}})^* U_{ij}\psi_j^{\mathrm{ex}}(\psi_k^{\mathrm{gr}})^* U_{kl}\psi_l^{\mathrm{ex}}]\!]\\ &= \sum_{ijkl} U_{ij}U_{kl}[\![(\psi_i^{\mathrm{gr}})^*\psi_j^{\mathrm{ex}}(\psi_k^{\mathrm{gr}})^*\psi_l^{\mathrm{ex}}]\!]\ .\end{aligned} \tag{C.27}$$

This sum may have non-zero addends for $k = i$ and $l = j$ only, else we again integrate over a quantity linear in the coordinates, leading to zero. We consider such a term

$$\begin{aligned}&[\![(\psi_i^{\mathrm{gr}})^*(\psi_i^{\mathrm{gr}})^*\psi_j^{\mathrm{ex}}\psi_j^{\mathrm{ex}}]\!]\\ &= \frac{\int_{\mathrm{AR}}(\psi_i^{\mathrm{gr}})^*(\psi_i^{\mathrm{gr}})^*\prod_n \mathrm{d}\eta_n^{\mathrm{gr}}\mathrm{d}\xi_n^{\mathrm{gr}}\int_{\mathrm{AR}}\psi_j^{\mathrm{ex}}\psi_j^{\mathrm{ex}}\prod_m \mathrm{d}\eta_m^{\mathrm{ex}}\mathrm{d}\xi_m^{\mathrm{ex}}}{Z(\sqrt{\langle\psi^{\mathrm{gr}}|\psi^{\mathrm{gr}}\rangle}, 2N_1^{\mathrm{c}}, 0, 0)Z(\sqrt{\langle\psi^{\mathrm{ex}}|\psi^{\mathrm{ex}}\rangle}, 2N_0^{\mathrm{c}}, 0, 0)}\ ,\end{aligned} \tag{C.28}$$

and the integrals over the subspace gr in the nominator

$$
\begin{aligned}
\int_{\mathrm{AR}} (\psi_i^{\mathrm{gr}})^* (\psi_i^{\mathrm{gr}})^* \prod_n \mathrm{d}\eta_n^{\mathrm{gr}} \mathrm{d}\xi_n^{\mathrm{gr}} &= \int_{\mathrm{AR}} (\eta_i^{\mathrm{gr}})^2 - 2\mathrm{i}\eta_i^{\mathrm{gr}} \xi_i^{\mathrm{gr}} - (\xi_i^{\mathrm{gr}})^2 \prod_n \mathrm{d}\eta_n^{\mathrm{gr}} \mathrm{d}\xi_n^{\mathrm{gr}} \\
&= \int_{\mathrm{AR}} (\eta_i^{\mathrm{gr}})^2 - (\xi_i^{\mathrm{gr}})^2 \prod_n \mathrm{d}\eta_n^{\mathrm{gr}} \mathrm{d}\xi_n^{\mathrm{gr}} = 0 \, .
\end{aligned}
\tag{C.29}
$$

Also the second integral in the nominator is zero, and therefore all addends in (C.27) are zero. With the same result also for the second term of (C.26), we turn to the remaining Hilbert space average

$$
\begin{aligned}
[\![\alpha\alpha^*]\!] &= [\![\sum_{ij} U_{ij} (\psi_i^{\mathrm{gr}})^* \psi_j^{\mathrm{ex}} \sum_{kl} U_{kl} (\psi_k^{\mathrm{ex}})^* \psi_l^{\mathrm{gr}}]\!] \\
&= \sum_{ijkl} U_{ij} U_{kl} [\![(\psi_i^{\mathrm{gr}})^* \psi_j^{\mathrm{ex}} (\psi_k^{\mathrm{ex}})^* \psi_l^{\mathrm{gr}}]\!] \, .
\end{aligned}
\tag{C.30}
$$

Again, the addends of the sum are non-zero in the case of $l = i$ and $k = j$ only, and we find

$$
\begin{aligned}
[\![|\psi_i^{\mathrm{gr}}|^2 \, |\psi_j^{\mathrm{ex}}|^2]\!] &= \frac{\int_{\mathrm{AR}} |\psi_i^{\mathrm{gr}}|^2 \prod_n \mathrm{d}\eta_n^{\mathrm{gr}} \mathrm{d}\xi_n^{\mathrm{gr}} \int_{AR} |\psi_j^{\mathrm{ex}}|^2 \prod_m \mathrm{d}\eta_m^{\mathrm{ex}} \mathrm{d}\xi_m^{\mathrm{ex}}}{Z(\sqrt{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle}, 2N_1^{\mathrm{c}}, 0, 0) Z(\sqrt{\langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}, 2N_0^{\mathrm{c}}, 0, 0)} \\
&= 4 \frac{Z(\sqrt{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle}, 2N_1^{\mathrm{c}}, 2, 0) Z(\sqrt{\langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}, 2N_0^{\mathrm{c}}, 2, 0)}{Z(\sqrt{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle}, 2N_1^{\mathrm{c}}, 0, 0) Z(\sqrt{\langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}, 2N_0^{\mathrm{c}}, 0, 0)} \\
&= \frac{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle \langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}{N_1^{\mathrm{c}} N_0^{\mathrm{c}}} \, .
\end{aligned}
\tag{C.31}
$$

Thus we get

$$
[\![\alpha\alpha^*]\!] = \frac{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle \langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}{N_1^{\mathrm{c}} N_0^{\mathrm{c}}} \sum_{ij} U_{ij} U_{ji} = \frac{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle \langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}{N_1^{\mathrm{c}} N_0^{\mathrm{c}}} \mathrm{Tr}_{\mathrm{gr}} \{ \hat{U}_1^2 \} \, .
\tag{C.32}
$$

Finally, we find for the Hilbert space variance of the special quantity

$$
\Delta_{\mathrm{H}}^2(-\mathrm{i}(\alpha - \alpha^*)) = 2 \frac{\langle \psi^{\mathrm{gr}} | \psi^{\mathrm{gr}} \rangle \langle \psi^{\mathrm{ex}} | \psi^{\mathrm{ex}} \rangle}{N_1^{\mathrm{c}} N_0^{\mathrm{c}}} \mathrm{Tr}_{\mathrm{gr}} \{ \hat{U}_1^2 \} \, .
\tag{C.33}
$$

To calculate $\Delta_{\mathrm{H}}^2(\kappa^{\mathrm{gr}})$ and $\Delta_{\mathrm{H}}^2(\kappa^{\mathrm{ex}})$ we can use the Hilbert space variance of an expectation value of a Hermitian operator from (C.17), with $\hat{A} = \hat{U}_1^2$, and the respective dimensions of subspaces,

$$
\Delta_{\mathrm{H}}^2(\kappa^{\mathrm{gr}}) = \frac{1}{N_1^{\mathrm{c}} + 1} \left(\frac{\mathrm{Tr}\{\hat{U}_1^4\}}{N_1^{\mathrm{c}}} - \left(\frac{\mathrm{Tr}\{\hat{U}_1^2\}}{N_1^{\mathrm{c}}} \right)^2 \right) ,
\tag{C.34}
$$

$$
\Delta_{\mathrm{H}}^2(\kappa^{\mathrm{ex}}) = \frac{1}{N_0^{\mathrm{c}} + 1} \left(\frac{\mathrm{Tr}\{\hat{U}_1^4\}}{N_0^{\mathrm{c}}} - \left(\frac{\mathrm{Tr}\{\hat{U}_1^2\}}{N_0^{\mathrm{c}}} \right)^2 \right) .
\tag{C.35}
$$

D Power of a Function

Here we show that the k-th power of any function with a global maximum will essentially be a Gaussian. For this purpose we consider a function $f(x)$ with a global maximum at $x = 0$. Because of the positivity of $f(x) > 0$ we can rewrite the function

$$f(x) = \mathrm{e}^{g(x)} \quad \text{with} \quad g(x) = \ln f(x) \ . \tag{D.1}$$

Since the logarithm is a monotonous function, we consider instead of $f(x)$ the expansion of the function $g(x)$ around the global maximum $x = 0$,

$$g(x) = \sum_i C_i x^i = C_0 - C_2 x^2 + C_3 x^3 + \cdots \tag{D.2}$$

with $C_2 > 0$ and thus

$$f(x) = \mathrm{e}^{C_0}\, \mathrm{e}^{-C_2 x^2} \mathrm{e}^{C_3 x^3} \cdots \tag{D.3}$$

with some constants C_i. Since multiplying the function with itself will amplify the maximum in the center, we can truncate the decomposition in this way. Multiplying the function k times with itself we get

$$\Big(f(x)\Big)^k = \mathrm{e}^{kC_0} \mathrm{e}^{-kC_2 x^2} \mathrm{e}^{kC_3 x^3} \cdots \ . \tag{D.4}$$

The value of x, for which the quadratic part will have reduced the function to half maximum, i.e., for which

$$\exp\big(-kC_2 x_\mathrm{h}^2\big) = \frac{1}{2} \tag{D.5}$$

is

$$x_\mathrm{h} = \pm\sqrt{\frac{\ln 2}{kC_2}} \ . \tag{D.6}$$

Evaluating the third order part of $(f(x))^k$ at x_h then yields

$$\exp\big(kC_3 x_\mathrm{h}^3\big) = \exp\left(\frac{C_3 \left(\frac{\ln 2}{C_2}\right)^{3/2}}{\sqrt{k}} \right) \ , \tag{D.7}$$

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 265–266 (2004)
`http://www.springerlink.com/`

which tends to 1 if k approaches infinity. For the relevant region, in which the function is peaked, we can thus approximate

$$\Big(f(x)\Big)^k \approx \mathrm{e}^{kC_0}\mathrm{e}^{-kC_2x^2}\ , \tag{D.8}$$

which is essentially a Gaussian.

E Local Temperature Conditions for a Spin Chain*

In this chapter we present the technical details of the application of the local temperature conditions (18.32) and (18.34) to a spin chain (18.38).

The entire chain with periodic boundary conditions may be diagonalized via successive Jordan–Wigner, Fourier and Bogoliubov transformations [49]. The relevant energy scale is introduced via the thermal expectation value (without the ground state energy)

$$\overline{E} = \frac{N N^{\mathrm{G}}}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}k \, \frac{\omega_k}{\exp(\beta \, \omega_k) + 1} \, , \tag{E.1}$$

where

$$\omega_k = 2\varDelta E \sqrt{(1 - K \cos k)^2} \, , \tag{E.2}$$

with $K = \lambda / \varDelta E$. The ground state energy E_0 is given by

$$E_0 = - \frac{N N^{\mathrm{G}}}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}k \, \frac{\omega_k}{2} \, . \tag{E.3}$$

Since $N^{\mathrm{G}} \gg 1$, the sums over all modes have been replaced by integrals.

If one partitions the chain into N^{G} groups of N subsystems each, the groups may also be diagonalized via a Jordan–Wigner and a Fourier transformation [49] and the energy E_a reads

$$E_a = 2\varDelta E \sum_{\nu=1}^{N^{\mathrm{G}}} \sum_{k} (1 - K \cos(k)) \left(n_k^a(\nu) - \frac{1}{2} \right) \, , \tag{E.4}$$

where $k = \pi l/(N+1)$ $(l = 1, \ldots, N)$ and $n_k^a(\nu)$ is the fermionic occupation number of mode k of group number ν in the state $|a\rangle$. It can take on the values 0 and 1.

For the model at hand one has $\varepsilon_a = 0$ for all states $|a\rangle$, while the squared variance $\varDelta_a^2$ reads

$$\varDelta_a^2 = \sum_{\nu=1}^{N^{\mathrm{G}}} \varDelta_\nu^2 \, , \tag{E.5}$$

---

[*] Based on [46, 47, 48, 49] by Hartmann et al.

J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics,* Lect. Notes Phys. **657**, 267–269 (2004)
http://www.springerlink.com/

with

$$\begin{aligned}\varDelta_\nu^2 =& \frac{\varDelta E^2\, K^2}{2} - \\ &- \frac{8\,\varDelta E^2\, K^2}{(N+1)^2} \sum_k \sin^2(k) \left(n_k^a(\nu) - \frac{1}{2} \right) \sum_p \sin^2(p) \left(n_p^a(\nu+1) - \frac{1}{2} \right) \; .\end{aligned} \tag{E.6}$$

We now turn to analyze conditions (18.32) and (18.34). According to equation (E.6), $\varDelta_\nu^2$ cannot be expressed in terms of $E_{\nu-1}$ and E_ν. We therefore approximate (18.32) and (18.34) by simpler expressions.

Let us first analyze condition (18.32). Since it cannot be checked for every state $|a\rangle$, we make the following approximations.

For the present model with $|K| < 1$ all occupation numbers $n_k^a(\nu)$ are zero in the ground state and thus $\varDelta_a^2 = 0$ as well as $E_a - E_0 = 0$. Therefore (18.32) cannot hold for this state. However, if one occupation number is changed from 0 to 1, $\varDelta_a^2$ changes at most by $4\varDelta E^2 K^2/(n+1)$ and E_a changes at least by $2\varDelta E(1-|K|)$. Therefore (18.32) will hold for all states except the ground state if

$$N > 2\varDelta E \beta \frac{K^2}{1-|K|} \; . \tag{E.7}$$

If $|K| > 1$, occupation numbers of modes with $\cos(k) < 1/|K|$ are zero in the ground state and occupation numbers of modes with $\cos(k) > 1/|K|$ are equal to one. $\varDelta_a^2$ for the ground state then is $\left[\varDelta_a^2\right]_{\mathrm{gs}} \approx \left[\varDelta_a^2\right]_{\max}/2$ (in this entire chapter, $[x]_{\min}$ and $[x]_{\max}$ denote the minimal and maximal values x can take on). We therefore approximate (18.32) with the stronger condition

$$\frac{E_{\min} - E_0}{N^{\mathrm{G}}} > \beta \, \frac{\left[\varDelta_a^2\right]_{\max}}{N^{\mathrm{G}}} \; , \tag{E.8}$$

which implies that (18.32) holds for all states $|a\rangle$ in the energy range $[E_{\min}, E_{\max}]$ (see (18.36) and (18.37)). Equation (E.8) can be rewritten as a condition on the group size N

$$N > \beta \, \frac{\left[\delta_a^2\right]_{\max}}{e_{\min} - e_0} \; , \tag{E.9}$$

where $e_{\min} = E_{\min}/(NN^{\mathrm{G}})$, $e_0 = E_0/(NN^{\mathrm{G}})$ and $\delta_a^2 = \varDelta_a^2/N^{\mathrm{G}}$.

We now turn to analyse condition (18.34). Equation (E.6) shows that the $\varDelta_a^2$ do not contain terms that are proportional to E_a. One thus has to determine when the $\varDelta_a^2$ are approximately constant, which is the case if

$$\beta \, \frac{\left[\varDelta_a^2\right]_{\max} - \left[\varDelta_a^2\right]_{\min}}{2} \ll [E_a]_{\max} - [E_a]_{\min} \; . \tag{E.10}$$

As a direct consequence, we get $|c_1| \ll 1$ which means that temperature is intensive.

Defining the quantity $e_a = E_a/(NN^{\mathrm{G}})$, we can rewrite (E.10) as a condition on N,

$$N \geq \frac{\beta}{2\,\varepsilon} \frac{\left[\delta_a^2\right]_{\max} - \left[\delta_a^2\right]_{\min}}{\left[e_a\right]_{\max} - \left[e_a\right]_{\min}} \,, \tag{E.11}$$

where the accuracy parameter $\varepsilon \ll 1$ is equal to the ratio of the left hand side and the right hand side of (E.10).

Since (E.10) does not take into account the energy range (18.36), its application needs some further discussion.

If the occupation number of one mode of a group is changed, say from $n_k^a(\nu) = 0$ to $n_k^a(\nu) = 1$, the corresponding Δ_a^2 differ at most by $4\Delta E^2 K^2/(n+1)$. On the other hand, $\left[\Delta_a^2\right]_{\max} - \left[\Delta_a^2\right]_{\min} = N^{\mathrm{G}} \Delta E^2 K^2$. The state with the maximal Δ_a^2 and the state with the minimal Δ_a^2 thus differ in nearly all occupation numbers and, therefore, their difference in energy is close to $[E_a]_{\max} - [E_a]_{\min}$. On the other hand, states with similar energies E_a also have a similar Δ_a^2. Hence the Δ_a^2 only change quasi-continuously with energy and (E.10) ensures that the Δ_a^2 are approximately constant even locally, i.e., on any part of the possible energy range.

To compute the required group size $N_{\min}$, we need to know the maximal and minimal values E_a and Δ_a^2 can take on. For E_a they are given by

$$\left\{ \begin{matrix} [E_a]_{\max} \\ [E_a]_{\min} \end{matrix} \right\} = \left\{ \begin{matrix} + \\ - \end{matrix} \right\} N^{\mathrm{G}} N \Delta E \,, \tag{E.12}$$

for $|K| \leq 1$ and by

$$\left\{ \begin{matrix} [E_a]_{\max} \\ [E_a]_{\min} \end{matrix} \right\} = \left\{ \begin{matrix} + \\ - \end{matrix} \right\} N^{\mathrm{G}} N \Delta E \frac{2}{\pi} \left[\sqrt{K^2 - 1} + \arcsin\left(\frac{1}{|K|} \right) \right] \,, \tag{E.13}$$

for $|K| > 1$, where the sum over all modes k has been approximated by an integral. The maximal and minimal values of Δ_a^2 are given by

$$\left\{ \begin{matrix} \left[\Delta_a^2\right]_{\max} \\ \left[\Delta_a^2\right]_{\min} \end{matrix} \right\} = \left\{ \begin{matrix} N^{\mathrm{G}}\, K^2\, \Delta E^2 \\ 0 \end{matrix} \right\} . \tag{E.14}$$

Plugging these results into (E.11) as well as (E.1) and (E.3) (taking into account (18.36) and (18.37)) into (E.9) for $|K| > 1$ and using (E.7) for $|K| < 1$, the minimal number of systems per group can be calculated.

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions.* Dover Publ., New York, 9th edition, 1970.

[2] G. Adam and O. Hittmair. *Wärmetheorie.* Vieweg, Braunschweig, Wiesbaden, 4th edition, 1992.

[3] R. Alicki and M. Fannes. *Quantum Dynamical Systems.* Oxford University Press, Oxford, 2001.

[4] A. E Allahverdyan and Th. M Nieuwenhuizen. Extraction of Work from a Single Thermal Bath in the Quantum Regime. *Phys. Rev. Lett.*, 85:1799–1802, 2000.

[5] A. E. Allahverdyan and Th. M. Nieuwenhuizen. Optimizing the Classical Heat Engine. *Phys. Rev. Lett.*, 85:232–235, 2000.

[6] A. E. Allahverdyan and Th. M. Nieuwenhuizen. Testing the Violation of the Clausius Inequality in Nanoscale Electric Circuits. *Phys. Rev. B*, 66:115309, 2002.

[7] H. Araki and E. H. Lieb. Entropy Inequalities. *Commun. Math. Phys.*, 18:160–170, 1970.

[8] P. Ball. Transitions still to be made. *Nature*, 402:C73–C76, 1999.

[9] L. B. Ballentine. *Quantum Mechanics.* World Scientific, Singapore, 2nd edition, 1998.

[10] F. J. Belinfante. *A Survey of Hidden-Variables Theories.* Pergamon Press, Oxford, 1973.

[11] Y. Ben-Menahem and I. Pitowsky. Introduction. *Stud. Hist. Phil. Mod. Phys.*, 32:503–510, 2001.

[12] P. Bergmann and J. Lebowitz. New Approach to Nonequilibrium Processes. *Phys. Rev.*, 99:578–587, 1955.

[13] J. Biggus. Sketching the History of Statistical Mechanics and Thermodynamics. *http://history.hyperjeff.net/statmech.html*, 2002.

[14] G. D. Birkhoff. Proof of the Ergodic Theorem. *Proc. Nat. Acad. Sc. USA*, 17:656–660, 1931.

[15] J. M. Blatt. An Alternative Approach to the Ergodic Problem. *Progr. Theor. Phys.*, 22(6):745–756, 1959.

[16] K. Blum. *Density Matrix Theory and Applications.* Plenum Press, New York, London, 2nd edition, 1996.

[17] P. Bocchieri and A. Loinger. Ergodic Foundation of Quantum Statistical Mechanics. *Phys. Rev.*, 114:948–951, 1959.

[18] L. Boltzmann. *Lectures on Gas Theory.* University of California Press, Los Angeles, 1964.

[19] P. Borowski, J. Gemmer, and G. Mahler. On the Concept of Pressure in Quantum Mechanics. *Europhys. Lett.*, 62:629–635, 2003.

[20] P. Borowski, J. Gemmer, and G. Mahler. Relaxation to Equilibrium under Pure Schrödinger Dynamics. *Eur. Phys. J. B*, 35(2):255–259, 2003.

[21] D. Braun, F. Haake, and W. T. Strunz. Universality of Decoherence. *Phys. Rev. Lett.*, 86:2913–2917, 2001.

[22] W. Brenig. *Statistische Theorie der Wärme*, Volume 1. Springer, Berlin, 4th edition, 1996.

[23] Th. Brody. *The Philosophy behind Physics.* Springer, Berlin, Heidelberg, New York 1993.

[24] H. B. Callen. *Thermodynamics and an Introduction to Thermostatistics.* Wiley, New York, 2nd edition, 1985.

[25] V. Cápek. Zeroth and Second Laws of Thermodynamics Simultaneously Questioned in the Quantum Microworld. *Eur. Phys. J. B*, 25(1):101–113, 2002.

[26] M. C. Caves, C. A. Fuchs, and R. Schack. Quantum Probabilities as Bayesian Probabilities. *Phys. Rev. A*, 65:022305, 2002.

[27] A. Cook. *The Observational Foundations of Physics.* Cambridge University Press, Cambridge, 1994.

[28] J. Y. Dalibard, Y. Castin, and K. Mølmer. Wave Function Approach to Dissipative Processes in Quantum Optics. *Phys. Rev. Lett.*, 68(5):580–583, 1992.

[29] P. Davies. *The Physics of Time Asymmetry.* University of California Press, Berkeley, 1974.

[30] B. Diu, C. Guthmann, D. Lederer, and B. Roulet. *Elements de Physique Statistique.* Hermann Editeurs des Sciences et des Arts, Paris, 1989.

[31] P. Ehrenfest. Adiabatische Invarianten und Quantentheorie. *Ann. Phys.*, 51:327–352, 1916.

[32] P. Ehrenfest and T. Ehrenfest. *The Conceptual Foundations of the Statistical Approach in Mechanics.* Cornell University Press, Ithaca, 1959.

[33] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47:777–780, 1935.

[34] J. Eisert. Exact Decoherence to Pointer States in Free Open Quantum Systems Is Universal. *Phys. Rev. Lett.,* 92:210401, 2004.

[35] T. Feldmann and R. Kosloff. The Quantum Four Stroke Heat Engine: Thermodynamic Observables in a Model with Intrinsic Friction. *quant-ph/0303046*, 2003.

[36] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics.* Addison-Wesley, 1963.

[37] E. Fick and G. Sauermann. *The Quantum Statistics of Dynamic Processes.* Springer, Berlin, Heidelberg, New York, 1990.

[38] C. A. Fuchs. Quantum Foundations in the Light of Quantum Information. *quant-ph/0106166*, 2001.

[39] J. Gemmer and G. Mahler. Entanglement and the Factorization-Approximation. *Euro. Phys. J. D*, 17:385–393, 2001.

[40] J. Gemmer and G. Mahler. Limits of Effective Potential Approach: Nucleus-Electron Entanglement in a Model of the He^{+}-Ion. *Europhys. Lett.*, 59:159–165, 2002.

[41] J. W. Gibbs. *Elementary Principles in Statistical Mechanics.* Dover Publications, New York, 1960.

[42] D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H. D. Zeh. *Decoherence and the Appearance of a Classical World in Quantum Theory.* Springer, Berlin, Heidelberg, New York, 2nd edition, 2003.

[43] P. Gong, J. Brumer. When is Quantum Decoherence Dynamics Classical? *Phys. Rev. Lett.*, 90(5):050402, 2003.

[44] K. Gustafson. Professor Ilya Priogine: A Personal and Scientific Remembrance. *Mind and Matter*, 1:9–13, 2003.

[45] L. Hackermüller, K. Hornberger, B. Brezger, A. Zeilinger, and M. Arndt. Decoherence of Matter Waves by Thermal Emission of Radiation. *Nature*, 427:711–714, 2004.

[46] M. Hartmann, J. Gemmer, G. Mahler, and O. Hess. Scaling Behavior of Interactions in a Modular Quantum System and the Existence of Local Temperature. *Europhys. Lett.*, 65:613–619, 2004.

[47] M. Hartmann, G. Mahler, and O. Hess. On the Existence of Temperature on the Nanoscale. *Phys. Rev. Lett.*, 93:080402, 2004.

[48] M. Hartmann, G. Mahler, and O. Hess. Gaussian Quantum Fluctuations in Interacting Many Particle Systems. *math-ph/0312045*, 2004. acc. for publ. in Lett. Math. Phys.

[49] M. Hartmann, G. Mahler, and O. Hess. Local Versus Global Thermal States: Correlations and the Existence of Local Temperatures. *quant-ph/0404164*, 2004.

[50] M. Hartmann, G. Mahler, and O. Hess. Spectral Densities and Partition Functions of Modular Quantum Systems as Derived from Central Limit Theorem. *cond-mat/0406100*, 2004.

[51] F. Heidrich-Meisner, A. Honecker, D. C. Cabra, and W. Brenig. Thermal Conductivity of One-Dimensional Spin-1/2 Systems. *cond-mat/0406378*, 2004.

[52] F. Heidrich-Meisner, A. Honecker, D.C. Cabra, and W. Brenig. Zero-Frequency Transport Properties of One Dimensional Spin-1/2 Systems. *cond-math/0304595*, 2003.

[53] M. Hemmo and O. Shenker. Can We Explain Thermodynamics by Quantum Decoherence? *Stud. Hist. Phil. Mod. Phys.*, 32:555–568, 2001.

[54] M. Henrich. Dynamische und thermische Eigenschaften abgeschlossener modularer quantenmechanischer Systeme. Master Thesis, Universität Stuttgart, Stuttgart, 2004.

[55] T. Holstein and H. Primakoff. Field Dependence of the Intrinsic Domain Magnetization of a Ferromagnet. *Phys. Rev.*, 58:1098–1113, 1940.

[56] E. Jaynes. Information Theory and Statistical Mechanics. *Phys. Rev.*, 106:620–630, 1957.

[57] E. Jaynes. Information Theory and Statistical Mechanics. II. *Phys. Rev.*, 108:171–190, 1957.

[58] R. Jensen and R. Shankar. Statististical Behavior in Deterministic Quantum Systems with Few Degrees of Freedom. *Phys. Rev. Lett.*, 54:1879–1882, 1985.

[59] E. Joos. Decoherence through Interaction with the Environment. In D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H. D. Zeh, editors, *Decoherence and the Appearance of a Classical World in Quantum Theory*, Chap. 3, page 41. Springer, Berlin, Heidelberg, New York, 2nd edition, 2003.

[60] A. N. Jordan and M. Büttiker. Entanglement Energetics at Zero Temperature. *Phys. Rev. Lett.*, 92(24):247901, 2004.

[61] B.; Glaser S. J. Khaneja, N.; Luy. Boundary of Quantum Evolution Under Decoherence. *Proc. Natl. Acad. Sci. USA*, 100(23):13162–13166, 2003.

[62] Ch. Kittel. *Quantum Theory of Solids.* Wiley, New York, 1963.

[63] J. R. Klauder and E. C. Sudarshan. *Fundamentals of Quantum Optics.* Benjamin, New York, 1968.

[64] C. Kostoglou. Zum Quanten-Carnot-Prozess. Master Thesis, University of Stuttgart, Stuttgart, 2004.

[65] H. J. Kreuzer. *Nonequilibrium Thermodynamics and Its Statistical Foundation.* Clarendon Press, Oxford, 1981.

[66] R. Kubo. Statistical-Mechanical Theory of Irreversible Processes. I. *J. Phys. Soc. Jpn.*, 12:570–586, 1957.

[67] R. Kubo, M. Toda, and N. Hashitsume. *Statistical Physics II: Nonequilibrium Statistical Mechanics.* Number 31 in Solid-State Sciences. Springer, Berlin, Heidelberg, New-York, 2nd edition, 1991.

[68] J. Kupsch. Open Quantum Systems. In D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H.D. Zeh, editors, *Decoherence and the Appearance of a Classical World in Quantum Theory*, Chap. 7, page 317. Springer, Berlin, Heidelberg, New York, 2nd edition, 2003.

[69] L. Landau and E. Lifshitz. *Statistical Physics, Part 1*, Volume 5 of *Course of Theretical Physics.* Pergamon Press, Oxford, 3rd edition, 1980.

[70] L. Landau and I. Lifshitz. *Quantum Mechanics*, Volume 3 of *Course of Theoretical Physics.* Pergamon Press, Oxford, 3rd edition, 1977.

[71] J. L. Lebowitz. Boltzmann's Entropy and Time's Arrow. *Phys. Today*, 46(9):32–38, 1993.

[72] G. Lindblad. On the Generators of Quantum Dynamical Semigroups. *Commun. Math. Phys.*, 48:119–130, 1976.

[73] G. Lindblad. *Non-equilibrium Entropy and Irreversibility*, Volume 5 of *Mathematical Physics Studies.* Reidel, Dordrecht, 1983.

[74] S. Lloyd. Quantum-Mechanical Maxwell's Demon. *Phys. Rev. A*, 56(5):3374–3382, 1997.

[75] R. Loudon. *The Quantum Theory of Light.* Oxford University Press, Oxford, 1985.

[76] E. Lubkin. Entropy of an n-System from Its Correlation with a k-Reservoir. *J. Math. Phys.*, 19:1028–1031, 1978.

[77] J. M. Luttinger. Theory of Thermal Transport Coefficients. *Phys. Rev.*, 135(6A):A1505–A1514, 1964.

[78] G. D. Mahan. *Many-Particle Physics.* Plenum Press, New York, London, 3rd edition, 2000.

[79] G. Mahler and V.A. Weberruß. *Quantum Networks.* Springer, Berlin, Heidelberg, New York, 2nd edition, 1998.

[80] F. Mandl. *Statistical Physics.* Wiley, London, New York, 1971.

[81] H. Michel and M. Wagner. Non-thermal Simulation of Three-Level Impurities by Phonons. *Phys. Status Solidi B*, 75:507–517, 1976.

[82] H. Michel and M. Wagner. On the Use of Generalized Langevin Equations in Phonon Transport Problems. *Phys. Status Solidi B*, 85:195–201, 1978.

[83] H. Michel and M. Wagner. Phonon Scattering at Multi-Level Impurities. *Ann. Phys.*, 35:425–440, 1978.

[84] M. Michel, M. Hartmann, J. Gemmer, and G. Mahler. Fourier's Law Confirmed for a Class of Small Quantum Systems. *Eur. Phys. J. B*, 34:325–330, 2003.

[85] A. Münster. *Statistische Thermodynamik.* Springer, Berlin, Heidelberg, New York, 1956.

[86] A. Münster. *Prinzipien der statistischen Mechanik*, Volume III/2 of *Handbuch der Physik.* Springer, Berlin, Heidelberg, New York, 1959.

[87] J. von Neumann. Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik. *Z. Phys.*, 57:30–70, 1929.

[88] J. von Neumann. Proof of the Quasi-ergodic Hypothesis. *Proc. Nat. Acad. Sci. USA*, 18:70–82, 1932.

[89] J. von Neumann. *Mathematischen Grundlagen der Quantenmechanik.* Die Grundlehren der Mathematischen Wissenschaften. Springer, Berlin, Heidelberg, New York, 1968.

[90] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000.

[91] Th. M. Nieuwenhuizen and A. E. Allahverdyan. Statistical Thermodynamics of Quantum Brownian Motion: Construction of Perpetuum Mobile of the Second Kind. *Phys. Rev. E*, 66:036102, 2002.

[92] A. Otte. *Separabilität in Quantennetzwerken.* PhD Thesis, Universität Stuttgart, Stuttgart, 2001.

[93] A. Otte and G. Mahler. Adapted-Operator Representations: Selective Versus Collective Properties of Quantum Networks. *Phys. Rev. A*, 62:012303, 2000.

[94] D. N. Page. Average Entropy of a Subsystem. *Phys. Rev. Lett.*, 71(9):1291–1294, 1993.

[95] R. K. Pathria. *Statistical Mechanics.* Pergamon Press, Oxford, 1972.

[96] W. Pauli and Z. Fierz. Über das H-Theorem in der Quantenmechanik. *Z. Phys.*, 106:572–587, 1937.

[97] H. Paus. *Physik in Experimenten und Beispielen.* Hanser, München, Wien, 2nd edition, 2002.

[98] R. Peierls. On the Kinetic Theory of Thermal Conduction in Crystals. *Ann. Phys. (N. Y.)*, 3:1055–1101, 1929.

[99] R. Peierls. Zur kinetischen Theorie der Wärmeleitung in Kristallen. *Ann. Phys. (Leipzig)*, 395:1055–1101, 1929. Folge 5 Nr. 3.

[100] R. E. Peierls. *Quantum Theory of Solids.* Clarendon Press, Oxford, 2001.

[101] R. Penrose. *The Emperor's New Mind.* Oxford University Press, New York, 1990.

[102] R. Penrose. Wavefunction Collapse as a Real Gravitational Effect. In A. Fokas, A. Grigoryan, T. Kibble, and B. Zegarlinski, editors, *Mathematical Physics.* Imperial College, London, 2000.

[103] A. Peres. What's Wrong with These Observables? *Found. Phys.*, 33(10):1543–1547, 2003.

[104] H. Pick, J. M. Spaeth, and H. Paus. *Kursvorlesung Experimentalphysik I.* Universität Stuttgart, 1969.

[105] H. Primas. Time-Entanglement between Mind and Matter. *Mind and Matter*, 1(1):81–119, 2003.

[106] E. G. Raine, D. J.; Thomas. *An Introduction to the Science of Cosmology.* IOP Publishing, London, 2001.

[107] G. Rickayzen and J. G. Powles. Temperature in the Classical Microcanonical Ensemble. *J. Chem. Phys.*, 114:4333–4334, 2001.

[108] H. Römer and T. Filk. *Statistische Mechanik.* VCH, Weinheim, New York, Basel, Cambridge, Tokyo, 1994.

[109] R. Rucker. *Mind Tools: The Five Levels of Mathematical Reality.* Mariner Books, 1988.

[110] H. H. Rugh. Dynamical Approach to Temperature. *Phys. Rev. Lett.*, 78:772–774, 1997.

[111] K. Saito. Strong Evidence of Normal Heat Conduction in a One-Dimensional Quantum System. *Europhys. Lett.*, 61:34–40, 2003.

[112] K. Saito, S. Takesue, and S. Miyashita. Thermal Conduction in a Quantum System. *Phys. Rev. E*, 54:2404–2408, 1996.

[113] L. I. Schiff. *Quantum Mechanics.* McGraw-Hill, Duesseldorf, 3rd edition, 1968.

[114] F. Schlögl. *Probability and Heat.* Vieweg, Braunschweig, Wiesbaden, 1989. p 217 ff.

[115] J. Schnack. Thermodynamics of the Harmonic Oscillator Using Coherent States. *Eurphys. Lett.*, 45:647–652, 1999.

[116] E. Schrödinger. *Statistical Thermodynamics.* A Course of Seminar Lectures, Jan.–March 1944. Cambridge University Press, Cambridge, 1948.

[117] M. O. Scully. Quantum Afterburner: Improving the Efficiency of an Ideal Heat Engine. *Phys. Rev. Lett.*, 88:050602, 2002.

[118] M. O. Scully and M. S. Zubairy. *Quantum Optics.* Cambridge University Press, Cambridge, 1997.

[119] M. O. Scully, M. S. Zubairy, G. S. Agarwal, and H. Walther. Extracting Work from a Single Heat Bath via Vanishing Quantum Coherence. *Science*, 299:862–864, 2003.

[120] D. P. Sheehan, editor. *Quantum Limits of the 2nd Law*, San Diego, CA, 28-31 July 2002. AIP Conference Proceedings 643, New York 2002. First Int'l Conf. on Quantum Limits to the 2nd Law.

[121] C. P. Snow. *The Two Cultures and the Scientific Revolution.* Cambridge University Press, New York, 1959.

[122] H. P. Stapp. *Mind, Matter, and Quantum Mechanics.* Springer, Berlin, Heidelberg, New York, 2nd edition, 2004.

[123] L. Szilard. Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Z. Phys.*, 53:840–856, 1929.

[124] H. Tasaki. From Quantum Dynamics to the Canonical Distribution: General Picture and a Rigorous Example. *Phys. Rev. Lett.*, 80:1373–1376, 1998.

[125] H. Tasaki. Inevitable Irreversibility in a Quantum System Consisting of Many Non-interacting "Small" Pieces. */cond-mat/0008420*, 2000.

[126] W. R. Theis. *Grundzüge der Quantentheorie.* Teubner, Stuttgart, 1985.

[127] M. Toda, R. Kubo, and N. Saito. *Statistical Physics I: Equilibrium Statistical Mechanics.* Number 30 in Solid-State Sciences. Springer, Berlin, Heidelberg, New York, 2nd edition, 1978.

[128] Tribus and McIrvine. Energy and Information. *Sci. Am.*, 225(9):179–189, 1971.

[129] D. F. Walls and G. J. Milburn. *Quantum Optics.* Springer, Berlin, Heidelberg, New York, 2nd edition, 1995.

[130] X. Wang. Threshold Temperature for Pairwise and Many-Particle Thermal Entanglement in the Isotropic Heisenberg Model. *Phys. Rev. A*, 66:044305, 2002.

[131] W. Weidlich. *Thermodynamik und statistische Mechanik.* Akademische Verlagsgesellschaft, Wiesbaden, 1976.

[132] E. P. Wigner. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Commun. Pure Appl. Math.*, 13:1, 1960.

[133] J. M. Ziman. *Electrons and Phonons.* Oxford: Clarendon Press, Oxford, 2001.

[134] W. H. Zurek. Decoherence and the Transition from Quantum to Classical. *Physics Today*, 41:Oct 36–44, 1991.

[135] W. H. Zurek. Decoherence, Einselection, and the Quantum Origin of the Classical. *Rev. Mod. Phys.*, 75:715–775, 2003.

[136] W. H. Zurek and J. P. Paz. Decoherence, Chaos, and the Second Law. *Phys. Rev. Lett.*, 72:2508–2511, 1994.

[137] R. Zwanzig. Memory Effects in Irreversible Thermodynamics. *Phys. Rev.*, 124:983–992, 1961.

Index