Pohl's Introduction to Physics

Klaus Lüders · Robert O. Pohl
Editors

Pohl's Introduction to Physics

Mechanics, Acoustics and Thermodynamics, Vol. 1

Springer

Editors

Klaus Lüders
Fachbereich Physik
Freie Universität Berlin
Berlin, Germany

Robert O. Pohl
Department of Physics
Cornell University
Ithaca, NY, USA

Translated by
Prof. William D. Brewer, PhD,
Fachbereich Physik,
Freie Universität Berlin,
Berlin, Germany

Additional material to this book can be downloaded from
http://extra.springer.com.

ISBN 978-3-319-40044-0   ISBN 978-3-319-40046-4 (eBook)
DOI 10.1007/978-3-319-40046-4

Library of Congress Control Number: 2017944730

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface to the Second English Edition

The first English edition of Pohl's "*Physical Principles of Mechanics and Acoustics*" appeared in 1932 (published by Blackie & Son, Ltd., London and Glasgow). It was based on the second edition of Pohl's "*Einführung in die Physik, Mechanik und Akustik*" (Julius Springer, 1931). The present new, second English edition, based on the 21st edition of "*Pohls Einführung in die Physik*", Vol. 1, (*Mechanik, Akustik und Wärmelehre*) (Springer Spektrum, 2017), has now been published after nearly 85 years!

Following R.W. Pohl's death in 1976 and the posthumous appearance of the 18th edition in 1983, since 2004 we have edited three new revised and updated editions, based mainly on the 16th edition (1964), the 13th edition (1955), and the 18th edition. The major change in this new series was the addition of 74 videos demonstrating many of the experiments that R.W. Pohl had developed and used. It is also augmented by comments in the margins when they appeared to be helpful as additional explanations or were needed to provide more modern information (see the Preface to the 19th edition). We have in addition included the collection of exercises which Pohl provided for the first English edition, and we have supplemented these (see the Preface to the 20th edition). The exercises were not included in any of the first 18 German editions. We have furthermore modified some mathematical formulations, symbols, and units so that they conform to the recommendations of the International System of Units (SI).

We gratefully acknowledge the help of Professor W.D. Brewer of the Physics Department of the Free University of Berlin, not only for carrying out the translation of the text with great quality and speed, but also, and this is probably even more important, for his help with the identification and clarification of unclear parts in the text and in our comments. The English-language readers will appreciate the numerous links he added for further information.

We also wish to thank Dr. T. Schneider and Ms. D. Mennecke-Buehler of the Springer-Verlag for making this edition possible, and for their generous help in carrying out its preparation and production.

Berlin and Göttingen, March 2017

K. Lüders
R. O. Pohl

Preface to the Twenty-First German Edition

One of the most extensive changes in this new edition concerns its format, and is intended to make the books more readable. “Pohl” will now be published for the first time as an e-book, but also as a printed version with a new format. The numbering of the chapters, figures, equations etc. now conforms to the usual system in modern textbooks. The relevant exercises are given at the end of each chapter.

We have also made major changes to the accompanying videos. In the e-book format, they are now more readily accessible and can be called up directly using the appropriate links in the text. All those videos which were produced in cooperation with the Institute for Scientific Films (IWF) in Göttingen are now available in their original quality and with a spoken text. The remaining videos were to some extent supplemented and in one case replaced by an improved version. Two new videos have been added: “Kepler Ellipses” (an excerpt from the opening show of the 2009 “*Highlights der Physik*” in Cologne), and “The Magdeburg Hemispheres” (an excerpt from a Lichtenberg Lecture given by Prof. G. Beuermann in Göttingen).

At the same time, we have taken advantage of the opportunity to review all of the text critically. This has led to a number of clarifications, both in the text and in the figures, including the addition of several new figures, notably in Chapters 12 and 19.

We owe special thanks to Prof. K. Samwer from the First Physics Institute of the University of Göttingen for his committed and helpful support of the preparation of this new edition in a variety of ways. We also wish to give particular thanks to Prof. G. Beuermann, J. Feist and C. Mahn from that Institute for their various and dedicated assistance. Furthermore, we wish especially to thank Dr. J. Kirstein from the Physics Didactics group at the Free University of Berlin for his professional and speedy editing of the videos. We are once again indebted to the Physics Department of the Free University and its administration for providing working facilities and for the helpful efforts of many of its members in solving technical problems, especially in connection with computer technology. Finally, we heartily thank the Springer-Verlag, and in particular Dr. V. Spillner, Ms. M. Maly and Ms. B. Saglio for their stimulating and agreeable cooperation.

Berlin and Göttingen, July 2015

K. Lüders
R. O. Pohl

From the Preface to the 20th Edition (2008)

The many positive comments from readers of the 19th edition of POHL's *Introduction to Mechanics, Acoustics and Thermodynamics* have encouraged us to publish a new, revised edition. This also gave us the opportunity to include some supplementary material which we believe to be important. In addition to new or revised marginal comments and a few factual clarifications within the text, this new material consists in particular of additional videos and a set of exercises for the readers. Also, the sections on osmosis and diffusion from earlier editions have now been included here.

This time, the videos were filmed under our own direction in the new lecture hall in Göttingen, in addition to several filmed in cooperation with the Physics Didactics group at the Free University in Berlin. In choosing the topics, we have again been guided on the one hand by our attempt to present 'lively' illustrations of physics, and on the other by our intention to document typical demonstration experiments in the tradition of POHL, which in some cases are no longer being shown even in Göttingen.

The major portion of the exercises originates with an earlier English-language edition (from 1932!); they are thus POHL's original exercises. However, we found it desirable to add some more exercises which deal with questions that either relate directly to the videos or illustrations, or that complement the experiments, which are sometimes described rather briefly in the text due to lack of space. These exercises are thus not problem sets in the usual sense, but rather they are intended to help the reader achieve a better understanding of the sometimes difficult physical concepts described in this volume, and furthermore they provide additional information.

From the Preface to the 19th Edition (2004)

For over thirty years, from 1919 to 1952, R.W. POHL gave the introductory lectures in experimental physics at the University of Göttingen for students of a variety of major subjects. The three-volume set of textbooks based on those lectures pursued a double goal for many years: On the one hand, they were intended to arouse the readers' interest in physics; and on the other, they served as textbooks for teaching basic physics to interested students. Even though in more recent decades, physics education at the university level has adjusted more and more to the needs of various professions and now includes many specialized courses, the goals of POHL's works are still valid and topical. We are therefore convinced that these books still convey a fascination for the experimental investigation of physical phenomena and deserve a place on the bookshelves of modern-day students. That is the reason for the present new edition, initially covering the fields of mechanics, acoustics and thermodynamics. A second volume will present the most important topics from electrodynamics and optics.

For many readers, the most noticeable characteristic of POHL's books is the large number of experiments which they illustrate and describe in detail; these demonstrate how one must ask questions of Nature in order to uncover her secrets. The presentation of demonstration experiments using shadow projections, which fix the attention of the observer on the essentials of the demonstration, is an integral part of this program. But in addition, we want to provide readers with the opportunity to experience the demonstrations just as they have been presented in the Göttingen lecture hall for more than 80 years. For this reason, we have complemented this edition with two CD-ROMs containing short videos. The first of these is an original film of a lecture delivered by R.W. POHL in 1952 (**Video 1**).[1] We hope that our readers will enjoy watching these videos as much as we have enjoyed filming them.

In order to retain the liveliness of the "POHLs", as the books are often called, it seemed important to us to maintain the manner of presentation of their original author as nearly as possible. Since, however, the first volume alone was available in no fewer than fourteen different editions, we had to make choices. This book is based mainly on the 16th edition, which appeared in 1964. Occasionally, however, we refer to other editions, in particular the 13th (1955) and the 18th (1983).

---

[1] **Video 1:**
"R.W. POHL Lecturing"
http://tiny.cc/fpqujy
This film, shot by FRITZ LUETY (now professor emeritus at the University of Utah in Salt Lake City) while he was a graduate student in 1952, for the summer celebration of the Göttingen physics institute, shows a lecture on oscillatory motion given by POHL, with several demonstration experiments that are described in Chap. 11 of this book.

We have as far as possible avoided making changes to the text. Among the exceptions are our more frequent use of vectors and integrals, i.e. mathematical objects with which today's readers are in general familiar. Furthermore, we have adapted the symbols and units to modern usage, in order to spare our readers the unnecessary annoyance of conversion. Our own attempts at enriching the text are limited to comments in the margins, which contain both direct explanations of material in the text and references to newer developments in the areas of physics treated.

From the Preface to the First Edition (1930)

This book contains the first part of my lectures on experimental physics. An effort has been made to present them as simply as possible. This is intended to make the book accessible not only to students and teachers, but also to other readers with an interest in physics.

Basic experiments occupy the most prominent place in the presentation. They serve in particular to clarify the concepts and to provide an overview of the magnitudes of the quantities involved. Quantitative details are not emphasized.

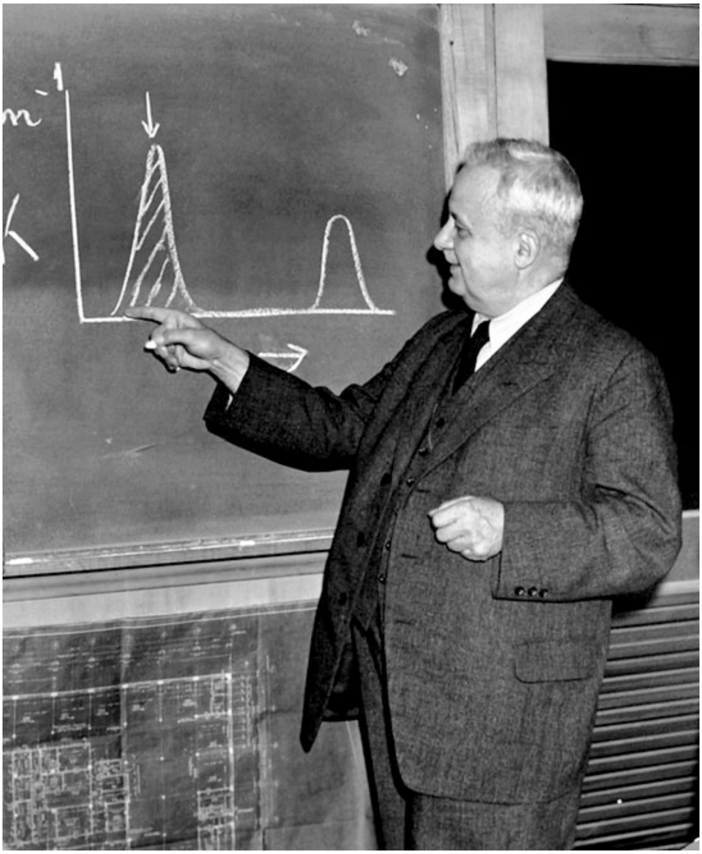
A large collection of demonstration experiments occupies considerable space. In our lecture hall in Göttingen, we have a smooth-floored area of $12 \times 5\,\mathrm{m}^2$. The cumbersome accessory of earlier lecture halls, i.e. the heavy, stationary demonstration table, has long since been dispensed with. Instead, smaller tables are set up as needed, and they are no more anchored to the floor than is the furniture in a living room. The clarity of the experimental arrangement and the accessibility of the individual experimental setups are enhanced considerably by the use of these convenient tables. Most of them can be rotated around their vertical axis and they are readily adjustable in height. Thus, the annoying overlap of perspective between different setups can be avoided. The setup currently being demonstrated can be highlighted and made visible to every member of the audience by panning the tables.

The apparatus used is as simple as possible and consists of a moderate number of devices. Many of the setups are described here for the first time. They can be obtained, as can other accessories for lecture demonstrations, from the Spindler & Hoyer company in Göttingen.

The main portion of the illustrations in the book are based on photographs. Many of the images are presented as silhouettes. This method of presentation is especially suitable for reproduction in book form; in addition, it often provides some indication of the dimensions of the experimental setup. Finally, showing the experiments as silhouettes makes them visible even in large lecture halls, which demand clear-cut outlines, not interrupted by incidental details such as laboratory stands, frames etc.

Göttingen, March 1930 R. W. Pohl

R.W. Pohl (1884–1976)



R.W. POHL (1884–1976) discussing color centers (F-centers), elementary crystal lattice defects which were discovered at his institute and investigated there for many years. He is shown during a visit to the Ansco Research Laboratory in Binghamton, NY in the year 1951. Details of POHL's life and work can be found on the website http://rwpohl.mpiwg-berlin.mpg.de of the Max Planck Institute for the History of Science (MPIWG). There, one can find links to other literature, scientific institutions and websites which offer information and documents on the teaching and research of the famous physicist in Göttingen. In addition, the documentary video "Simplicity is the Mark of Truth" by Ekkehard SIEKER (Video 1 from Vol. 2) can be found on the MPIWG web site, together with all the other videos from both volumes and other audiovisual materials, available both for videostreaming or as downloads.

Contents

Part I Mechanics

1 Introduction; Distance and Time Measurements . . . 3
1.1 Introduction . . . 3
1.2 Distance and Length Measurements. Direct Distance Measurements 4
1.3 The Meter as a Unit of Length . . . 6
1.4 Indirect Length Measurements of Very Large Distances . . . 7
1.5 Angle Measurements . . . 9
1.6 Time Determinations. True Time Measurements . . . 10
1.7 Clocks and Graphical Registration . . . 11
1.8 Measurement of Periodic Sequences of Equal Times and Lengths . . 13
1.9 Indirect Time Measurements . . . 15
Exercises . . . 16

2 The Description of Motion: Kinematics . . . 17
2.1 Definition of Motion. Frames of Reference . . . 17
2.2 Definition of Velocity. Example of a Velocity Measurement . . . 18
2.3 Definition of Acceleration: The Two Limiting Cases . . . 20
2.4 Path Acceleration and Linear Motion . . . 22
2.5 Constant Radial Acceleration and Circular Orbits . . . 26
2.6 Distinguishing Physical Quantities and Their Numerical Values . . . 28
2.7 Base Quantities and Derived Quantities . . . 29
Exercises . . . 30

3 Fundamentals of Dynamics . . . 31
3.1 Force and Mass . . . 31
3.2 Measurements of Force and Mass. NEWTON's Fundamental Equation of Motion . . . 34

3.3 The Units of Force and Mass. Expressions Containing Physical Quantities . . . 38
3.4 Density and Specific Volume . . . 38
Exercises . . . 39

4 Applications of NEWTON's Equation . . . 41
4.1 Constant Acceleration in a Straight Line . . . 41
4.2 Circular Motion and Radial Forces . . . 44
4.3 Sinusoidal Oscillations: The Gravity Pendulum as a Special Case . . . 49
4.4 Motions Around a Central Point . . . 54
4.5 Elliptical Orbits and Elliptically Polarized Oscillations . . . 56
4.6 LISSAJOUS Orbits . . . 57
4.7 KEPLER's Elliptical Orbits and the Law of Gravity . . . 58
4.8 The Gravitational Constant . . . 60
4.9 The Law of Gravity and Celestial Mechanics . . . 62
Exercises . . . 65

5 Three Useful Concepts: Work, Energy, and Momentum . . . 67
5.1 Preliminary Remarks . . . 67
5.2 Work and Power . . . 67
5.3 Energy and Its Conservation . . . 71
5.4 First Applications of the Conservation Law of Mechanical Energy . . . 74
5.5 Impulse and Momentum . . . 75
5.6 Momentum Conservation . . . 76
5.7 First Applications of Momentum Conservation . . . 77
5.8 Momentum and Energy Conservation During Elastic Collisions of Objects . . . 79
5.9 Momentum Conservation in Inelastic Collisions and the Ballistic Pendulum . . . 80
5.10 Non-Central Collisions . . . 82
5.11 Motions Against Dissipative Forces . . . 82
5.12 The Production of Forces with and without Consuming Power . . . 86
5.13 Closing Remarks . . . 87
Exercises . . . 88

6 Rotational Motion of Rigid Bodies . . . 91
6.1 Introductory Remarks . . . 91
6.2 Definition of Torque . . . 91
6.3 The Production of Known Torques, the Constant D^*, and the Angular Velocity ω . . . 94
6.4 The Moment of Inertia, Equation of Motion for Rotations, Torsional Oscillations . . . 97
6.5 The Physical Pendulum and the Beam Balance . . . 102
6.6 Angular Momentum . . . 104
6.7 Free Axes . . . 108
6.8 Free Axes of Humans and Animals . . . 111
6.9 Definition of the Spinning Top and Its Three Axes . . . 112
6.10 The Nutation of a Force-Free Top and Its Fixed Spin Axis . . . 115
6.11 Tops Acted on by Torques. Precession of the Angular-Momentum Axis . . . 117
6.12 Precession Cone with Nutation . . . 121
6.13 A Top with Only Two Degrees of Freedom . . . 123
Exercises . . . 125
7 Accelerated Frames of Reference . . . 129
7.1 Preliminary Remarks. Inertial Forces . . . 129
7.2 Frames of Reference with Only Path Acceleration . . . 130
7.3 Frames of Reference with Radial Acceleration. Centrifugal and CORIOLIS Forces . . . 133
7.4 Vehicles as Accelerated Frames of Reference . . . 141
7.5 The Gravity Pendulum as a Plumb Bob in Accelerated Vehicles . . . 144
7.6 Earth as an Accelerated Frame of Reference. Centrifugal Acceleration of Bodies at Rest . . . 146
7.7 Earth as an Accelerated Frame of Reference. CORIOLIS Force on Moving Bodies . . . 148
Exercises . . . 151
8 Some Properties of Solids . . . 153
8.1 Preliminary Remarks . . . 153
8.2 Elastic Deformation, Flow and Solidification . . . 153
8.3 HOOKE's Law and POISSON's Relation . . . 155
8.4 Shear Stress . . . 156

8.5 Normal, Shear and Principal Stress 157

8.6 Bending and Twisting (Torsion) 160

8.7 Time Dependence of Deformation.
Elastic Aftereffects and Hysteresis 165

8.8 Rupture Strength and Specific Surface Energy of Solids 167

8.9 Sticking and Sliding Friction 169

8.10 Rolling Friction 172

Exercises 173

9 Liquids and Gases at Rest 175

9.1 The Free Displacements of Liquid Molecules 175

9.2 Pressure in Liquids. Manometers 178

9.3 The Isotropy of Pressure and Its Applications 179

9.4 The Pressure Distribution in a Gravitational Field. Buoyancy 182

9.5 Cohesion of Liquids: Tensile Strength,
Specific Surface Energy, and Surface Tension 185

9.6 Gases as Low-Density Liquids Without Surfaces.
BOYLE's Law 193

9.7 A Model Gas. Pressure Due to Random Molecular Motions
(Thermal Motion) 195

9.8 The Fundamental Equation of the Kinetic Theory of Gases.
Velocity of the Gas Molecules 196

9.9 The Earth's Atmosphere. Atmospheric Pressure in Demonstration
Experiments 198

9.10 The Pressure Distribution of Gases in the Gravitational Field.
The Barometric Pressure Formula 202

9.11 Static Buoyancy in Gases 205

9.12 Gases and Liquids in Accelerated Frames of Reference 206

Exercises 208

10 Motions in Liquids and Gases 211

10.1 Three Preliminary Remarks 211

10.2 Internal Friction and Boundary Layers 211

10.3 Laminar Flow: Fluid Motions Which Occur when Friction Plays
a Decisive Role 214

10.4 The REYNOLDS Number 217

10.5 Frictionless Fluid Motion and BERNOULLI's Equation 220

10.6 Flow Around Obstacles. Sources and Sinks. Irrotational or Potential Flows 226

10.7 Rotations of Fluids and Their Measurement. The Irrotational Vortex Field 228

10.8 Vortices and Separation Surfaces in Nearly Frictionless Fluids 232

10.9 Flow Resistance and Streamline Profiles 234

10.10 The Dynamic Transverse Force or Lift 237

10.11 Applications of the Transverse Force 241

Exercises 244

Part II Vibrations and Waves

11 Vibrations 247

11.1 Preliminary Remarks 247

11.2 Producing Undamped Vibrations 247

11.3 The Synthesis of Non-Sinusoidal Periodic Processes from Sine Curves 251

11.4 The Spectral Representation of Complex Oscillatory Processes 256

11.5 Elastic Transverse Vibrations of Linear Solid Bodies Under Tensile Stress 259

11.6 Elastic Longitudinal and Torsional Vibrations of Stressed Linear Solid Bodies 263

11.7 Elastic Vibrations in Columns of Liquids and Gases 265

11.8 Normal Modes of Stiff Linear Bodies 269

11.9 Normal Modes of 2-Dimensional and 3-Dimensional Bodies. Thermal Vibrations 271

11.10 Forced Oscillations 273

11.11 Energy Transfer Stimulated by Resonance 278

11.12 The Importance of Resonance for the Detection of Pure Sinusoidal Oscillations. Spectral Apparatus 279

11.13 The Importance of Forced Oscillations for Distortion-Free Recording of Non-Sinusoidal Oscillations 281

11.14 The Amplification of Oscillations 282

11.15 Two Coupled Oscillators and Their Forced Oscillations 283

11.16 Damped and Undamped Wobble Oscillations 286

11.17 Relaxation (or Toggle) Oscillations 288

Exercises 289

12 Travelling Waves and Radiation . . . 291

12.1 Travelling Waves . . . 291
12.2 The Doppler Effect . . . 293
12.3 Interference . . . 295
12.4 Interference with Two Slightly Different Source Frequencies . . . 296
12.5 Standing Waves . . . 297
12.6 The Propagation of Travelling Waves . . . 299
12.7 Reflection and Refraction . . . 302
12.8 Image Formation . . . 303
12.9 Total Reflection . . . 304
12.10 Shockwaves when the Wave Velocity Is Exceeded . . . 307
12.11 Huyghens' Principle . . . 308
12.12 Model Experiments on Wave Propagation . . . 309
12.13 Quantitative Results for Diffraction by a Slit . . . 311
12.14 Fresnel's Zone Construction . . . 315
12.15 Narrowing of the Interference Fringes by a Lattice Arrangement of the Wave Sources . . . 317
12.16 Interference of Wave Trains of Limited Length . . . 321
12.17 The Production of Longitudinal Waves and Their Velocities . . . 321
12.18 High-Frequency Longitudinal Waves in Air. The Acoustic Replica Method . . . 323
12.19 The Radiation Pressure of Sound. Sound Radiometers . . . 326
12.20 Reflection, Refraction, Diffraction and Interference of 3-Dimensional Waves . . . 328
12.21 The Origin of Waves on Liquid Surfaces . . . 336
12.22 Dispersion and the Group Velocity . . . 341
12.23 The Excitation of Waves by Aperiodic Processes . . . 345
12.24 The Energy of a Sound Field. The Wave Resistance for Sound Waves 347
12.25 Sound Sources . . . 351
12.26 Aperiodic Sound Sources and Supersonic Velocities . . . 354
12.27 Sound Receivers . . . 355
12.28 The Sense of Hearing . . . 356
12.29 Phonometry . . . 360
12.30 The Human Ear . . . 362
Exercises . . . 365

Part III Thermodynamics

13 Fundamentals . . . 369
13.1 Preliminary Remarks. Definition of the Concept 'Amount of Substance' . . . 369
13.2 The Definition and Measurement of Temperature . . . 370
13.3 The Definitions of the Concepts of Heat and Heat Capacity . . . 374
13.4 Latent Heat . . . 377
Exercise . . . 380

14 The First Law and the Equation of State of Ideal Gases . . . 381
14.1 Work of Expansion and Technical Work . . . 381
14.2 Thermal State Variables . . . 384
14.3 The *Internal Energy U* and the First Law . . . 384
14.4 The State Function *Enthalpy*, *H* . . . 386
14.5 The Two Specific Heats, c_p and c_V . . . 388
14.6 The Thermal Equation of State of Ideal Gases. Absolute Temperatures . . . 390
14.7 Addition of Partial Pressures . . . 394
14.8 The Caloric Equations of State of Ideal Gases. GAY-LUSSAC's Throttle Experiment . . . 395
14.9 Changes of State of Ideal Gases . . . 398
14.10 Applications of Polytropic and Adiabatic Changes of State. Measurements of κ . . . 403
14.11 Pneumatic Motors and Gas Compressors . . . 406
Exercises . . . 407

15 Real Gases . . . 409
15.1 Phase Changes of Real Gases . . . 409
15.2 Distinguishing the Gas from the Liquid . . . 411
15.3 The VAN DER WAALS Equation of State for Real Gases . . . 414
15.4 The JOULE-THOMSON Throttle Experiment . . . 416
15.5 The Production of Low Temperatures and Liquefaction of Gases . . . 418
15.6 Technical Liquefaction Processes and the Separation of Gases . . . 420
15.7 Vapor Pressure and Boiling Temperature. The Triple Point . . . 421
15.8 Hindrance of the Phase Transition Liquid → Solid: Supercooled Liquids . . . 424
15.9 Hindrance of the Phase Transition Liquid ↔ Gas: The Tensile Strength of Liquids . . . 425
Exercise . . . 426

16 Heat as Random Motion . . . 429
16.1 Temperature on the Molecular Scale . . . 429
16.2 The Recoil of Gas Molecules Upon Reflection. The "Radiometer Force" . . . 433
16.3 The Velocity Distribution and the Mean Free Path of the Gas Molecules . . . 435
16.4 Molar Heat Capacities in a Molecular Picture. The Equipartition Principle . . . 437
16.5 Osmosis and Osmotic Pressure . . . 440
16.6 The Experimental Determination of BOLTZMANN's Constant *k* from the Barometric Equation . . . 445
16.7 Statistical Fluctuations and the Particle Number . . . 447
16.8 The BOLTZMANN Distribution . . . 449

17 Transport Processes: Diffusion and Heat Conduction . . . 453
17.1 Preliminary Remarks . . . 453
17.2 Diffusion and Mixing . . . 453
17.3 FICK's First Law and the Diffusion Constant . . . 454
17.4 Quasi-Stationary Diffusion . . . 457
17.5 Non-Stationary Diffusion . . . 459
17.6 General Considerations on Heat Conduction and Heat Transport . . . 460
17.7 Stationary Heat Conduction . . . 463
17.8 Non-Stationary Heat Conduction . . . 464
17.9 Transport Processes in Gases and Their Lack of Pressure Dependence . . . 465
17.10 Determination of the Mean Free Path . . . 468
17.11 The Mutual Relations of Transport Processes in Gases . . . 470

18 The State Function Entropy, *S* . . . 475
18.1 Reversible Processes . . . 475
18.2 Irreversible Processes . . . 477
18.3 Measurement of the Irreversibility Using the State Function *entropy* . . . 479
18.4 Entropy in a Molecular Picture . . . 482
18.5 Examples of the Calculation of the Entropy . . . 483
18.6 Application of Entropy to Reversible Changes of State in Closed Systems . . . 487
18.7 The *H-S* or MOLLIER Diagram with Applications. Supersonic Gas Jets . . . 488
Exercise . . . 493

19 Converting Heat into Work. The Second Law . . . 495

19.1 Heat Engines and the Second Law . . . 495

19.2 The CARNOT Cycle . . . 497

19.3 The STIRLING Engine . . . 498

19.4 Technical Heat Engines . . . 500

19.5 Heat Pumps (Refrigeration Devices) . . . 501

19.6 The Thermodynamic Definition of Temperature . . . 504

19.7 Pneumatic Motors. Free and Bound Energy . . . 504

19.8 Examples of Applications of the Free Energy . . . 506

19.9 The Human Body as an Isothermal Engine . . . 509

Exercise . . . 510

Table of Physical Constants . . . 511

Solutions to the Exercises . . . 512

Index . . . 518

List of Videos

All of the videos for this volume, and also those for Vol. 2, may be downloaded from http://extra.springer.com.

1 R.W. POHL Lecturing
http://tiny.cc/fpqujy . . . . viii, 261, 264, 267, 274, 282, 283

Part I Mechanics

2.1 "Free fall"
http://tiny.cc/jpqujy 23, 24

3.1 "Action = Reaction"
http://tiny.cc/ppqujy 33

4.1 "MAXWELL's wheel"
http://tiny.cc/9pqujy 43

4.2 "Dynamic stability of a bicycle chain"
http://tiny.cc/cqqujy 49

4.3 "Circular oscillations"
http://tiny.cc/1pqujy 57

4.4 "LISSAJOUS figures"
http://tiny.cc/6pqujy 58

4.5 "KEPLER's elliptical orbits"
http://tiny.cc/tpqujy 59

5.1 "Dancing steel ball"
http://tiny.cc/3qqujy 75

5.2 "Conservation of linear momentum"
http://tiny.cc/tqqujy 77, 78

5.3 "Elastic collisions in slow motion"
http://tiny.cc/xqqujy 79

5.4 "Elastic collisions"
http://tiny.cc/xrqujy 80, 89

5.5 "Measuring the velocity of a bullet"
http://tiny.cc/qrqujy 81

5.6 “Non-central collisions”
http://tiny.cc/jqqujy 82

6.1 “Twisting a steel bar”
http://tiny.cc/csqujy 96, 163, 173

6.2 “Moments of inertia”
http://tiny.cc/gsqujy 102

6.3 “Conservation of angular momentum using a rotating chair”
http://tiny.cc/4rqujy 105

6.4 “Angular momentum as a vector”
http://tiny.cc/ltqujy 107

6.5 “The physics of swinging”
http://tiny.cc/ssqujy 108

6.6 “Rotation around free axes”
http://tiny.cc/itqujy 109

6.7 “Supple shaft as stable axis of rotation”
http://tiny.cc/5squjy 46, 109

6.8 “Free rotation of a rectangular parallelepiped”
http://tiny.cc/1squjy 110

6.9 “The three axes of a top”
http://tiny.cc/2tqujy 113

6.10 “Precession of a spinning top”
http://tiny.cc/wtqujy 119, 127

6.11 “Precession of a rotating wheel”
http://tiny.cc/otqujy 119, 122

6.12 “Physics of riding a bicycle with no hands”
http://tiny.cc/vsqujy 120

6.13 “Stabilization using a spinning top”
http://tiny.cc/5tqujy 125

7.1 “Freely falling frame of reference”
http://tiny.cc/1uqujy 133

7.2 “Simple pendulum in a rotating reference frame”
http://tiny.cc/euqujy 138, 139

7.3 “Gyrocompass”
http://tiny.cc/ququjy 140

7.4 “Torsional pendulum on a rotating table”
http://tiny.cc/cuqujy 142

7.5 “FOUCAULT’s pendulum”
http://tiny.cc/luqujy 54, 148, 149, 151

8.1 “Elastic deformation: HOOKE’s law”
http://tiny.cc/cvqujy 154, 173

8.2 "Bending a rod"
http://tiny.cc/4uqujy 160, 162, 173

8.3 "Plastic deformation, rupture strength"
http://tiny.cc/kvqujy 167, 168

8.4 "Reducing sliding friction"
http://tiny.cc/mvqujy 172

9.1 "BROWNian motion"
http://tiny.cc/4wqujy 175

9.2 "The Compressibility of water"
http://tiny.cc/6vqujy 182

9.3 "Buoyancy in a model fluid"
http://tiny.cc/fwqujy 184

9.4 "The tensile strength of water"
http://tiny.cc/3vqujy 169, 187, 426, 431

9.5 "Surface work"
http://tiny.cc/tvqujy 188

9.6 "Coalescence of Hg droplets"
http://tiny.cc/mwqujy 192, 209

9.7 "A Model Gas: Its barometric density distribution"
http://tiny.cc/rwqujy 196, 204

9.8 "Magdeburg hemispheres" (Otto von Guericke's experiment)
http://tiny.cc/vwqujy 199

9.9 "BEHN's tube"
http://tiny.cc/0vqujy 206

10.1 "Model experiments for streamlines"
http://tiny.cc/8wqujy 215, 217, 221, 225, 239

10.2 "Turbulence of flowing water"
http://tiny.cc/cxqujy 217

10.3 "Fluid Flow around obstacles"
http://tiny.cc/fcgvjy 220, 226, 235, 237–239

10.4 "Smoke rings"
http://tiny.cc/ocgvjy 233

10.5 "MAGNUS effect"
http://tiny.cc/dcgvjy 240

Part II Vibrations and Waves

11.1 "Vibrations of a tuning fork"
http://tiny.cc/idgvjy 249

11.2 "Transverse normal modes of a stretched rubber band"
http://tiny.cc/negvjy 261

11.3 "Transverse vibrations of a string"
http://tiny.cc/odgvjy 262, 263

11.4 "Longitudinal vibrations of a helical spring"
http://tiny.cc/ydgvjy 270

11.5 "HELMHOLTZ resonators"
http://tiny.cc/0cgvjy 273

11.6 "Free and forced oscillations of a torsional pendulum (POHL's pendulum)"
http://tiny.cc/uegvjy 274, 275, 290

11.7 "Forced oscillations with a pocket watch"
http://tiny.cc/zegvjy 279

11.8 "Coupled pendulums: Force coupling"
http://tiny.cc/pfgvjy 283, 285

11.9 "Acceleration coupling and chaotic oscillations"
http://tiny.cc/8cgvjy 284, 290

11.10 "Coupled oscillations with damping"
http://tiny.cc/5egvjy 285

11.11 "Antirolling tank"
http://tiny.cc/jfgvjy 286

11.12 "Wobble oscillations"
http://tiny.cc/qcgvjy 287

12.1 "Model of a travelling wave"
http://tiny.cc/2fgvjy 292

12.2 "Experiments with water waves"
http://tiny.cc/tfgvjy 293, 294, 295–299, 300–303, 305, 341, 347

12.3 "The acoustic radiometer"
http://tiny.cc/gggvjy 327

12.4 Fresnel's zones
http://tiny.cc/fggvjy 331

Part III Thermodynamics

13.1 "Model experiment on thermal expansion and evaporation"
http://tiny.cc/nggvjy 370

14.1 "Adiabatic changes of state"
http://tiny.cc/yggvjy 400

14.2 "Compressed-air motor"
http://tiny.cc/0ggvjy 406

15.1 "Liquefaction of oxygen"
http://tiny.cc/8ggvjy 419

15.2 "Liquid and solid nitrogen"
http://tiny.cc/ihgvjy 422

15.3 "Solid carbon dioxide (dry ice)"
http://tiny.cc/lhgvjy 423

16.1 "Model experiments on diffusion and osmosis"
http://tiny.cc/xhgvjy 175, 196, 432, 443

16.2 "Crookes' radiometer (light mill)"
http://tiny.cc/qhgvjy 433

19.1 "Operation of a STIRLING engine"
http://tiny.cc/gigvjy 498, 500, 502

19.2 "Technical Version of a STIRLING engine"
http://tiny.cc/kigvjy 500

19.3 "Heat pump/refrigerator"
http://tiny.cc/1hgvjy 502

Mechanics

I

1 Introduction; Distance and Time Measurements

Part I

1.1 Introduction

Physics is an empirical science. It is based on experimentally-determined facts. The facts remain, their interpretations may change in the course of the historical progress of the science. Facts are obtained from observations; sometimes chance observations, but usually carefully-planned observations. Observing must be learned; the inexperienced may be readily deceived. We offer two examples of this:

a) *The colored shadows*. In Fig. 1.1, we see a white wall W, an incandescent gas lamp[C1.1] and an electrical incandescent lamp. P is an arbitrary opaque object, e.g. a cardboard square. – Initially, only the electric lamp is turned on. It illuminates the white wall *except for* the shadow of the cardboard square, S_1. The shadow is marked, for example with a snippet of paper pinned to the wall. – Then only the gas lamp is ignited. Again, the wall appears white, this time *including* the marked area S_1. A black shadow from the cardboard square can now be seen at S_2. – Now we come to the actual experiment: While the gas lamp is burning, the electric lamp is switched on. This causes not the least physical or objective change to the area S_1. Nevertheless, to our eyes the image undergoes a fundamental change. At S_1,

C1.1. In the incandescent gas lamp, a net permeated with oxides of Th and Ce (the 'mantle') is heated by a gas flame (C. Auer, 1885, see Vol. 2, Sect. 28.6).

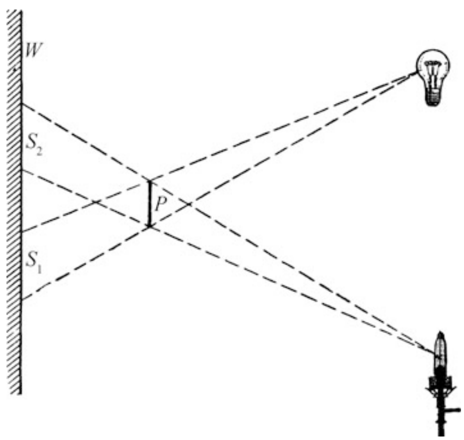
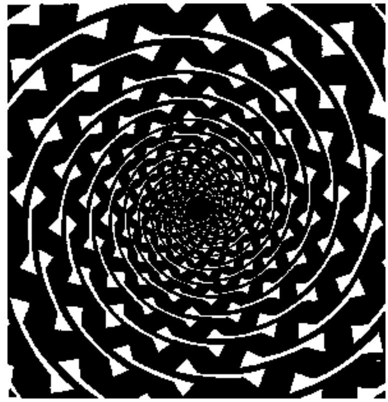


Figure 1.1 Colored shadows

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_1

Figure 1.2 Spiral illusion



we see a lively *olive-green* shadow. It looks quite different from the (now reddish-brown) shadow at S_2. But still, only the light from the gas lamp can reach our eyes from the area S_1. That area is however surrounded by a bright *halo* due to the light from the electric lamp. The presence of this halo by itself gives rise to the very noticeable change in the color of the area S_1.

This demonstration is instructive for every beginner: *Colors* are not physical properties, but rather results of psychology and physiology! Not taking this fact into account has given rise to a good deal of useless effort in the past.

b) *The spiral illusion*. In Fig. 1.2, everyone sees a system of spirals curving around a common midpoint. However, in fact it consists of concentric circles. One can verify this immediately by following one of the circles with the point of a pencil [C1.2].

C1.2. Figure 11.42c shows another, similar optical illusion.

These and many other phenomena which result from the way our sensory organs function seldom cause difficulties for practiced observers. But they still warn us to be careful. How many other as yet unrecognized subjective influences may be lurking in our physical observations of Nature!? In particular, the most general concepts which have been developed since earliest times in the course of human experience, such as space, time, forces etc., must be considered suspect. Physics may yet have to deal with many a prejudice and some misinterpretations.

1.2 Distance and Length Measurements. Direct Distance Measurements

Without doubt, experiments and observations have yielded new knowledge, often knowledge of great import, even when they were carried out only in a *qualitative* manner. Nevertheless, experiments and observations attain their full value only when all the quantities involved are determined in terms of precise numerical values with units. Measurements play an important role in physics. The art of physical measurement is highly developed, the number of applicable techniques is large, and they are the subject of an extensive specialized literature and their own technical field (metrology).

Among the manifold types of physical measurements, those involving *lengths* and *times* occur particularly often – sometimes alone, frequently paired with measurements of other quantities. It is therefore expedient for us to begin by discussing the measurement of lengths and of times, and to elucidate their fundamentals, while not concerning ourselves with the technical details of how they are carried out.

We learn the usage of the words 'length' and 'distance' as children. *Every direct measurement of a length is based on applying and comparing to a ruler or other length standard*. One counts how many lengths of the ruler are contained within the length to be measured. This may seem trivial, but it is often not adequately taken into account. The procedure of measurement itself, i.e. comparison with a length standard, is not sufficient; in addition, a *unit of length* must be defined.

Every definition of physical units is completely arbitrary. The most important requirement is always an international agreement, which must be as all-inclusive as possible. Furthermore, ready reproducibility is desirable, along with convenient numerical values for the most frequently applied measurements in everyday life.

Length or distance measurements are based on the unit of length, the *meter*. The meter was previously (before 1960) defined by the length of a metal bar (the "archival meter"), kept at the *Bureau des Poids et Mésures* in Sèvres; it is also called the "standard meter". The modern definition of the meter will be given below.

For calibration purposes, length standards are commercially available. They take the form of *gauge blocks*; these are rectangular steel blocks with planar, parallel, highly polished end surfaces. When struck together, they stick to each other (cf. Fig. 9.19). They can be used to reproduce lengths to a precision within $10^{-3}\,\mathrm{mm} = 1\,\mu\mathrm{m}$ (termed 1 micrometer, or 1 micron).

For practical length measurements, one uses *divided* length scales or rulers, and various forms of measuring instruments. Rulers should have division marks whose length is $2\,\frac{1}{2}$ times as great as their spacing; then fractional distances can be estimated most accurately.

Measurement instruments for lengths facilitate reading off the fractional distances by means of mechanical or optical arrangements. The mechanical instruments make use of length conversions using some sort of arrangement of levers, or screws ("screw micrometers"), or gears ("dial indicators"), or spirals. *Vernier calipers* are also frequently used.

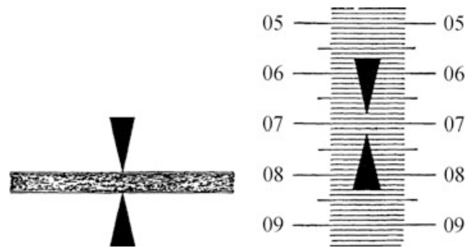


Figure 1.3 Length measurements using a microscope

C1.3. The optical microscope, referred to here, is discussed in detail in Vol. 2, Sects. 18.11 and 18.12.

Observations with a microscope[C1.3]: The microscope is among the most important optical instruments for length determinations. These are direct length measurements. As an example which can be demonstrated to a large audience, we could measure the thickness of a human hair.

With a simple microscope, an image of the hair is projected onto a screen. The thickness of the hair is marked on the screen with two arrows, cf. Fig. 1.3, left side. Then the hair is removed and replaced on the microscope's object stage by a small scale etched into a glass plate (object micrometer), showing e.g. one millimeter divided into 100 scale marks. The field of view now shows the image seen in Fig. 1.3, right side. We can read off four scale divisions between the arrows; the thickness of the hair was thus $4 \cdot 10^{-2}$ mm or 40 µm.

C1.4. Today, with optical methods, distances can be measured with uncertainties down to about 20 nm (1 nm $= 10^{-9}$ m, i.e. 1 nanometer). Because of the requirements of manufacturing technology, efforts are being made to reduce this limit even further (cf. https://www.ptb.de/cms/en/presseaktuelles/journals-magazines/ptb-news/ptb-news-ausgaben/ptb-news/news07-2/high-precision-length-measurements.html).

The error limits for length measurements with optical methods can be reduced to around ± 0.1 µm.[C1.4] Mechanical methods can be used down to ± 1 µm. The naked eye is limited to around ± 50 to ± 30 µm (i.e. the diameter of a hair!).

1.3 The Meter as a Unit of Length

For direct length measurements, one can employ a ruled scale with extremely fine divisions, which are no longer visible to the naked eye. This is demonstrated in Fig. 1.4. – One scale is attached to the fixed part and one to the movable part of a screw caliper. Both scales are glass platelets with fine division marks. As seen by the viewer, they are one behind the other, and overlap over a large region. The black division marks and the transparent gaps between them have the same width (more precisely, each has a width of e.g. $\frac{1}{20}$ mm).

In the zero position of the instrument, the marks on one scale just cover the gaps on the other, so that the whole overlap region is opaque and appears dark. If one now moves the caliper with its scale slowly to the right, the overlap region appears periodically bright and dark. Each new darkening means that the distance $a - b$ has been increased by the spacing of the division marks (in this example 1/10 mm). As

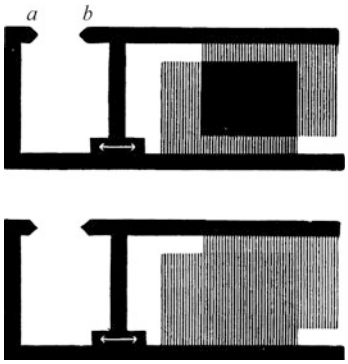


Figure 1.4 An interference micrometer, enlarged for clarity

a result, by counting the number of dark-light intervals of the invisibly fine scales, one can carry out a direct length measurement. This is, put succinctly, a length measurement using geometric *interference*.

There is an optical analog to this length measurement by interference: In optics, one can replace the man-made scales by naturally given scales. One uses the light waves of a particular spectral line emitted by excited atoms of the krypton isotope $^{86}_{36}$Kr. Their wavelength in vacuum (the "division mark spacing") has been compared to the standard meter in Sèvres, and by international agreement, the 1 650 765.73-fold multiple of this wavelength has been defined as one meter [C1.5].

C1.5. The unit of length 'meter' was redefined in 1983: "The meter is the length of the distance which light traverses during a time period of (1/299 792 458) seconds". This fixes the numerical value of the velocity of light in vacuum and relates the meter to the unit of time, the second (see e.g. http://en.wikipedia.org/wiki/History_of_the_metre, and http://physics.nist.gov/cgi-bin/cuu/Info/Units/meter.html).

In this way, the hope is that the precise sense of the word 'meter' will be passed on to later generations more securely than by using the "archival meter" as a prototype for defining the unit. An archival meter rod, in spite of all the care that may be taken, remains an impermanent object. All rulers change their lengths in the course of long time periods. This is the result of internal material changes which occur within all solid bodies.

1.4 Indirect Length Measurements of Very Large Distances

Baseline methods, stereogrammetry: Very long distances are often no longer accessible to *direct* measurement methods. Consider for example the distance between two mountain peaks, or the distance of a celestial object from the earth. One must then apply an indirect method of length measurement, e.g. the well-known baseline method as indicated in Fig. 1.5. The length BC of the baseline is found when possible by a direct length measurement. Then the angles β and γ are determined. From the length of the baseline and these angles, the required distance x is obtained graphically or by calculation.

Part I

This method, well known from school mathematics, is however not free of fundamental difficulties. It assumes without further consideration that the light rays used to measure the angles can be identified with the straight lines of Euclidian geometry. That is a presumption, and its admissibility can in the end be confirmed only empirically. – Fortunately, we need not concern ourselves with such fundamental uncertainties in the case of the usual physical measurements on the earth. They become important only in special cases, e.g. in determining the enormous distances which are relevant to astronomy. Nevertheless, the beginning student of physics should be aware of such potential difficulties; otherwise, he or she may see no problems at all with length measurements and hold them to be the most trivial sort of measurements in general. This opinion however is valid only for *direct* length measurements, i.e. when applying a ruler and comparing lengths.

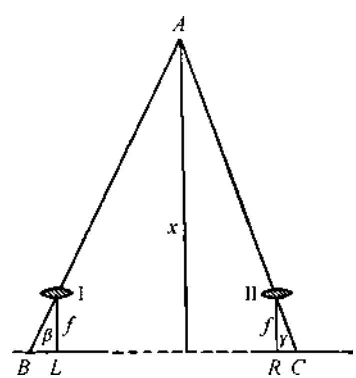
To end our brief discussion of length measurements, we mention an elegant technical variation of the baseline method of length determination, the so-called *stereogrammetry*[C1.6]. It is the preferred method for surveying terrain, especially in mountainous regions. In physics, it is used among other things for determining complex 3-dimensional orbits or paths, e.g. those of lightning discharges.

C1.6. The method of stereogrammetry, which is today employed using digital data processing technology, is applied in many fields; in addition to surveying terrain, also e.g. in architecture and in medicine (X-ray stereograms). Modern surveying methods use in addition the global positioning system (GPS), based on the arrival of radio signals from satellites.

In Fig. 1.5, the angles β and γ were determined using some sort of protractor (e.g. a telescope with divisions on a circular scale). Stereogrammetry replaces the two angular measurements at the ends of the baseline by two cameras. Their objectives are denoted as *I* and *II*. The images *B* and *C* of the same object *A* are shifted at their centers (different viewing angles) by the distances *BL* or *CR*. From *BL* or *CR* on the one hand, and the overall length *BC* on the other, the desired distance x of the object *A* can be computed. This can be understood simply by applying geometry. For a given baseline $I - II$ and a given focal length f of the lenses, a calibration table can be compiled.

Thus far, the method offers nothing notable. But now we encounter a serious difficulty: It would be time consuming and often impossible for example to determine the individual segment lengths in the complicated, zig-zag path of a lightning discharge from the corresponding images *B* and *C*. This difficulty can however be avoided. One combines the two photographic images in the well-known manner in a stereoscope to give *one* image field which appears 3-dimensional. We see in Fig. 1.6 how the two

Figure 1.5 Length measurements using a baseline, and stereogrammetric length measurements[C1.6]



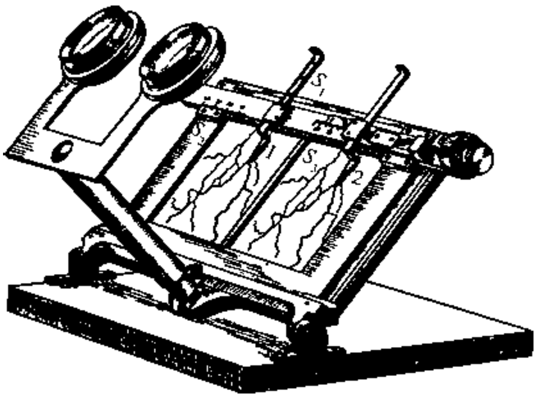


Figure 1.6 A stereoscope with a moveable marker. The images represent forked lightning discharges

individual photographs are combined in a stereoscope. And now comes the decisive trick: the use of a "movable marker".

This movable marker or cursor is obtained by employing two identical pointers *1* and *2*. They can be moved over the image surface horizontally and vertically. The distances through which they have been moved can be read off the scales S_1 and S_2. In addition, the distance between the two pointers can be varied systematically in a measurable way (S_3, with a scaled screw drive).

Looking into the stereoscope, we see these two pointers superimposed as *one*, apparently floating freely in space. If we change their spacing (by turning the screw S_3), then the cursor appears to move towards us or away from us in the 'image space'. By using all three possible motions (S_1, S_2, S_3), the cursor can be adjusted to indicate any given point in the image, thus to point to a mountain peak, to an arbitrary position along the crooked path of a lightning stroke, etc. This demonstration is very impressive. A calibration table then allows us to read off the distances (height, depth, width) which determine the location of the given point (i.e. its three spatial coordinates) from the values of the scale readings S_1, S_2, and S_3.

1.5 Angle Measurements

Proceeding from length measurements, we can determine surface areas, volumes and angles. We note only a few aspects of angle measurements:

Planar angles (Fig. 1.7) are defined by the ratio $\dfrac{\text{Arc length } b}{\text{Radius } r}$; *solid* angles (Fig. 1.8) by the ratio $\dfrac{\text{Spherical surface sector } A}{(\text{Radius } r)^2}$. Then all angles are determined as pure (dimensionless) *numbers*.

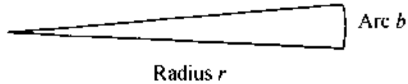


Figure 1.7 The definition of the planar angle Ω

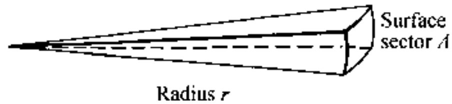


Figure 1.8 The definition of the solid angle φ

The symbol °, representing the word *degree*, is simply a numerical unit, defined by the equation

$$1^\circ = \frac{1/360 \text{ circumference}}{\text{radius}} = \frac{2\pi r/360}{r} = \frac{\pi}{180}$$
$$= 0.01745\ldots \tag{1.1}$$

π is the abbreviation for the number 3.1415... Correspondingly, ° is an abbreviation for the number 0.01745... Thus e.g. $\alpha = 100^\circ$ is identical to $\alpha = 100 \cdot 0.0175 = 1.75$.

The unit of all angles is the number 1. It is often expedient to denote this number 1 in referring to a planar angle as the unit *radian* (abbreviated *rad*); or in referring to a solid angle as the unit *steradian* (abbreviated *sr*). If these names for the number 1 occur in some combination of units, one sees immediately that the determination of an angle is included in the implied measurement procedure.

The equation: 1 radian = 57.3° simply expresses the identity

$$1 \text{ radian} = 57.3 \cdot 0.0175 = 1.$$

A cone with the opening angle α intersects a surface sector A on a sphere centered at its apex, with $A = 2\pi r^2(1 - \cos\alpha)$.
For $\alpha = 32.8^\circ$, the corresponding solid angle is $\Omega = 1 =$1 steradian. It intersects a surface sector of area $A = r^2$ on the sphere, i.e. the fraction $r^2/4\pi r^2 = 1/4\pi = 7.96\,\%$ of the surface area of a sphere of radius r.
Example: Radiation intensity can be referred to the solid angle subtended by the beam of radiation. See Vol. 2, Chap. 19.

1.6 Time Determinations. True Time Measurements

The word *time* has two meanings: either an *interval* of time, or a *moment* in time. Just as a length is limited by its two end points, a time interval is limited by the two moments in time at its beginning and its end. Just as every direct length measurement is associated with a comparison to a *length standard*, so is every direct time measurement associated with a comparison to a *clock*. The most important clocks are based on counting uniformly repeated processes (usually rotations or oscillations). Here, we cannot define "uniform" as a concept, but rather only experimentally: One compares many clocks of all possible different types with each other and with periodically recurring astronomical phenomena. This comparison leads to a "struggle for survival": Clocks whose behavior deviates from that of the majority are eliminated, and the regularity of operation of the surviving clocks is declared to be "uniform".

Just like the definition of the unit of length, the definition of the unit of time is a matter for international agreement. The unit called the *second* was originally defined in terms of astronomical phenomena (initially by the *rotation* of the earth around its axis, and later by its annual orbit around the sun). These definitions, which in the final analysis are *mechanical*, have proved to be inadequate[1]. Thus, they have been replaced by an *electrodynamic* definition. It is based on an *electromagnetic wave* emitted by $^{133}_{55}$Cs atoms under certain excitation conditions, with a wavelength λ in the range of 3 cm. Since 1967, the second has been defined as the time in which 9 192 631 770 oscillations (wave crest + wave trough) of these waves occur (and can be counted) ("atomic clock").

1.7 Clocks and Graphical Registration

Clocks which can be used for practical time measurements are well known. They make use of mechanical oscillation phenomena. Either a hanging pendulum oscillates in the gravitational field of the earth (e.g. as in wall clocks or upright clocks), or a balance wheel oscillates rotationally around a spiral spring (e.g. in a wristwatch or pocket watch)[C1.7]. We still have to show that the oscillations of such a pendulum or balance wheel can be referred to a uniform rotation.

C1.7. Today, most small clocks and watches utilize the mechanical oscillations of quartz crystals (Sect. 11.8).

The swinging of a pendulum, put concisely, is similar to *a circular motion as seen from the side.* Looking in the plane of the circular motion, we see an object on a circular path as though it were moving back and forth. Its motion over time is precisely the same as that of a swinging pendulum. An optical recording can show this in a particularly graphical way; it converts the temporal sequence into a spatial diagram and represents the motion in terms of a curve.

In order to register this motion, we use the arrangement which is illustrated in Fig. 1.9: A slit S is imaged onto a screen P by the lens L. The light source which illuminates the slit (an arc lamp) is not shown.

[1] This is due to the non-constancy of the earth's rate of rotation. The frictional forces associated with the ocean (and surface) tides increase the rotational period of the earth (with a power dissipation of ca. 10^9 kilowatts!), adding about $1.5 \cdot 10^{-3}$ s per century. As a result, each century lasts about 30 s longer than the preceding one. Furthermore, the rotational period varies within the course of a year; for not clearly understood reasons, it is ca. $2 \cdot 10^{-3}$ s longer in May than in July. Finally, additional random fluctuations in the period of the earth's rotation have been observed. See **Exercise 1.1**.

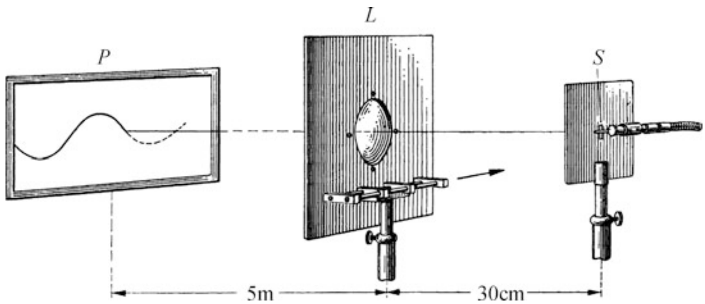


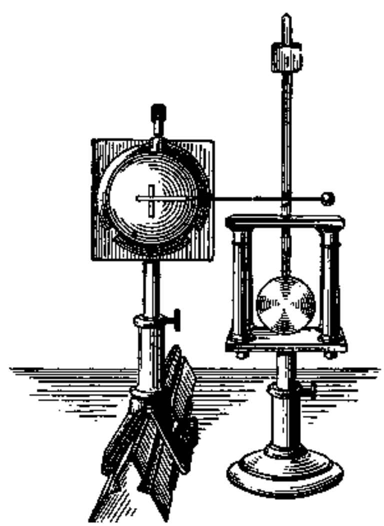
Figure 1.9 The relation between circular motion and a sine curve. In front of the vertical slit *S*, there is a horizontal pin attached to the edge of a cylinder which can rotate around a horizontal axis. The cylinder can be made to rotate around this axis using a flexible shaft; the axis is parallel to the plane of the slit.

The lens *L* is moved by a slider in the direction of the arrow; this causes the image of the slit to move uniformly across the screen *P*. This screen is coated with a phosphorescent powder, which glows for a long time after being illuminated briefly. In front of the vertical slit *S*, we arrange one behind the other:

1. a metal pin which moves in a circle around the surface of a cylinder with a horizontal axis which is parallel to the plane of the slit (Fig. 1.9), and

2. a wire pointer which is fastened to the side of a pendulum (cf. Fig. 1.10, metronome pendulum). Its maximum deflection is adjusted to be the same as the radius of the cylinder which carries the metal pin for the first experiment.

Figure 1.10 A metal pin connected to a metronome pendulum in front of a slit. This arrangement replaces *S* in Fig. 1.9 for the second experiment.



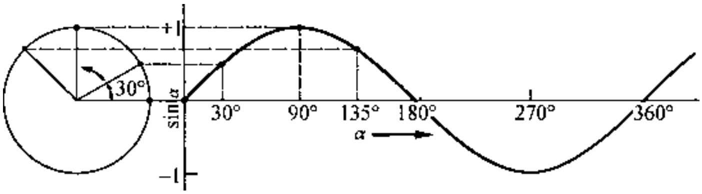


Figure 1.11 A sine curve gives the dependence of the angular function $\sin\alpha$ on the angle α

In both cases, we obtain the *same* curve, deep black on a bright green glowing background: a sine curve, see Fig. 1.11. This close connection between circular motion, the oscillation of a pendulum and a sine curve plays an important role in many different areas of physics. We will come back to this topic in Sect. 4.3.

Graphical registration is useful for many rapidly-occurring processes, and in some cases it is indispensable[C1.8]. For this purpose, e.g. oscilloscopes are useful; they can also be employed for measuring short times. They can be used to register periods of time down to 10^{-9} s.

C1.8. Images of this type are often obtained using CCD cameras (charge-coupled devices), which also permit the images to be recorded in color. Today, of course, there are also convenient computer programs which generate graphical representations from measured data.

Today, in order to synchronize and calibrate clocks for science and technology, time signals are broadcast; they are controlled by precision standard clocks (Cs atomic clocks).

1.8 Measurement of Periodic Sequences of Equal Times and Lengths

Let us assume that within a given time t, N identical processes occur in sequence, each one lasting a time T; e.g. oscillations or rotations. One then defines in general

$$t/N = T \quad \text{as the } \textit{period} \tag{1.2}$$

and (from the Latin *frequentia*)

$$N/t = 1/T = \nu \quad \text{as the } \textit{frequency} \tag{1.3}$$

(the unit of frequency is 1/s = 1 Hertz (Hz)). Within a given length l, N identical forms occur in sequence, each one of length D. Then we define[2]

$$l/N = D \qquad \text{as the period of length,} \tag{1.4}$$

$$N/l = 1/D = \nu^* \qquad \text{as the frequency of length.} \tag{1.5}$$

[2] Generally accepted terminology is lacking. For D, terms like 'wavelength' and 'lattice constant' are in use. In optics, $1/D$ is called the *'wavenumber'*. This term is however a poor choice, as is the term *'number of revolutions'* used in technology. An electric motor, for example, has a rotational frequency of ν = 3000/min = 50/s = 50 Hz. Reciprocal lengths and reciprocal times are not numbers, but rather dimensioned quantities with units.

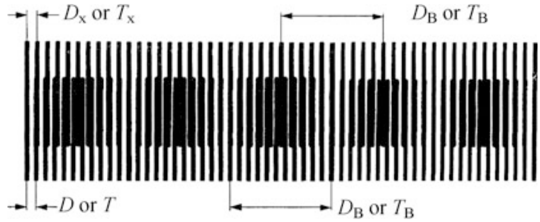


Figure 1.12 Measurement of a periodic sequence of identical lengths D_x or identical times T_x. One can imagine two parallel combs whose teeth partially overlap (an arrangement like that shown in Fig. 1.4). The upper comb has the period D_x (T_x), and the lower comb has the period D (T). The beats produced by interference have the period D_B (T_B). In the figure, the length periods (or frequencies) are $D_x = 0.12\,\text{cm}$ ($\nu_x^* = 8.3/\text{cm}$), $D = 0.11\,\text{cm}$ ($\nu^* = 9.1/\text{cm}$) and $D_B = 1.2\,\text{cm}$ ($\nu_B^* = 0.83/\text{cm}$). We find (for $D < D_x$) that: $N \cdot D_x = (N+1)\,D = D_B$ (where N is equal to 10 in the example shown in the figure). Equations (1.6) and (1.7) follow from this (for the general case of $D \gtrless D_x$ and $T \gtrless T_x$).

A periodic sequence of identical lengths D_x or of identical times T_x can be measured using the same scheme, as illustrated in Fig. 1.12. In the upper part of the figure, we see the periodic sequence of D_x or T_x. Superposed on it is a second periodic sequence of known lengths D or known times T, which differ only slightly from D_x or from T_x. The superposition produces a third periodic sequence (by "interference") of "enlarged" lengths D_B or "stretched" times T_B, which can readily be counted or measured[3]. Quantitatively, for the measurement of lengths, we find

$$1/D_x = 1/D \pm 1/D_B \quad \text{or} \quad \nu_x^* = \nu^* \pm \nu_B^* \tag{1.6}$$

and for the measurement of times

$$1/T_x = 1/T \pm 1/T_B \quad \text{or} \quad \nu_x = \nu \pm \nu_B \tag{1.7}$$

(the minus sign applies when $D < D_x$ or $T < T_x$.)

Out of the many practical examples we mention only one, the *stroboscopic measurement* of a frequency ν_x or a period T_x[C1.9].

C1.9. Modern high-speed stroboscopes attain time resolutions of 10^{-15} s.

Figure 1.13 shows a leaf spring F; we cause it to oscillate at a very high, unknown frequency ν_x, as shown in Fig. 11.45. This oscillation is projected onto the wall by intermittent light pulses, a uniform sequence of individual flashes. This kind of illumination can be produced simply by using a rotating disk with, for example, 20 slits. It is placed in the light beam from a suitable source of light.

The frequency of illumination ν is obtained from the rotational frequency of the disk, ν_D. One can use a stopwatch to count the number of revolutions of the disk N within the time t. Then $N/t = \nu_D$ is the frequency of the disk and $\nu = 20\,\nu_D$ is the frequency of the light pulses.

[3] The index B refers to the word *'beats'*, defined later in the text, e.g. in Fig. 11.10 and in Sect. 12.4.

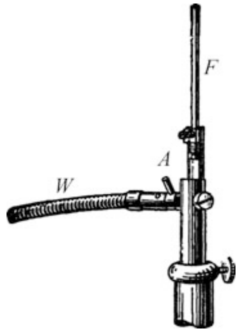


Figure 1.13 A leaf spring F for demonstrating stroboscopic time measurement. The oscillations of this spring are shown in Fig. 11.45. The driving force is provided by a flexible shaft W and a holder which is tapped on one end by the pin A. More details are given in Sect. 11.10 under "forced oscillations".

We start with a *high* light pulse frequency ν and reduce it gradually. The *image* of the oscillating leaf spring seems to move, and the frequency of its apparent oscillations becomes lower and lower as the light pulse frequency decreases (stroboscopic time dilation). When the image appears to move very slowly, one can finally determine ν_{B} conveniently, for example $\nu_{\mathrm{B}} = 1.5$ Hz. We insert ν_{B} and ν into Eq. (1.7) and for the example, we find $\nu_{\mathrm{x}} = 50$ Hz. – In the limiting case $\nu_{\mathrm{B}} = 0$, the image of the leaf spring is stationary and Eq. (1.7) gives $\nu_{\mathrm{x}} = \nu$.

1.9 Indirect Time Measurements

Instead of today's usual "direct" or true time measurements, i.e. measurements based on counting periodic motions, in earlier times non-periodic phenomena were often used, for example in hour glasses and water clocks. These types of clocks played an important role in the early history of mechanics (e.g. in the work of GALILEI, Sect. 2.4). Today, they are found only in the puny form of "egg timers". But a modern variant which makes use of radioactive decays for determining the age of historical objects is of considerable importance[C1.10].

One concluding remark: We have described only some measurement procedures for lengths and times, but have not attempted to define these two concepts in words or sentences. Both of them have developed over long times from extremely diverse experiences and observations. Physicists base their use on only a narrow selection of these. For 'time', for example, they might say the following:

Every physical measurement requires at least two "readings"; for length measurements, the beginning and the end of the length must be "read out"; electrical measurement instruments show the difference between the zero point (ground potential or zero current) and the actual reading, etc. Between the first and the second readout, our heart beats or the clock ticks. All observations can be ordered into one of two groups: In the first group, the results of a measurement

C1.10. This refers for example to the so-called C-14 method. Here, the radioactive carbon isotope ${}^{14}_{6}\mathrm{C}$ (half-life ca. 5700 years) is employed; its concentration in living organisms has a roughly constant equilibrium value, while in no-longer living objects, it decreases due to the radioactive decay, so that the age of such objects (since their death) can be estimated (see e.g. http://en.wikipedia.org/wiki/Radiocarbon_dating.)

depend upon how often our heart beats or the clock ticks between the first and the second readout. In the second group, by contrast, that is of no significance for the result of the measurement. Then we can say that the processes in the first group depend on a quantity that we call time, and which we measure by counting heartbeats or the ticks of the clock. This by no means exhaustively encompasses the concept of *time*, but it is at least not just a meaningless phrase[C1.11].

"This by no means exhaustively encompasses the concept of time, but it is at least not just a meaningless phrase".

C1.11. There is a voluminous literature on the question raised here, which is only too difficult to answer; namely, "What is time?". It has been written not just by physicists, but also by scientists and scholars from many other disciplines. To name just one (arbitrarily chosen) example, we mention the paperback book by GENE YERGER, "The Meaning of Time", whose subtitle is "A Theory of Nothing" (Perfect Paperback Press, 2008).

Exercises

1.1 In the footnote in Sect. 1.7, it is pointed out that the coming century will be about 30 s longer than the one just past. Find the precise number using the information given in the text. (Sect. 1.7)

1.2 A swinging pendulum is observed through a rotating disk with four slits, which is rotating at five revolutions per second. At which oscillation period T does the pendulum seem to be at rest to the observer? (Sect. 1.8)

The Description of Motion: Kinematics

2

Part I

2.1 Definition of Motion. Frames of Reference

Motion refers to a change in the location of an object with time, as seen from a fixed, rigid frame, the "frame of reference". This supplementary specification is quite essential. We can see this from a randomly-chosen example: A bicyclist looks down at her feet and sees them moving in circular paths with the pedals. An observer standing on the sidewalk sees a very different picture of the motion of the bicyclist's feet; for her or him, the feet follow a wavelike path, namely the cycloids which are sketched in Fig. 2.1.

The rigid solid body which is our frame of reference for the description of motion in the rest of this chapter is the earth or the floor of the room where we are located. We leave the daily rotation of the earth out of consideration. (In reality, we are practicing physics on a large carousel. The earth is also not really rigid, but instead is deformable.)

Later, we will occasionally change the standpoint of our observations, i.e. our frame of reference. We will take the earth's rotation into account in some discussions, and sometimes also the deformation of the earth. This will always be mentioned explicitly. Otherwise we would have an endless confusion, especially when we treat rotational motions.

"Otherwise we would have an endless confusion, especially when we treat rotational motions".

For the description of all motions, also called *kinematics*, we require the concepts of *velocity* and *acceleration*. We will begin with them.

Figure 2.1 The path followed by the pedals of a bicycle as seen by a stationary observer

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_2

2.2 Definition of Velocity. Example of a Velocity Measurement

Suppose that an object moves through a distance Δl within the time interval Δt. Then we define

$$u_{\mathrm{m}} = \frac{\text{Distance moved } \Delta l}{\text{Time interval } \Delta t} \tag{2.1}$$

as the *mean* velocity along the direction of the distance Δl. This quotient *changes* in general if one successively decreases the distance Δl. However, the changes gradually decrease to below the precision of the measurements. The value of u_{m} which is then measured, which depends only on the starting point, is denoted as the velocity u at the starting point. Mathematically, one thus finds the velocity u as the limiting value of u_{m} by taking the limit $\Delta t \to 0$. The symbol Δ is conventionally replaced by d, giving for the definition of the velocity

$$u = \frac{\mathrm{d}l}{\mathrm{d}t} \tag{2.2}$$

i.e. the differential quotient of the distance travelled divided by the time interval.

This definition in many cases requires the measurement of rather short times. As an example, we consider the measurement of the *muzzle velocity* of a bullet from a pistol.

Figure 2.2 shows a suitable setup for this measurement. The distance interval Δl is fixed by two thin cardboard disks; its length could be for example 22.5 cm. The time measurement is performed in a straightforward manner by referring to the basis of all time measurements,

C2.1. For this velocity measurement, the defining equation (2.2) or (2.1) is used directly. This principle is also applied e.g. to the measurement of the velocities of molecules in a molecular beam. For macroscopic objects, however, a much more elegant method is the one described in Sect. 5.9 (Fig. 5.19), which makes use of momentum conservation.

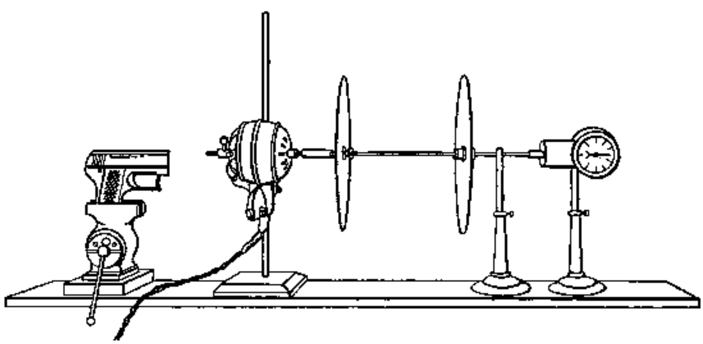


Figure 2.2 Measurement of the velocity of a bullet from a pistol with a simple "time recorder". On the right is the tachometer which registers the rotational frequency[C2.1].

a uniform rotation. The time markers are registered automatically. For this purpose, an electric motor causes the two disks on a common shaft to rotate at a uniform, rapid rate. Their rotational frequency ν, i.e. the quotient of the number N of rotations/time t, is determined by a tachometer which measures rotational frequencies[1], for example $\nu = 50\,\mathrm{Hz}$.

The bullet first passes through the left disk, and the bullet hole that it leaves is our first time marker. While it travels over the distance of 22.5 cm to the second cardboard disk, a certain time elapses, and the bullet hole or time marker in the second disk is shifted relative to the first marker by a certain angle, corresponding to the rotation of the shaft during this elapsed time. After stopping the rotation, we measure an angle of ca. 18° or 1/20 of the circumference of the disks.

By pushing a pin through the two bullet holes, we make the angular shift visible for distant observers in the silhouette of the apparatus.

The time delay Δt was thus $\frac{1}{20} \cdot \frac{1}{50}\,\mathrm{s} = 10^{-3}\,\mathrm{s}$. From this, we find the velocity

$$u = \frac{22.5\,\mathrm{cm}}{10^{-3}\,\mathrm{s}} = \frac{0.225\,\mathrm{m}}{10^{-3}\,\mathrm{s}} = 225\frac{\mathrm{m}}{\mathrm{s}}.$$

We then repeat the experiment with a smaller distance Δl between the disks of only 15 cm. The end result is the same. Thus, the distance chosen in the first experiment was already short enough. It permitted us to measure the true muzzle velocity and not the smaller mean value over a longer trajectory[C2.2].

C2.2. The word 'trajectory' originally meant 'flight path'. It is now used in a more general sense to mean the path (location *vs*. time or geometrical form) followed by *any* moving object. When this path is circular (like the path of a satellite around a planet), it is called an 'orbit'. We will use the simpler term 'path' here in most cases.

Only in cases with a constant or uniform velocity can we choose the quantities Δl (measured distance) and Δt (elapsed time) freely to allow the most convenient measurement. In such cases, one can write the velocity in abbreviated form as $u = l/t$.

"One should early on adopt the habit of always writing the units after every numerical value of a physical quantity. It belongs to good physics practice!"

One should early on adopt the habit of always writing the units after every numerical value of a physical quantity. It belongs to good physics practice! This spares the reader from having to deduce the units meant from the physical context, and spares oneself from committing frequent computational errors. When different units are used, the *numerical values* of the measured quantities also change. Their recalculation is automatically reliable if the quantitative results are always given as numerical values *and* units[C2.3].

C2.3. POHL indicates here the advantage of using *physical-quantity equations*, which he consistently employs. In a preliminary remark on writing physical equations (included in the volume on 'Mechanics' since the 12th edition), he writes among other things, "For each symbol, we write both the numerical value *and* the unit. The choice of units is free. Those mentioned under some equations are simply examples" (see also Sect. 2.6).

Example

The velocity $u = 225\,\mathrm{m/s}$ is to be recalculated in terms of kilometers per hour. We have $1\,\mathrm{m} = 10^{-3}\,\mathrm{km}$ and $1\,\mathrm{s} = (1/3600)$ hour; then

$$u = 225\frac{10^{-3}\,\mathrm{km}}{(1/3600)\,\mathrm{hour}} = 810\,\mathrm{km/h}.$$

[1] If no tachometer is available, one can make use of a simple reducing gear: a pulley of circumference d is mounted on the motor shaft. An endless belt made of string with length $L \gg d$ drives another pulley a few meters away; its knot serves as a marker. The number N' of revolutions of the string in a time t can be counted using this marker. Then $\nu = \frac{N'}{t} \cdot \frac{L}{d}$ is the rotational frequency.

Part I

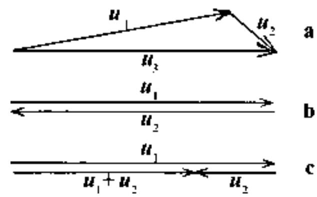


Figure 2.3 The geometrical addition of vectors, e.g. of two velocities

Well-formulated units can often be seen as a compact form of measurement instructions. – We will encounter this in many places in this book.

In everyday life, we often make do with the *magnitude* of a velocity, e.g. 10 m/s. In physics, however, the magnitude is only *one* of the two determining quantities for a velocity. The second is its *direction*. In physics, velocities are always directed quantities, i.e. mathematically, they are *vectors* (represented graphically as an arrow). This can be most clearly seen in the process, well known also to non-physicists, of *addition of two velocities*.

In Fig. 2.3a, the high velocity $\boldsymbol{u}_1$ (e.g. the velocity of an aircraft relative to the surrounding air) and the much lower velocity $\boldsymbol{u}_2$ (e.g. the local wind velocity), which generally points in a different direction, are vectorially combined to yield the "resultant" velocity $\boldsymbol{u}_3$ (the groundspeed of the aircraft).

Vectors which point in opposite directions differ in their signs; for example, Fig. 2.3b is described by the equation $\boldsymbol{u}_1 = -\boldsymbol{u}_2$ or $\boldsymbol{u}_1 + \boldsymbol{u}_2 = 0$. – Therefore, $\boldsymbol{u}_1 + \boldsymbol{u}_2$ in Fig. 2.3c refers to the geometrical addition or combination of the two oppositely-directed vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. The resultant vector has a magnitude (length of the arrow) of $|\boldsymbol{u}_1 + \boldsymbol{u}_2| = |\boldsymbol{u}_1| - |\boldsymbol{u}_2|$. Magnitudes are denoted by vertical bars on both sides of the symbol[C2.4].

C2.4. The rules for vector addition are explained here in an intuitive manner using the example of combining velocities. Vector quantities are denoted by printing their symbols in boldface, while their magnitudes are represented by normal type, dispensing with the vertical bars for simplicity.

2.3 Definition of Acceleration: The Two Limiting Cases

Motions with a constant velocity are rare. In general, the magnitude and the direction of the velocity change along the path of the motion.

In Fig. 2.4, the vector $\boldsymbol{u}_1$ indicates the velocity of a body *at the beginning* of a time interval Δt. *During* the time interval, the body is supposed to gain an additional velocity $\Delta\boldsymbol{u}$ in an arbitrary direction, represented by the short second arrow. *At the end* of the time interval Δt, the body has the velocity $\boldsymbol{u}_2$. It is determined graphically in Fig. 2.4 as the arrow $\boldsymbol{u}_2$.

Then we define

$$\boldsymbol{a}_{\mathrm{m}} = \frac{\text{Velocity increase } \Delta\boldsymbol{u}}{\text{Time interval } \Delta t}$$

Figure 2.4 The general definition of the acceleration

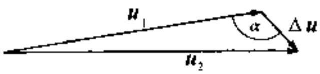
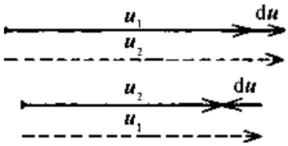


Figure 2.5 The definition of the path acceleration



as the *mean* acceleration. The time interval Δt is chosen so that the quotient no longer changes measurably when Δt is further decreased. Mathematically, one carries out the limit $\Delta t \to 0$, replaces the symbol Δ by d, and thus obtains for the *acceleration*:

$$\boldsymbol{a} = \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t}\,. \tag{2.3}$$

Just like the velocity, the acceleration is also a vector. The *direction* of this vector is the same as that of the increase in the velocity $\Delta\boldsymbol{u}$ (Fig. 2.4).

In Fig. 2.4, the angle α between the increase of the velocity $\Delta\boldsymbol{u}$ and the initial velocity $\boldsymbol{u}_1$ was arbitrary. We now consider *two limiting cases:*

1. $\alpha = 0$ or $\alpha = 180°$, Fig. 2.5. The velocity increase lies along the same direction as the original velocity. Then only the *magnitude*, not the direction of the velocity changes. In this case the acceleration is referred to as a *path acceleration* with the magnitude

$$a = \frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\mathrm{d}^2 l}{\mathrm{d}t^2}. \tag{2.4}$$

2. $\alpha = 90°$, Fig. 2.6. The velocity increase points in a direction perpendicular to the original velocity $\boldsymbol{u}$. Now, only the *direction* and not the magnitude of the velocity changes; within the time interval $\mathrm{d}t$ through the angle $\mathrm{d}\beta$. In this case, one refers to $\mathrm{d}\boldsymbol{u}/\mathrm{d}t$ as the *transverse acceleration* or *radial acceleration* $\boldsymbol{a}_\mathrm{r}$. From Fig. 2.6, we see immediately that the following relation[2] holds:

$$\mathrm{d}\beta = \frac{\mathrm{d}u}{u} \quad \text{or} \quad \mathrm{d}u = u \cdot \mathrm{d}\beta\,.$$

[2] *Example:* $\mathrm{d}\beta = 4.5°$, $° = 0.0175$, $\mathrm{d}t = 0.1\,\mathrm{s}$, $\omega = \dfrac{\mathrm{d}\beta}{\mathrm{d}t} = \dfrac{4.5 \cdot 0.0175}{0.1\,\mathrm{s}} = 0.79/\mathrm{s}$.

Part I

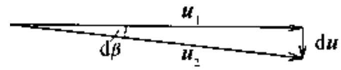


Figure 2.6 The definition of the radial acceleration

The quotient

$$\frac{d\beta}{dt} = \omega \tag{2.5}$$

is called the angular velocity[C2.5], and the radial acceleration becomes

$$a_r = \omega \cdot u \,. \tag{2.6}$$

C2.5. The angular velocity is also a vector. It points in the direction of the axis of rotation: if we look along this direction, we see the rotation as clockwise. Then Eq. (2.6) has the general form: $\boldsymbol{a}_r = \boldsymbol{\omega} \times \boldsymbol{u}$ (for the definition of the vector product, see Chap. 6). In all cases considered here, with $\omega \perp u$, the magnitude equation is sufficient.

The word *acceleration*, according to the above definitions, is used in physics in a quite different sense than in everyday language. First of all, in everyday life an accelerated motion usually means a motion at a higher speed, e.g. the accelerated circulation of a document or file. – Secondly, the word acceleration in everyday language leaves the direction completely out of consideration.

For the majority of motions, path accelerations $\boldsymbol{a}$ and transverse accelerations $\boldsymbol{a}_r$ are both present at the same time; along the path of the motion, both the magnitude and the direction of the velocity are changing during the motion. Nevertheless, we will limit ourselves for the moment to the limiting cases of pure path acceleration (straight-line motion) or pure radial acceleration (circular orbit).

2.4 Path Acceleration and Linear Motion

(G. GALILEI, 1564–1642.) The path acceleration changes only the magnitude, not the direction of the velocity. As a result, the motion follows a straight-line path; its trajectory is linear.

A path acceleration is in principle easy to measure: We determine the velocity at two times with the time interval Δt; these velocities are u_1 and u_2. Then we compute $\Delta u = (u_2 - u_1)$ (positive or negative) and form the quotient $\Delta u/\Delta t = a$[C2.6].

C2.6. Here, we see among other things that POHL is concerned not only with the definition of a quantity, but rather he always gives a compact instruction for its measurement; an aspect which is not emphasized in many textbooks. It is in this sense that the detailed descriptions of measurements of velocities and accelerations given here are to be understood.

Δt, as we have already pointed out, must be chosen to be sufficiently small. The result of the measurement should not change on a further decrease of Δt. Practically, this requirement usually means that rather short time intervals Δt must be used. The latter are measured with some sort of *registration procedure*. That is, the course of the motion is first recorded automatically and the record is then evaluated after the motion is complete. But there is also a much simpler procedure. For example, time marks can be imprinted on the moving

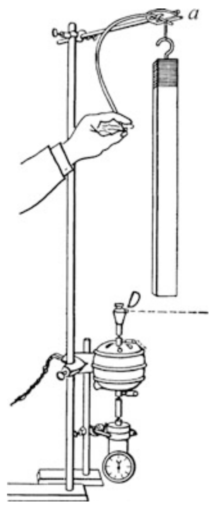


Figure 2.7 Measurement of the acceleration of a freely-falling body **(Video 2.1)**

Video 2.1:
"Free fall"
http://tiny.cc/jpqujy

object by a clock. Of course, the imprinting process must not disturb the motion itself. We give a practical example: The acceleration of a freely-falling wooden bar is to be determined. Fig. 2.7 shows a suitable arrangement. It can also be used for many other types of acceleration measurements[C2.7].

The essential element is a fine ink jet which is rotating in the horizontal plane. The jet is sprayed out of a nozzle D on the side of a rotating ink container (Fig. 2.8) (electric motor with its shaft vertical). The frequency, e.g. $\nu = 50\,\mathrm{Hz}$, is determined by a frequency meter. Here again, the time measurement is referred to a uniform rotation.

The wooden bar is wrapped in a jacket of white paper and hung at the point a. A cable trigger lets it fall at the desired time. The bar then falls through the rotating ink jet and on to the floor. – Figure 2.9 shows the result: a clean sequence of time marks at time intervals of 1/50 of a second.

C2.7. For example, with modern apparatus, the accelerated motion of a freely-falling body can be conveniently recorded electronically using photoelectric triggers. Beams of light which are directed at photocells are interrupted briefly by the passing object, and the resulting signals control an electronic stopwatch.

The object continues falling while the ink jet swishes past; this causes the curvature of the time marks.

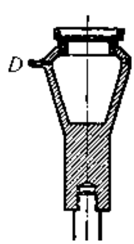
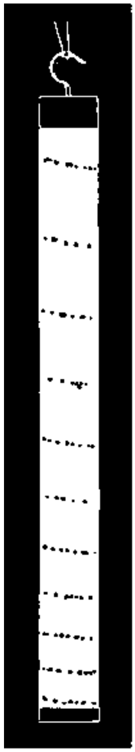


Figure 2.8 The ink jet used in Fig. 2.7, at half its actual size



Velocity $u = \dfrac{\Delta l}{\Delta t}$ (cm/s)	Velocity change Δu in $\Delta t = \dfrac{1}{50}$ s (cm/s)	Acceleration $a = \dfrac{\Delta u}{\Delta t}$ (m/s^2)
285.50		
	22.50	11.25
263.00		
	17.50	8.75
245.50		
	18.00	9.00
227.50		
	21.25	10.63
206.25		
	21.25	10.63
185.00		
	18.50	9.25
166.50		
	19.00	9.50
147.50		
	18.00	9.00
129.50		
	19.50	9.75
110.00		
Mean values:	19.50 cm/s	9.8 m/s^2

Figure 2.9 Falling body with time marks and their evaluation, with the usual experimental and readout errors. This experiment also shows that the measurement of a second-order differential quotient is in general an awkward matter. (**Video 2.1**)

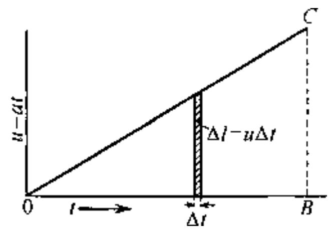
Video 2.1:
"Free fall"
http://tiny.cc/jpqujy

We can see with the unaided eye that the motion is accelerated: The spacing of the time marks, i.e. the distance Δl through which the object falls in the time $\Delta t = (1/50)$ s, increases continuously. The computed values of the velocity $u = \Delta l/\Delta t$ are written beside the marks; the velocity increases on the average within each (1/50) s by the amount 19.5 cm/s. Here, we ignore the inevitable errors in the individual values (scatter). This case of *free fall* is one of the rare examples of a *constant or uniform path acceleration*. For the magnitude of this constant acceleration, we find

$$a = 9.8\,\mathrm{m/s^2}.$$

Repetition of the experiment with an object made of a different material, for example a brass tube instead of the wooden bar, yields the same value for the acceleration. *The constant acceleration a for free fall is the same for all falling bodies. It is denoted by g.* A more precise value is $g = 9.81\,\mathrm{m/s^2}$. It is called the *acceleration of gravity*

Figure 2.10 The velocity u as a function of time for constant path acceleration



or *apparent gravity*[3]. These are incidental experimental facts at this point. Their great significance will become apparent later.

Our practical example of a measurement led us to the special case of a *constant* path acceleration. This is an important case.

Constant acceleration means that the increase in the velocity Δu is the same in equal time intervals Δt. The velocity u increases as shown in Fig. 2.10 linearly with time t. In each time interval Δt, the object travels through the distance Δl. Therefore, we have $\Delta l = u\Delta t$. Here, u is the average value of the velocity within each time interval Δt. Such an interval is shown in Fig. 2.10 as a shaded area. The entire triangular area *0BC* is the sum of all of the distance segments Δl traversed in the time t. Thus for the case of constant path acceleration, the overall distance l traversed in the total time t is given by the equation[C2.8]

$$l = \tfrac{1}{2}at^2\,, \tag{2.7}$$

i.e. the distance increases as the square of the time during which the acceleration acts.

C2.8. The formula explained graphically here (2.7) follows mathematically from the defining equations $u = \mathrm{d}l/\mathrm{d}t$ (2.2) and $a = \mathrm{d}u/\mathrm{d}t$ (2.3). For $a = $ const and from the rules of integral calculus, we obtain:
$l = \int u\,\mathrm{d}t = \int at\,\mathrm{d}t$
$= a\int t\,\mathrm{d}t = \tfrac{1}{2}at^2$.

If the body already had an initial velocity u_0 before the beginning of the acceleration, then instead of Eq. (2.7), the equation

$$l = u_0 t + \tfrac{1}{2}at^2 \tag{2.8}$$

would apply.

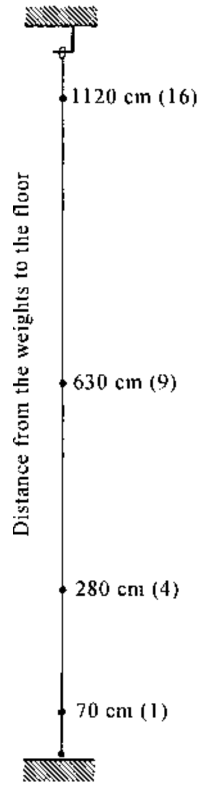
The origin of the constant path acceleration is completely irrelevant. It could be for example an electrical force instead of a mechanical force.

Usually, for verifying Eq. (2.7), one makes use of the constant acceleration $a = g$ which acts on a freely falling body. As an example, we mention the well-known falling rope experiment.

This experiment consists of a thin rope, hung perpendicularly, with a series of lead weights attached to it; cf. Fig. 2.11. The lowest weight nearly

[3] This numerical value holds near the surface of the earth; for most purposes, g can be considered to be constant. A more precise consideration shows that g depends weakly on the geographical latitude of the location (Sect. 7.6). Furthermore, it also depends on local variations in the surface features of the earth (e.g. deposits of heavy ores under the ground), and, although only weakly, on the altitude of the location where the observations are made.

Figure 2.11 Falling rope



touches the floor. The spacing of the other weights is arranged to correspond to the squares of whole numbers. When the upper end of the rope is released, the weights hit the floor one after another. The experimenter hears their impacts at equal time intervals.

Strictly speaking, observations of free fall should be carried out in a vacuum. Only then can perturbations due to air resistance be eliminated. In a highly evacuated glass tube, all bodies indeed fall at exactly the same speed[C2.9]. A lead ball and a fluffy feather fall at the same speed in a vacuum, and strike the ground at the same time. In room air, in contrast, the feather falls much more slowly, as we well know. But heavy bodies with a relatively small surface area are only weakly slowed by air resistance (cf. Fig. 5.20).

C2.9. Free fall can be used for carrying out experiments under conditions of weightlessness, e.g. in the evacuated drop tower at University of Bremen, which is over 120 m high (cf. Comment C7.2).

2.5 Constant Radial Acceleration and Circular Orbits

(C. HUYGHENS, 1629–1695.) The radial acceleration $\boldsymbol{a}_\mathrm{r}$ does not change the magnitude of a velocity $\boldsymbol{u}$, but only its direction. Let the radial acceleration $\boldsymbol{a}_\mathrm{r}$ be constant and let us assume that no other accelerations are present. Then the direction of $\boldsymbol{u}$ changes by the same angular increment $\mathrm{d}\beta$ in equal time intervals $\mathrm{d}t$. The resulting orbit is a circle. It is traversed with the constant angular velocity $\omega = \mathrm{d}\beta/\mathrm{d}t$.

The time required for a complete revolution was termed the (rotational) *period* T in Sect. 1.8, and its inverse was called the rotational (or mechanical) *frequency* ν; thus $\nu = 1/T$. Then for the orbital velocity u, we have

$$u = \frac{\text{Circumference}}{\text{Period}} = \frac{2\pi r}{T} = 2\pi r\nu \tag{2.9}$$

and the angular velocity is

$$\omega = \frac{2\pi \cdot \text{ Angle}}{\text{Period}} = 2\pi\nu \tag{2.10}$$

so that[C2.10]

$$u = \omega r\,. \tag{2.11}$$

C2.10. In general, Eq. (2.11) is written as a vector equation: $\boldsymbol{u} = \boldsymbol{\omega} \times \boldsymbol{r}$.

When a circular orbit is traversed with a constant velocity, the angular frequency ω is thus a factor of 2π greater than the mechanical frequency ν, that is $\omega = 2\pi\nu$; therefore, ω is often called the *circular frequency* (unit usually 1/s). This definition and these relations hold quite generally for periodic processes (e.g. the rotation of an electric motor shaft).

We combine Eqns. (2.6) and (2.11) and obtain the result[C2.11]

$$a_\mathrm{r} = \omega^2 r = \frac{u^2}{r}\,. \tag{2.12}$$

C2.11. Using the vector formulations of Eqns. (2.6) and (2.11), we obtain for the first part of Eq. (2.12) likewise a vectorial form: $\boldsymbol{a}_r = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})$ or, for $\boldsymbol{\omega} \perp \boldsymbol{r}$: $\boldsymbol{a}_r = \boldsymbol{\omega}^2 \boldsymbol{r}$. The radius vector $\boldsymbol{r}$ points from the center of the circular orbit outwards, while the acceleration vector points in the opposite direction towards the center.

This radial acceleration a_r must be present so that a body can traverse a circular orbit of radius r with the constant angular velocity (circular frequency) ω or the constant orbital velocity u.

Intuitively, the constant radial acceleration required for a circular orbit has the following interpretation (Fig. 2.12):

Suppose that a body traverses a circular segment *ac* within the time interval Δt. Imagine this orbit to be composed of two steps that occur one after the other, namely:

1. A motion along a (tangential) path *perpendicular* to the radius with the constant velocity u, $ad = u\Delta t$;

2. an accelerated motion along a (radial) path (anti-parallel to the radius), $l = \frac{1}{2}a_\mathrm{r}(\Delta t)^2$. The thin horizontal lines (time markers) in the figure permit us to see that the motion along l is accelerated, so that we can apply Eq. (2.7).

A numerical example can be useful. The earth's *moon* moves during the time $\Delta t = 1\,\mathrm{s}$ along the direction *ad*, that is *perpendicular* to

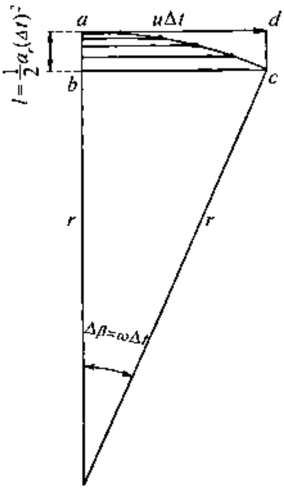


Figure 2.12 The explanation of radial acceleration

its orbital radius, by 1 km, thereby "increasing" its distance from the earth slightly. At the same time, it "approaches" the earth along the orbital radius in an accelerated motion, traversing the distance $l = \frac{1}{2}a_{\mathrm{r}}(\Delta t)^2 = 1.35\,\mathrm{mm}$. Thus, the net effect is that the radius remains unchanged, and the orbit is circular. The radial acceleration of the moon is found to be $a_{\mathrm{r}} = 2.70\,\mathrm{mm/s^2}$.

2.6 Distinguishing Physical Quantities and Their Numerical Values

In commerce[C2.12], the price of every item is a "quantity", i.e. the product of a *numerical value* and a *unit*. For example, a hat might cost 10 $ and a pencil 10 cents. No one would consider these two prices to be the same. The ratio of the two prices is rather

$$\frac{10\,\$}{10\,\mathrm{cent}} = \frac{10 \cdot 100\,\mathrm{cent}}{10\,\mathrm{cent}} = 100.$$

The same principle holds in physics: Distance l, time t, velocity u, acceleration a, frequency ν etc. are measured as *physical quantities*, i.e. as products of a *numerical value* with a *unit*. A velocity of $u = 7$ is meaningless. It becomes meaningful only when expressed as for example $u = 7\,\mathrm{m/s}$. The confusion of physical *quantities* (e.g. distance $l = 5\,\mathrm{km}$ and velocity $u = 5\,\mathrm{km/h}$) with their *numerical values* (in the example, the value of the distance is 5 and the value of the velocity is 5) gives rise to widespread but *incorrect* definitions, such as for example "the velocity is the distance traversed in a unit time"[C2.13]. The velocity is not a *distance*, but rather the quotient of distance/ time. – Or still worse: "The frequency is the number of oscillations in one second". First of all, a frequency is not a number, but

C2.12. The following two sections do not belong directly to the content of the rest of this chapter. POHL however added them here, after introducing the first important concepts, from the 10th edition on; possibly because of his experience with examinations. They deal with simple and indeed self-evident topics, which are however to some extent not consistently treated in most textbooks even today.

C2.13. Such "false definitions" can still be found in many textbooks!

rather the quotient number/time; e.g. the pulse frequency of a person is ca. 70/minute. Secondly, a physical quantity which is supposed to be *generally applicable* cannot be defined in terms of a particular *unit* such as the second.

2.7 Base Quantities and Derived Quantities

Some few physical quantities are termed *base quantities* and are measured in units defined especially for those quantities, the *base units*; for example, *time* with the unit *second*, or *temperature* with the unit *kelvin*. If one wishes to introduce a base quantity, it can be defined only in terms of axioms which are founded on extensive experiments or observations, and not on equations.

Most physical quantities are defined as *derived quantities*. This means that they and their units can be defined not in terms of axioms, but instead by means of equations which contain other quantities and their units. We recall the example of the velocity, $u = \mathrm{d}l/\mathrm{d}t$ and its units, chosen for example to be meters/second or kilometers/hour, etc.

The possibility of defining quantities and their units in terms of equations is the only point in which derived *quantities differ from the* base *quantities employed.*

No physical quantity is *in its essence* a base quantity; one could introduce many different quantities as base quantities. The number and type of the base quantities should be chosen insofar as possible so that no two derived quantities have the same defining equation. – In distinguishing between base quantities and derived quantities, one should in no case envision a hierarchy; base quantities are not imbued with a special aura, nor should limiting them to a certain number (e.g. three) be raised to the status of a dogma.

The currently agreed-upon base quantities in the international unit system (SI) and their base units are[C2.14]:

C2.14. In these books (Vol. 1 and Vol. 2), we use mainly (but not exclusively) the units defined within the SI. See for example http://physics.nist.gov/cuu/index.html or http://www.bipm.org/en/measurement-units/.

- Length (or distance), unit *meter* (m),
- Time, unit *second* (s),
- Mass, unit *kilogram* (kg),
- (Thermodynamic) temperature, unit *kelvin* (K),
- Electric current, unit *ampere* (A),
- Luminous intensity, unit *candela* (cd); and
- Amount of substance, unit *mole* (mol).

Note that the names of *units* and their abbreviations are printed in Roman type (m, s, kg, . . .), while *physical quantities* are printed in *italics* (*distance* = l or d, *mass* = M or m, etc.). Avoid confusing them; there are not enough letters in the Roman and Greek alphabets to allow us to use different symbols for all the quantities and units that we require[C2.15].

C2.15. Recommendations for writing physical-quantity expressions and equations and for naming or abbreviating units are given at http://physics.nist.gov/cuu/pdf/checklist.pdf.

Exercises

2.1 A person wants to row her boat along a straight-line path from A to B. A and B are points opposite each other on either side of the mouth of a river, which is 600 m wide; B is also 300 m upriver from A. The person rows at the same velocity as the upriver flow rate of the tide which is coming in. In which direction should she row? (Sect. 2.2)

2.2 Calculate the distance l which is traversed during the fourth second by an object in free fall which began falling from a resting position at the time $t = 0$. The acceleration of gravity is $g = 9.81\,\mathrm{m/s^2}$. (Sect. 2.4)

2.3 A parachute jumper initially falls freely over a distance of 50 m with the acceleration of gravity $g = 9.81\,\mathrm{m/s^2}$ after jumping (we may neglect air friction). Then his parachute opens, delaying his fall (i.e. it produces a negative acceleration, upwards) by $2\,\mathrm{m/s^2}$, so that he finally lands on the ground with a velocity of 3 m/s. How long (time t) was he in the air, and from what height h did he jump? (Sect. 2.4)

2.4 An object of mass $m = 20\,\mathrm{kg}$ increases its velocity from 15 m/s to 18 m/s, covering a distance of 20 m in the process. What is the magnitude of its constant acceleration and the corresponding force? (Sect. 2.4 and 3.2)

2.5 A particle begins at rest at the point P and moves along a straight-line path to the point O, a distance of 3 m away. There, its velocity is 6 m/s. Plot its velocity *vs.* position for (a) constant acceleration, and (b) a sinusoidal motion back and forth through the point O. How long does the motion from P to O take in each case? (Sect. 2.4 and 4.3)

2.6 How large is the radial acceleration a_r of a person who is at 51.5° north latitude (e.g. in London)? The radius of the earth is 6378 km. (Sect. 2.5)

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_2) contains supplementary material, which is available to authorized users.

3 Fundamentals of Dynamics

3.1 Force and Mass

In kinematics, the relevant quantities are the “velocity” and the “acceleration”; in *dynamics*, we require in addition the concepts of “force” and “mass”. These two terms, which have a variety of meanings in everyday language, first need to be defined as *technical terms* of physics.

The concept of *force* is based on the sense of our own muscles. A force is determined qualitatively by two characteristics: It can *deform* solid bodies which are held fixed, and it can *accelerate* movable bodies.

For the *deformation* of a solid body we give a clear example: Figure 3.1 shows an oak table with a thick frame Z. Two mirrors are set on this table, and a beam of light rays passes between them as shown in the sketch. It casts an image of the light source, an illuminated slit S, onto the wall. Any bending of the table top will tilt the mirrors in the direction shown by the small arrows. This “light-beam pointer” provides great sensitivity to small motions owing to its length (ca. 24 m); it acts as an ‘optical lever’. – We set down a metal weight at the point A, e.g. a 1 kg block. The table top will be deformed. In physics and technology, one says that a *force* is acting on the block, called its *weight*; *the deformed table top keeps the block from being accelerated downwards*. Then we press down on the block with our little finger, causing the table top to bend more. We could say: Now a second force is acting on the block, namely our muscular force. Finally, we replace the block with a long rod and slide our hand down the rod (Fig. 3.2). Again, the table top is

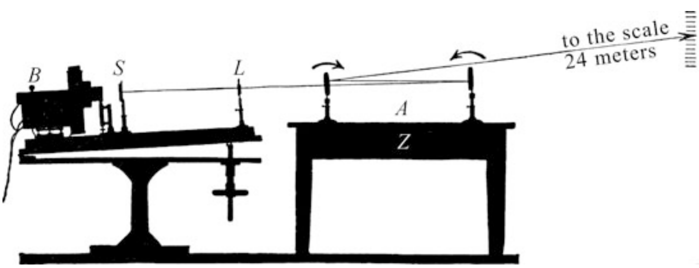
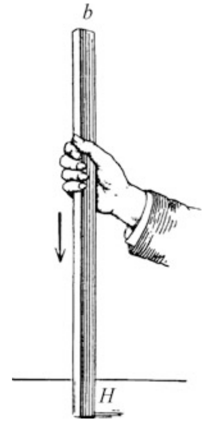


Figure 3.1 Optical detection of the deformation of a table top by small forces, e.g. a finger pressing down at the point A[C3.1]

C3.1. This striking experiment can be readily carried out today using a beam of laser light. (Compare also a torsional deformation as shown in Fig. 6.7.)

K. Lüders, R.O. Pohl (Eds.), *Pohl’s Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_3

Figure 3.2 The force due to external friction acts *downwards* on the rod, and its counter force acts in an upwards direction on the hand. The arrow indicates the direction in which the hand slides along the rod, and of its frictional force.



deformed and we can say that a force is acting on the rod, in addition to its own weight; it is called *external friction*[1], and is produced here by the slipping motion of our hand sliding down the rod.

Forces are *vectors*. They can be decomposed into their components in different directions. Figure 3.3 shows an example of this decomposition.

Forces always act in pairs: The two forces may act at two different places, but they are oppositely directed and have the same strength. In NEWTON's terms, this is called *actio* = *reactio*, or force = counter force. We offer three examples:

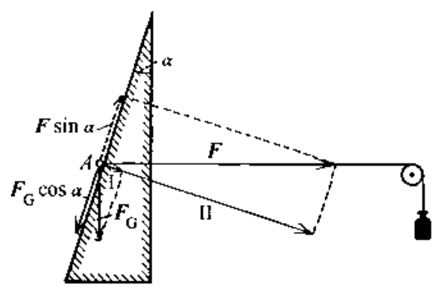


Figure 3.3 The decomposition of vectors into their components. A roller *A* is being held fixed on a steep ramp by a horizontal force $\boldsymbol{F}$. The vector $\boldsymbol{F}_{\mathrm{G}}$ denotes the weight of the roller. We decompose both $\boldsymbol{F}$ and $\boldsymbol{F}_{\mathrm{G}}$ into components parallel to the ramp and components perpendicular to it. The two components which are perpendicular to the surface of the ramp, represented by the arrows *I* and *II*, are held in equilibrium by the elastic force of the slightly deformed surface of the ramp. The components parallel to the ramp, $\boldsymbol{F}_{\mathrm{G}} \cos\alpha$ and $\boldsymbol{F} \sin\alpha$, pull the roller downwards and upwards, respectively. In equilibrium, we have $\boldsymbol{F} = -\boldsymbol{F}_{\mathrm{G}}/\tan\alpha$. If the ramp is very steep, α and $\tan\alpha$ are quite small, and thus the force $\boldsymbol{F}$ must be large.

[1] For internal friction (viscosity), cf. Sect. 10.2.

Figure 3.4 Deformation of a circular spring. In the center of the spring is a guide rod. This simple apparatus will be used later for demonstration experiments as an uncalibrated force meter.

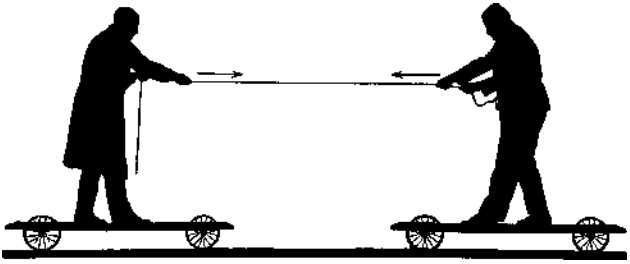


Figure 3.5 Force = counter force, *actio* = *reactio*[C3.2] **(Video 3.1)**

C3.2. At the left in the picture is master mechanic W. Sperber, and the author is on the right.

Video 3.1:
"Action = Reaction"
http://tiny.cc/ppqujy

1. In Fig. 3.4, at the left a stretched circular spring is shown being pulled by two hands. A force acts on each hand. If the spring is held by only one hand, no force and no deformation occurs; cf. Fig. 3.4 on the right.

2. In Fig. 3.5, we see two flat, nearly frictionless carts on a smooth, flat floor which supports their weight without noticeable deformation. The arrangement is completely symmetric, the carts and the men on each side have the same size, shape and mass. – Both of the men could pull at the same time, i.e. they could work together as a "motor"; or only the man on the left or the man on the right could pull. In each case, the two carts meet in the middle. Thus, always *two* forces occur simultaneously. They are oppositely directed and of equal strength. This is indicated by the arrows.

3. In the case of the force that we call "weight", the counter force would seem to be lacking. But that is simply due to our choice of the frame of reference. Figure 3.6 shows on the left the earth and a large stone. As seen from the sun or the moon, this picture would have to be drawn with *two* arrows. The earth pulls on the stone, and the stone pulls on the earth. Both bodies are accelerated towards each other. In Fig. 3.6 on the right, their approach is prevented by inserting a spring between them. In this case, two *new* forces occur, denoted by $\boldsymbol{F}_\mathrm{D}$ and $-\boldsymbol{F}_\mathrm{D}$. Now, equal but opposite forces act on each of the two bodies. Their sum is zero, and therefore the two bodies remain motionless relative to one another.

The concept of *mass* is even more ambiguous in everyday language than the word 'weight'. For example: dough is a mass that can be kneaded; the news media appeal to the broad mass of people, they consume a mass of paper, etc.

Part I

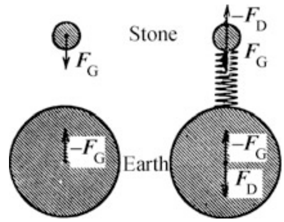


Figure 3.6 On the pairwise occurrence of forces, force = counter force

In physics, however, the concept of mass refers to two characteristics of every object, namely "gravitational" and "inertial" mass. "Gravitational" means that every massive object is attracted to the earth with a force which we call its "weight". – "Inertial" means that no object can change its velocity (magnitude and/or direction!) by itself; every change in velocity requires the action of a force.

3.2 Measurements of Force and Mass. NEWTON's Fundamental Equation of Motion

(ISAAC NEWTON, 1643–1727). For the measurement of masses and of forces, one can use the same basic apparatus, namely a *set of weights* (Fig. 3.7, upper part) and some sort of *scales* (e.g. a beam balance or a spring balance).

The measurement of mass is explained in Fig. 3.7: One defines the masses of two bodies as equal when they can be used interchangeably on a scale or balance[C3.3]. The *unit of mass* is defined by a primary standard made of noble metals[2]; it is internationally denoted as the *kilogram* (kg)[C3.4](1 kg = 10^3 grams (g), 10^3 kg = 1 metric ton (t)).

Like all massive bodies, the weights are attracted to the earth with a certain (gravitational) force. The *forces* which act on the weights

C3.3. In this case, mass is measured by making use of its "gravitational" property. Two masses are equal when they have the same weight. In principle, the "inertial" property can also be used, e.g. in collision experiments (Sect. 5.8) or with an oscillating spring-and-mass system (Sect. 4.3).

C3.4. This is the only unit which is still defined by a "prototype", and it seems to be showing signs of aging. Efforts are therefore being made to redefine the unit of mass (see e.g. http://en.wikipedia.org/wiki/Proposed_redefinition_of_SI_base_units and http://www.nist.gov/pml/newsletter/siredef.cfm).

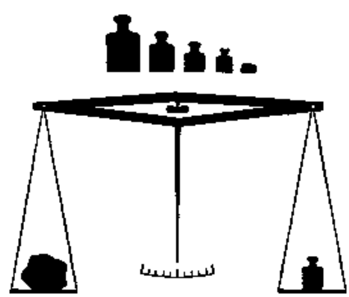


Figure 3.7 Measurement of a mass using a set of weights (*above*) and a beam balance. The masses of two objects are the same when both have the same weight at the same place, i.e. they are attracted to the earth with the same *forces*.

[2] It consists of 90 % Pt and 10 % Ir (wt.-%) and is kept at Sèvres, near Paris. See also http://www.bipm.org/en/publications/mises-en-pratique/kilogram.html.

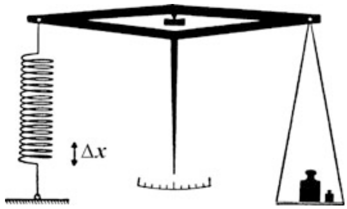


Figure 3.8 Measurement of a force using a set of weights and a beam balance. The force exerted by a spring which has been stretched by an amount Δx is compared to the force acting on a metal block from the earth's gravity, for short the *weight* of the block. (The weight of the relaxed spring is cancelled out by the weight of the small metal block.) For practical applications, a spring balance is more convenient to use than a beam balance; compare Fig. 3.9.

are called simply their *weights*[3]. These weights, i.e. *forces*, can also be used to measure forces in general. This is made clear by Fig. 3.8. There, the force of a spring is compared with another force, namely the weight of a kilogram. The unit of force is the *newton* (N) (cf. Eq. (3.4)). The weight of a kilogram block at a place where the acceleration of gravity is 9.81 m/s^2 (rounded off) is equal to 9.81 newton. Compare Fig. 3.9.

Why can we use the same apparatus, namely a set of weights and some sort of balance or scales, to measure both masses and forces? The answer is indicated in Fig. 3.7: The masses of two objects are defined to be equal when they have the same *weights* at the same place, i.e. when they are attracted with the same *forces* by the earth. The weight of an object depends first of all on a property of that object, called its *mass*, and secondly on the earth: the forces that we call weights are always directed vertically towards the center of the earth. The influence of the earth is the same on both sides of the balance and therefore cancels (more about the force that acts between masses, called the *gravitational force*, in Sect. 4.7).

The relation between force and mass is derived experimentally by varying the two quantities independently of one another and measuring them[C3.5]. The simplest experimental arrangement is shown in Fig. 3.10. As the known force, we make use of the force $F = F_{\mathrm{G}}$ which acts on the small block A and which we call the weight of the block. By means of a cord, this force is used to accelerate a loaded cart. The overall mass of the objects that are accelerated together, that of the cart and its load as well as that of the block A, is taken to be m. The acceleration is constant and therefore can be easily computed from the distance travelled, s, and the corresponding time t; it is $a = 2s/t^2$ from Eq. (2.7). *In this way, one finds that the acceleration a is proportional to the force F and inversely proportional to the*

C3.5. Now, having introduced forces in terms of the "gravitational" mass, we will derive the relation between force and "inertial" mass.

[3] The word 'weight' inevitably reminds those who are not steeped in physics of a *property* of the object instead of a force acting on it. This makes it more difficult to realize that these forces come about through the effect of the earth (or e.g. of the moon if the object were located on the lunar surface).

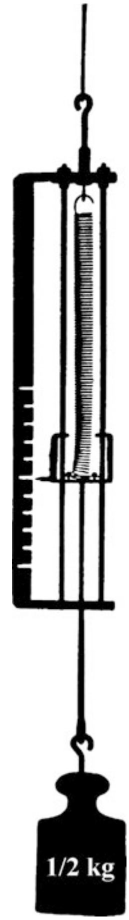


Figure 3.9 Calibration of a spring balance as a force meter. The known force is in this example the gravitational force that acts on a metal block of mass 1/2 kg. It is the same within the error limits of $\pm 0.3\,\%$ (cf. Sect. 7.6) at all points on the surface of the earth and in this case is equal to $\frac{1}{2} \cdot 9.81\,\mathrm{kg\,m/s^2} = 4.90\,\mathrm{newton}$.

mass m, that is

$$a = \frac{F}{m} \cdot \text{const.} \tag{3.1}$$

This finding forms the basis of mechanics. If one sets the proportionality constant = 1, then Eq. (3.1) takes on the simple and convenient form

$$a = \frac{F}{m}\,, \tag{3.2}$$

or, more generally using vector notation,

$$\boldsymbol{a} = \frac{\boldsymbol{F}}{m}\,. \tag{3.3}$$

One thus can dispense with measuring the mass and force *independently* of each other; or expressed positively, one makes use of the one quantity for the measurement of the other.

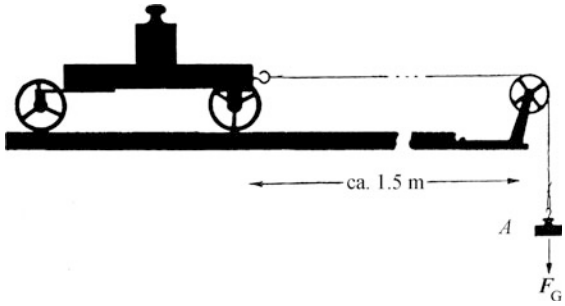


Figure 3.10 On the experimental derivation of the fundamental equation of motion. The force which produces the acceleration is the weight F_G of the small block A. The weight of the cart and its load is compensated by an unmeasurably small deformation of the horizontal surface (plate glass floor). – Friction and rotational inertia of the wheels reduce the acceleration by 5 to 10 %. To eliminate these sources of error, it is sufficient to incline the surface on which the cart moves by a small amount (a few tenths of a degree). This inclination is adjusted so that the cart, given a small push without the cord, moves along its path with practically constant velocity[C3.6].

C3.6. This basic experiment can be demonstrated today with a cart that moves almost without friction on an airtrack. POHL mentioned such airtracks in the later editions of his book, but did not describe any actual experiments with them.

Physics makes use of the mass for the measurement of forces. The resulting unit for force, kg m/s^2, is the *derived unit of force* which is called the newton, that is

$$1 \text{ newton (N)} = 1 \text{ kg m/s}^2. \tag{3.4}$$

Eq. (3.3) contains the great discovery of ISAAC NEWTON, the relation between force and acceleration. It will be found in later sections to withstand the most rigorous experimental tests. This is rather surprising: Eq. (3.3) contains the common property of all massive bodies, their *inertia* (Sect. 3.1) and the forces which are therefore necessary to *accelerate* them. However, the masses m of the objects were measured using the *other* property which all massive bodies have in common, namely their "gravitational mass"; and this in a state with the bodies *at rest*! – This fundamental fact is often described by the brief but frequently misunderstood statement: "gravitational mass = inertial mass" (equivalence principle).

Important applications of the equation $\boldsymbol{a} = \boldsymbol{F}/m$ will be treated in the next chapters. At the end of the present chapter, we first clarify some of the other concepts that will often be needed for what follows.

3.3 The Units of Force and Mass. Expressions Containing Physical Quantities

We begin by examining two results of Eq. (3.2) which will help to clarify the units of force and mass. To this end, we make use of *expressions with physical quantities*[C3.7], i.e. we replace each symbol with a numerical value *and* its unit.

C3.7. The consistent use of expressions and equations with physical quantities is a particular advantage of POHL's books!

1. In Eq. (3.2), we insert the force $F = 1\,\mathrm{N} = 1\,\mathrm{kg\,m/s^2}$ and the mass $m = 1\,\mathrm{kg}$. The result is:

$$a = \frac{1\,\mathrm{kg\,m/s^2}}{1\,\mathrm{kg}} = 1\,\frac{\mathrm{m}}{\mathrm{s}^2}\,.$$

Or in words: 1 newton is the force which can give to an object with a mass of 1 kilogram an acceleration of $1\,\mathrm{m/s^2}$, when *only* this force is acting on the object (the counter force is the inertial reaction force of the object).

2. In Eq. (3.2), we insert the force $F = 9.81\,\mathrm{N}$ (previously called 1 kilopond[C3.8]) and the mass $m = 1\,\mathrm{kg}$. The result is:

$$a = \frac{9.81\,\mathrm{kg\,m/s^2}}{1\,\mathrm{kg}} = 9.81\,\frac{\mathrm{m}}{\mathrm{s}^2}.$$

In words: If a body of mass[4] 1 kg is acted upon only by its weight, then it experiences an acceleration $a = 9.81\,\mathrm{m/s^2}$ (free fall!).

C3.8. The unit of force "kilopond", which is no longer used, was initially included in POHL's book to aid in making the distinction between force and mass units, following a suggestion by the German standards laboratories, the Physikalisch-Technische Reichsanstalt (today the PTB). It has now been suppressed in favor of the unit *newton*. POHL himself wrote (from the 17th edition on): "In physics, the kilopond as a unit of force alongside the newton is just as dispensable or superfluous as is the energy unit 'calorie' alongside the watt second."

3.4 Density and Specific Volume

In the cases of very large and clumsy or of very small bodies, the mass can often not be measured directly with a scale or balance, but the volume V is known from the dimensions of the body. In such cases, one can compute its mass by making use of the helpful concept of *mass density* or more briefly, simply *density*. Assume a body of mass m to have the volume V. Then we define

$$\text{Mass density } \varrho = \frac{\text{Mass } m}{\text{Volume } V}\,. \tag{3.5}$$

This quantity, for fixed external conditions (pressure, temperature), is a constant that characterizes the material of which the body is composed.

[4] The use of the word mass in the place of 'body' or 'object' is apparently ineradicable. Over and over, one finds for example a 'mass' hanging on a string, rather than a body, i.e. instead of the object, one of its properties! In English, the word 'weight' is used in a similar way: a 'weight' is hanging on a string, rather than a body with a certain weight. This is, however, officially tolerated by the SI.

Sometimes, instead of the mass density ϱ, one uses its reciprocal, the

$$\text{Specific volume } V_s = \frac{\text{Volume } V}{\text{Mass } m}\,. \tag{3.6}$$

Specific in general denotes a physical quantity when the quantity itself is not meant, but rather a quotient of the 'specific' quantity and another quantity. This makes it unnecessary to introduce a new name for the quotient. The 'specific volume' is thus the quotient of volume and mass, or the volume referred to the mass.

Exercises

3.1 An object sitting on a frictionless, horizontal plane on the earth's surface is subject to the following forces: $F_1 = 50\,\text{N}$ in the direction of the azimuthal angle $\alpha = 155°$ (geoscientific notation, i.e. north: $\alpha = 0$; east: $\alpha = 90°$ etc.); $F_2 = 30\,\text{N}$ in the direction $\alpha = 230°$; and $F_3 = 20\,\text{N}$ to the north. Which additional force to the northeast ($\alpha = 45°$) would be required to make the resultant total force F_{tot} lie along the east-west axis? What is the magnitude of this total force and in what direction does it act? (Sect. 3.1)

3.2 An object with the weight F_G is on a rough inclined plane with an inclination angle of $\alpha = 40°$. The coefficient of sticking friction is $\mu_h = 0.3$. The object is to be just kept from sliding down the incline by a force F which makes an angle Θ to the vertical. a) Find the expression relating the force F to the angle Θ. b) By variation of Θ, find the direction of the smallest force that will still hold the object from sliding. (Sects. 3.1 and 8.9)

3.3 A body of mass $m_1 = 0.5\,\text{kg}$ is sitting on a planar table at a distance of 1 m from its edge. A cord attached to the body is hanging over the edge of the table and carries a weight of mass $m_2 = 20\,\text{g}$ (see Fig. 3.10). The system begins to move starting from this configuration. How large are its velocity u and its kinetic energy E_{kin} when the body arrives at the edge of the table? What is the magnitude of the force along the cord, F_F? (Friction and the mass of the cord are negligible.) (Sect. 3.2 and 5.3)

For Sect. 3.2 see also Exercise 2.4.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_3) contains supplementary material, which is available to authorized users.

Applications of NEWTON's Equation

4

Part I

4.1 Constant Acceleration in a Straight Line[C4.1]

C4.1. "Applications" in the chapter title does not mean that we will derive specific motions from Newton's equation (also known as the "equation of motion"), but rather that we will demonstrate its validity with selected examples. These are instructive and are not always found in other textbooks.

We start by introducing a useful convention: We take the force $\boldsymbol{F}$ to be the cause of the acceleration $\boldsymbol{a}$ and write the equation of motion in the form

$$\boldsymbol{a} = \frac{\boldsymbol{F}}{m}\,. \tag{3.3}$$

Our convention is completely arbitrary: In speaking, the concepts 'cause' and 'effect' are convenient and familiar, and sometimes they are even useful. *However, in the equations of physics, cause and effect play no role*.

There are several variations on our first application of Eq. (3.3); it concerns the acceleration of a body in the vertical direction. In all these experiments, we should keep a basic fact in mind: for each force, one can specify only its *point of action*, its *magnitude* and its *direction*, but never its *point of origin*. For example, the force of a spring means only "the force related to the deformation (compression, tension) of a spring".

Let us assume that two forces act on a body (more precisely, they act at its center of gravity S; cf. Sect. 5.6 and 6.2, and compare Fig. 4.1): The first force, $\boldsymbol{F}_G$, is the force of gravity, directed downwards, which we call the *weight* of the body; the second, $\boldsymbol{F}_1$, is produced by compressing a force meter (spring). This force is directed upwards, and its magnitude can be read off the scale of the force meter.

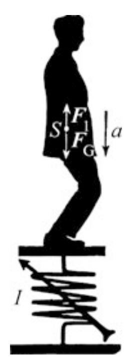


Figure 4.1 The downwards acceleration $\boldsymbol{a}$ of a man doing knee-bends

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_4

We observe accelerations only so long as $\boldsymbol{F}_1$ *and* $\boldsymbol{F}_\mathrm{G}$ *have differing magnitudes.* The direction of the acceleration (upwards or downwards) depends on whether the force $\boldsymbol{F}_1$ or $\boldsymbol{F}_\mathrm{G}$ has the larger magnitude.

The test body in the first experiment is a man, as sketched in the figure; the force meter is a common spring balance. $\boldsymbol{F}_\mathrm{G}$ is the downwards-directed *weight* of the man, and $\boldsymbol{F}_1$ is the oppositely-directed (upwards) force of the spring. We can make three observations, in logical order:

1. The man stands at rest. The spring balance shows his weight (e.g. 687 newton $\widehat{=}$ 70 kg[C4.2]). The weight $\boldsymbol{F}_\mathrm{G}$ and the force $\boldsymbol{F}_1$ indicated by the spring balance are equal and opposite, so that their resultant force is zero.

C4.2. The unit of force, the 'newton', has not become common in everyday use. Commercially available scales continue to use the unit 'kilogram', i.e. they indicate the mass which corresponds to the measured weight: $m = F_\mathrm{G}/g$ with $g = 9.81\,m/s^2$. A body of mass 1 kg has a weight of 9.81 N.

2. The man does an accelerated knee-bend. During his downward acceleration, the upward force read off the spring balance is smaller than his downwards-directed weight. As a result, the resultant force, and the acceleration, are directed *downwards*.

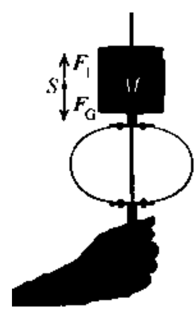
3. The man accelerates back up to his full height. During this motion, the upward force read off the spring balance is greater than his weight. The resultant force, and the acceleration, are directed *upwards*.

A variation on this experiment is often encountered in the form of a trick question: Instead of the large spring as in Fig. 4.1, we now use a sensitive spring balance; on its weighing pan is a closed bottle containing a live fly. Does the balance indicate the weight of the fly?

> The answer: When the fly maintains a *constant* altitude, or is flying upwards or downwards with a *constant* velocity, the scale reading includes the weight of the fly – case 1. (The fly can then be considered to be similar to a somewhat oversized air molecule.) During an *accelerated* downwards motion ("the fly is falling freely"), the balance indicates a weight which is too small – case 2. When the fly carries out an *accelerated* upwards motion ("flying up faster and faster"), the balance indicates a weight which is too high – case 3.

Another variant: An experimenter holds the force meter made of a circular spring, which we have already encountered, with its spring vertical (Fig. 4.2). At the upper end of the force meter is a body of

Figure 4.2 The origin of the 'elevator feeling'



mass M, acted upon by its weight $\boldsymbol{F}_{\mathrm{G}}$. If the hand holding the force meter is moving with a constant velocity upwards or downwards, the indicator reading (the compression) of the circular spring is the same as when the hand is at rest. When, however, the hand is *accelerating* upwards or downwards, the spring will be more or less strongly compressed, i.e. the upwardly-directed force $\boldsymbol{F}_1$ is larger or smaller than the downwards-directed weight $\boldsymbol{F}_{\mathrm{G}}$, respectively.

A similar experimental setup often plays an unpleasant role in our daily lives. The hand can be thought of as the floor of an elevator cabin, and we can think of the circular spring, in a drastically simplified version of anatomy, as our gut, while the body M is our stomach. When the elevator is accelerating downwards, the spring is stretched relative to its normal resting position. This relaxation is the physical basis for the much-disliked "elevator feeling", and when repeated periodically, it causes seasickness.

Finally, we consider the same experiment in a quantitative version. First, the principle: We hang a block of mass m on a beam balance (Fig. 4.3) and measure the force $\boldsymbol{F}_{\mathrm{G}}$, i.e. the weight of the block. Then an invisible mechanism causes the block to move downwards with a small acceleration. The balance deflects in a clockwise direction, its center of gravity is moved somewhat to the left and upwards. In this position, the force pulling the block upwards, $\boldsymbol{F}_1$, is smaller than the weights on the right balance pan. One would thus have to remove several small weights from the pan to return the balance to its zero position during the downwards acceleration of the block. Then, the force pulling the block upwards, $\boldsymbol{F}_1$, has the same magnitude as the remaining weights on the right pan. During the downwards acceleration, we thus have $|\boldsymbol{F}_{\mathrm{G}}| > |\boldsymbol{F}_1|$, i.e. the resultant force $\boldsymbol{F} = \boldsymbol{F}_{\mathrm{G}} + \boldsymbol{F}_1$ is directed downwards; its magnitude is $|\boldsymbol{F}_{\mathrm{G}}| - |\boldsymbol{F}_1|$. This downwards-

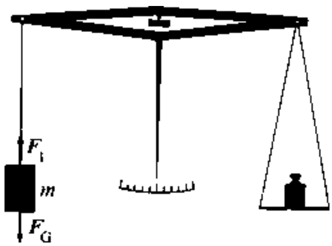


Figure 4.3 A beam balance holds a block of mass m, at rest. If the block is accelerated downwards, the indicator of the balance shows a deflection (the right pan sinks)

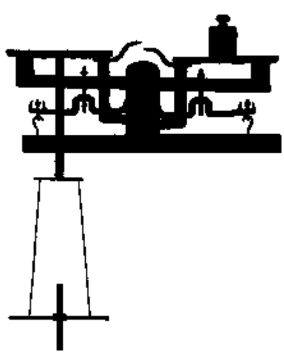


Figure 4.4 Continuation of Fig. 4.3. The object which experiences a constant downwards acceleration takes the form of a flywheel ("MAXWELL's wheel"). The balance is damped by an oil column (not shown). At the lowest point of the wheel, it changes its direction and moves upwards again. This causes a downward jerk (an impulse; cf. Sect. 5.5). This impulse is absorbed by briefly holding the pointer of the balance. **(Video 4.1)**

Video 4.1:
"MAXWELL's wheel"
http://tiny.cc/9pqujy
We can even follow the motion without intervention when the wheel reaches its upper reversal point and moves downwards again. The essential observation, namely that a body which is being accelerated downwards is lighter, no matter which direction its velocity vector is pointing, is somewhat obscured in the first part of the experiment by two effects: When the wheel initially becomes lighter, the damped oscillations of the balance make it hard to read its deflection accurately. The same problem occurs again when the balance stops moving briefly at the lower reversal point of the wheel. In the second part of the experiment (after about 75 s), when some weight has been added to the pan on the side of the balance where the MAXWELL wheel is hanging, it is in equilibrium during the downwards motion. Now, when the impulse at the lower reversal point is absorbed by the balance, it remains in equilibrium during the following upwards- and downwards motions, without the disturbing oscillations. Note that Fig. 4.4 shows the balance from the opposite side as in the video.

directed force causes the observed downwards acceleration of the block:

$$a = \frac{(|\boldsymbol{F}_\mathrm{G}| - |\boldsymbol{F}_1|)}{m}\,.$$

Practical demonstration (Fig. 4.4): The force meter is a simple kitchen scale. The object is a flywheel on a thin shaft. It is hung from two strings which are wound around the shaft. When released, it is accelerated downwards by its weight. The acceleration is measured by using the equation $s = \frac{1}{2}at^2$ and finding the time t required to travel the distance s with a stopwatch.

Numerical example

$m = 539.0\,\mathrm{g}$, $F_\mathrm{G} = 5.288\,\mathrm{N}$, $a = 0.048\,\mathrm{m/s^2}$, computed from $s = 0.83\,\mathrm{m}$ and $t = 5.9\,\mathrm{s}$. Here, $F_1 = 5.262\,\mathrm{N}$, so that $|\boldsymbol{F}_\mathrm{G}| - |\boldsymbol{F}_1| = 2.6 \cdot 10^{-2}\,\mathrm{N}$.

After the strings have rolled off to their ends, the flywheel continues rotating due to its *inertia*. The strings are again wound up, in the opposite sense, and the wheel climbs back up. One should not forget to repeat the observation in this direction of motion. Again, the indication of the force meter (balance) is *smaller* during the accelerated motion than at rest. The acceleration of the wheel is still directed *downwards*, since it is moving upwards more and more slowly, or with a "braked" motion (negative acceleration). This experiment often surprises even experienced physicists.

"This experiment often surprises even experienced physicists."

4.2 Circular Motion and Radial Forces

(Observer at rest!) First of all, as a preliminary remark, some good advice: Never draw any conclusions about circular or rotational motion before agreeing with your dialog partner (perhaps the author of the textbook!) on the frame of reference to be used. Our frame of reference was agreed upon in Sect. 2.1. It is the surface of the earth, or the floor of the lecture hall.

"…some good advice: Never draw any conclusions about circular or rotational motion before agreeing with your dialog partner on the frame of reference to be used."

Up to now, we have applied NEWTON's equation of motion only to the limiting case of a purely *path* acceleration (along a straight-line path). Now we do the same for the other limiting case, that is for a purely *radial* acceleration.

A body of mass m is supposed to move along a circular orbit of radius r with a constant angular velocity ω (the radius vector $\boldsymbol{r}$ points from the center point of the orbit towards the moving body). According to our kinematic treatment in Sect. 2.5, this motion is *accelerated*. The radial acceleration, directed towards the center point of the circular orbit, is

$$\boldsymbol{a}_\mathrm{r} = -\omega^2\boldsymbol{r}\,. \tag{2.12}$$

From the equation of motion, this acceleration of a body of mass m requires a force $\boldsymbol{F}$ directed towards the center of the orbit; we will call it the *radial force* (it is also called the *centripetal force*, 'towards

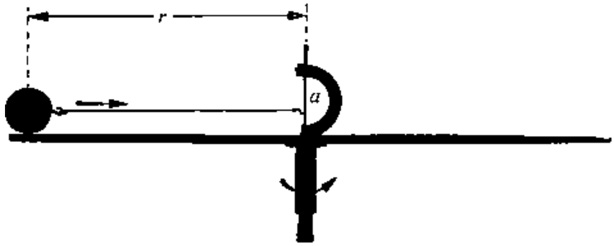


Figure 4.5 A ball on a rotating table, held by a leaf spring at the left of a. The vertical force which acts on the ball, its weight, is compensated by an invisibly small deformation of the tabletop.

the center'). Quantitatively, from the equation of motion, we must have

$$-\omega^2 \boldsymbol{r} = \frac{\boldsymbol{F}}{m} \tag{4.1}$$

(Circular frequency or angular velocity $\omega = 2\pi\nu$; ν is the mechanical frequency, i.e. number of rotations/time (e.g. rpm).)

For an experimental test of Eq. (4.1), we replace the angular velocity ω by the mechanical frequency ν and obtain

$$-4\pi^2\nu^2 \boldsymbol{r} = \frac{\boldsymbol{F}}{m}\,. \tag{4.2}$$

The radial force $\boldsymbol{F}$ could be produced by deformation of a spring; i.e. in brief, by an elastic force. We offer three examples:

1. A leaf spring produces the radial force on a ball at the outer rim of a carousel or rotating table (Fig. 4.5). It pulls in the direction of the axis of rotation (towards the center of the table) and can produce a *maximum force* of $F_{\max}$.

For this experiment, the spring is mounted below so that it can rotate, and its upper end is fixed by the holding pin a. To determine $F_{\max}$, a counter force is applied via a cord and weight. If the maximum force $F_{\max}$ is exceeded, the spring snaps out of its holder.

This spring can be used only up to a certain maximum frequency $\nu_{\max}$; this *critical* frequency can be computed from Eq. (4.2), giving

$$\nu_{\max} = \frac{1}{2\pi}\sqrt{\frac{F_{\max}}{mr}}\,. \tag{4.3}$$

Numerical example
$F_{\max} = 1.77\,\text{N}$; $m = 0.27\,\text{kg}$; $r = 0.22\,\text{m}$; then

$$\nu_{\max} = \frac{1}{2\pi}\sqrt{\frac{1.77\,\text{kg m s}^{-2}}{0.27\,\text{kg}\cdot 0.22\,\text{m}}} = 0.87\,\text{s}^{-1}$$

$$T_{\min} = \frac{1}{\nu_{\max}} = 1.15\,\text{s}\,.$$

C4.3. At the end of **Video 6.7, "Supple shaft…"** (http://tiny.cc/5squjy), we can see a grinding wheel spraying sparks tangentially. Up to the 11th edition, the following striking example could be found here:
"The spark-spraying grinding wheel seems to contradict our observations of the wheel of a car which is throwing off mud. One can cross a smooth road right behind a car with dirty wheels without being hit by the mud they throw off (as long as the wheels are not slipping!). The explanation is simple: For the observer in the moving car, the mud-throwing tires exhibit the same picture as the spark-spraying grinding wheel, i.e. tangential spraying in all directions. For the pedestrian behind the car, in contrast, the contact point of the tires with the road is the center of rotation. All the mud flies off perpendicular to the corresponding radii".

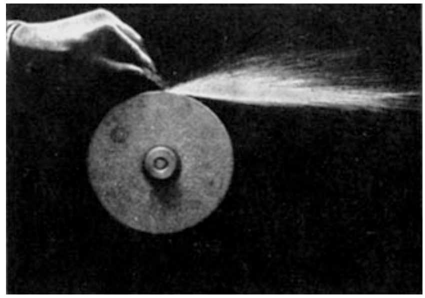


Figure 4.6 A grinding wheel, spraying sparks[C4.3]

If this limiting value is exceeded, the ball flies off. It leaves the carousel *tangentially*. When the radial acceleration is lost, the ball flies off on a *straight-line* path with constant velocity. Unfortunately, its weight generally disturbs our observation of this path. The weight converts the originally straight-line path into a parabolic falling curve; but this disturbance is less important at high orbital velocities. A good example of this phenomenon are the sparks flying off a grinding wheel. They exhibit very clearly the *tangential* escape paths. The glowing metal particles by no means fly *outwards*, away from the center of rotation (Fig. 4.6).

2. *Linear force law.* The force produced by a helical spring, F (Fig. 4.7), is directed towards the center point of the circular orbit and its magnitude is proportional to the radius of the orbit, that is

$$\boldsymbol{F} = -D\boldsymbol{r} \tag{4.4}$$

(D = spring constant).

Inserting this condition into the general equation (4.2) yields for the frequency

$$\nu = \frac{1}{2\pi}\sqrt{\frac{D}{m}}\,. \tag{4.5}$$

This means that *a body moves with a single frequency ν on a circular orbit.* The length of the orbital radius is totally unimportant. As long as this *critical* frequency ν is maintained, the body follows an arbitrary circular orbit once it has been established.

The linear force law can be implemented in various ways. In Fig. 4.7, the body is split symmetrically and mounted with as little friction as possible on two guide rods. These rods compensate the forces that we call the weights of the two bodies. The arrangement of the helical spring F allows its extension to be measured even during the rotation.

The helical spring must be pre-tensioned to the force $F = Dr_o$ while the apparatus is at rest; r_o is then the distance of the centers of gravity of the weights from their rest positions.

The experiment verifies the predictions. When the frequency is correctly adjusted, we can change the spacing r of the weights by tapping

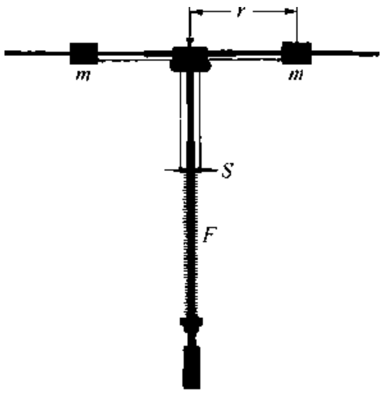


Figure 4.7 Circular motion with a linear force law. This figure also illustrates the scheme of an "astatic" frequency controller for all types of motors. When deviations from the critical rotation frequency occur, the two bodies move either all the way outwards or all the way towards the center. The disk *S* can then activate a control element of the machine and restore the critical rotation frequency.

the disk-shaped endplate *S* of the spring with a finger, increasing or decreasing it at will. The bodies follow their circular orbit for *any* radius. At this critical frequency ν, the weights are in an *indifferent* neutral equilibrium, similar to that of a ball which is at rest on a flat, horizontal table.

3. *Nonlinear force law*. Let the magnitude of a force produced by a spring and directed towards the center of the circular orbit be, for example, proportional to r^2, with a new force constant D'; that is

$$\boldsymbol{F} = -D' r^2 \frac{\boldsymbol{r}}{r} \tag{4.6}$$

($\boldsymbol{r}/r =$ is a unit vector in the direction of $\boldsymbol{r}$).

Inserting this condition into the general Eq. (4.2) for the radial force yields the frequency

$$\nu = \frac{1}{2\pi}\sqrt{\frac{D'}{m}\,r}\,. \tag{4.7}$$

The frequency ν is now dependent on the radius r. For each frequency, there is only *one* possible orbital radius r. In this orbit, the body is in a *stable equilibrium*, similar to a ball resting at the bottom of a bowl.

Experimentally, such a nonlinear force law can be implemented for example by using a circular spring as in Fig. 4.8. During rotation, the system can easily be perturbed, e.g. by tapping on the disk *S*. Following the perturbation, the correct value of r is immediately reestablished.

Up to now, our demonstration experiments for radial acceleration by means of a radial force involved rotating bodies with a very simple shape. They were "small" balls or blocks. We could neglect their diameters relative to the orbital radius r without making a significant error. They could be considered, put briefly, to be *pointlike* ("point masses"). Our final example illustrates the rotation of a more complex body, namely a chain ring.

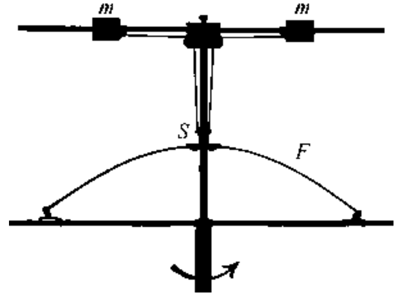


Figure 4.8 Circular motion with a nonlinear force law. This figure also illustrates the scheme of a rotational frequency meter or tachometer. Each frequency corresponds to a particular value of the radius r. The corresponding position of the disk S can be read off a scale with the aid of a pointer.

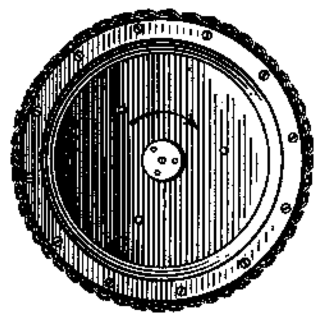


Figure 4.9 A chain on a flywheel, for demonstrating dynamic stability

In a preliminary experiment, a tightly-fitting chain is stretched around a flywheel (Fig. 4.9). If the individual links of the chain were not attached to one another, they would fly off tangentially like the sparks from a grinding wheel, once the flywheel was set in motion.

Linked together as a chain, however, they all behave in the same sense, namely by an expansion of the chain. This deformation produces forces F', which lead to a radial force F on each link, as derived in Fig. 4.10. This force F accelerates each link of the chain towards the center of the circular orbit. From Eq. (4.1), we obtain the magnitude of F (using the orbital velocity $u = \omega r$):

$$F = \frac{F'd}{r} = \frac{mu^2}{r} \,. \tag{4.8}$$

At a high rotational frequency of the flywheel, one can then push off the chain by tapping it from the side. It does not sink slackly together, but rather rolls like a stiff circular ring across the table. It even jumps over obstacles along the way. In this form, the experiment offers a good example of "dynamic stability".

A variation on this experiment is even more instructive. A long chain is hung on a rotating sprocket wheel (Fig. 4.11). Each section of this chain can be closely approximated by a circular segment with a variable radius r (the radius of curvature, see Sects. 4.4). Within these segments, the forces F' produce the radial force F directed towards their momentary centers of curvature (midpoints of the circles), as required for circular motion: Eq. (4.8). This force varies along the chain. F decreases as $1/r$; it thus

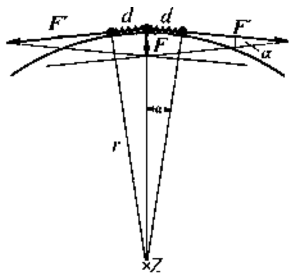


Figure 4.10 The origin of the radial force in a chain ring under tension. – We can imagine that a chain is stretched around a resting circular disk; the chain consists of balls with a spacing d connected by helical springs under tension. In the drawing, only 3 balls and 2 springs are shown. The long arrows begin at the center of gravity of the middle ball and represent the forces that act on it from the two springs, F'. Vector addition yields the net force F which is directed towards the center of the circular orbit. The quantitative relation between F' and F is found from the similarity of the isosceles triangles which have the same sides and an opening angle α: $F/F' = d/r$.

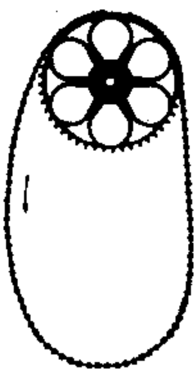


Figure 4.11 Oval shape of a bicycle chain before being thrown off the sprocket wheel **(Video 4.2)**

becomes smaller when the chain is more stretched out. *Therefore, the chain loop should be able to rotate in stable equilibrium not only as a circular ring, but also in any other arbitrary shape!* The results of the experiment verify this expectation. A bicycle chain can be conveniently used for this experiment. At a sufficiently high rotation frequency, it is thrown off the sprocket wheel.

In earlier times, this experiment was sometimes demonstrated involuntarily when a drive chain was accidentally thrown off its pulleys in a factory, etc.

Video 4.2:
"Dynamic stability of a bicycle chain"
http://tiny.cc/cqqujy
In order to demonstrate that "the chain loop should be able to rotate in stable equilibrium not only as a circular ring, but also in any other arbitrary shape", its rotation is shown in slow motion in the second part of the video. It is helpful to look at the still images one after another; one can then see how a bulge in the chain caused by colliding with an obstacle is maintained during continued rotation.
For other experiments involving chains, see e.g. https://en.wikipedia.org/wiki/Self-siphoning_beads.

4.3 Sinusoidal Oscillations: The Gravity Pendulum as a Special Case

In Chap. 2, we limited our considerations of kinematics to the simplest orbits, and likewise for the dynamics treated in the present chapter: We have treated only motion in a straight line and circular orbits.

For linear motion, there is only a linear (path) acceleration, and for a circular orbit with constant rotational velocity, only a radial acceleration. Sects. 4.3–4.9 will deal with linear pendulum oscillations and some motions around a central point. We treat the massive moving bodies to a good approximation as *point masses*. We start by describing these motions kinematically and then discuss how they are produced by forces (dynamics).

The simplest of all periodically-repeated motions takes place along a straight-line path (linear oscillation). A graphic representation of its time evolution yields the *curve of the oscillation* which can be described mathematically by a single sine function (Fig. 4.12). Oscillations of this type, and their curves, are termed *sinusoidal*, and in general one speaks of a *harmonic oscillator*.

We recall Sect. 1.8: If N oscillations occur within a time t, then $N/t = \nu$ is the *frequency* and $t/N = T$ is the *period* or oscillation time of the motion. – The abscissa of the sine function is an angle α, whose meaning can be read off Fig. 4.12. – In the oscillation curve, α is called the *phase angle* or simply the *phase*; it increases proportionally to the time, so that $\alpha = \omega t$. The significance of the proportionality factor ω can readily be seen: at $t = T$, $\alpha = 360° = 2\pi$ (cf. Sect. 1.5). For arbitrary phase angles, we have $\alpha = t \cdot 2\pi/T = \omega t$. Thus, $\omega = 2\pi/T = 2\pi\nu$, i.e. the 2π-fold of the frequency ν. ω is called the *circular frequency* (cf. Sect. 2.5).

Owing to the proportionality of the phase angle to time, one can plot the oscillation curve either against the phase angle α or the time $t = \alpha/\omega$. Both methods are illustrated in Fig. 4.12.

The ordinate of the graph of an oscillation curve is not simply the angular function $\sin\alpha$, but instead the *deflection* x (the displacement of the massive body of the pendulum from its rest position), which is proportional to a sine function, i.e.[C4.4]

C4.4. Since only the x direction occurs in the equations in this section, it suffices to use only the x component of the vector quantities *velocity, acceleration* and *force*. The sign however must be taken into account; a minus sign means 'directed opposite to the positive x direction'.

$$x = x_0 \cdot \sin\alpha = x_0 \cdot \sin\omega t. \tag{4.9}$$

x is the deflection (momentary value) at the time t, and x_0 is its maximum value. x_0 is often called the *amplitude* of the oscillation. (Instead of x_0, the symbol A is convenient when one wishes to distinguish several different amplitudes by using indices.)

In sinusoidal oscillations, not only the deflections x, but also the velocities $u = \mathrm{d}x/\mathrm{d}t$ and the accelerations $a = \mathrm{d}^2x/\mathrm{d}t^2$ can be represented by sine curves. By differentiating one and two times, one obtains

$$u = \frac{\mathrm{d}x}{\mathrm{d}t} = \omega x_0 \cos\omega t = \omega x_0 \sin\left(\omega t + \frac{\pi}{2}\right), \tag{4.10}$$

$$a = \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\omega^2 x_0 \sin\omega t = \omega^2 x_0 \sin(\omega t + \pi)\ . \tag{4.11}$$

Figure 4.12b shows u/ω, while Fig. 4.12c shows a/ω^2 graphically as functions of the time t.

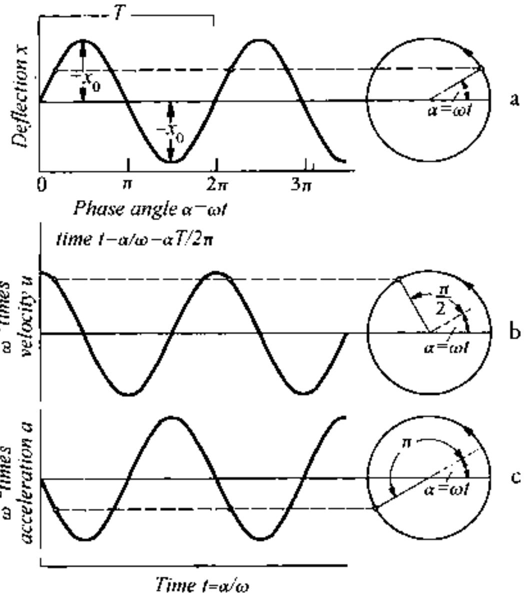


Figure 4.12 The time evolution of the deflection (position of the massive body of the pendulum), the velocity, and the acceleration for a sinusoidal oscillation. The period T is in general the time between two identical deflections.

The sinusoidal oscillation curve of the velocity precedes the curve of the deflection by a "phase shift" of $\pi/2 = 90°$; i.e. the positive, upwardly-directed values begin one-fourth of a period ($T/4$) earlier than those of x. At the time $t = 0$, $t = T/2$, $t = T$ etc., the oscillating body passes through its rest position (zero deflection). Then in Eq. (4.10), the sine function = 1, and the velocity has its maximum value

$$u_0 = \omega x_0. \tag{4.12}$$

The sinusoidal oscillation curve of the acceleration has a phase shift of $\pi = 180°$ relative to the deflection x. This means, in words: the direction of the acceleration is at every moment opposite to the direction of the deflection. As a result, Eqns. (4.9) and (4.11) can be combined to yield

$$a = -\omega^2 x. \tag{4.13}$$

This concludes the kinematic description. In order to treat the dynamical explanation of a sinusoidal oscillation, we have to consider also the equation of motion $a = F/m$; then we obtain

$$F = -m\omega^2 x$$

Part I

or, with the abbreviation

$$D = m\omega^2, \tag{4.14}$$
$$F = -Dx. \tag{4.15}$$

In words: To produce a sinusoidal oscillation, one requires a *linear* force law. The force which accelerates the massive body must be proportional to the magnitude of the deflection and oppositely directed[C4.5].

C4.5. The line of reasoning sketched here can be reversed and then leads to the following direct application of the equation of motion: Inserting the linear elastic force of a spring, $F_D = -Dx$, into the equation of motion $F = ma$, we obtain as the solution of the resulting differential equation a sinusoidal motion, $x = x_0 \sin \omega t$ (Eq. (4.9)), with the frequency $\nu = \frac{1}{2\pi}\sqrt{\frac{D}{m}}$ (Eq. (4.16)).

A linear force law can be obtained in various ways. The simplest is to make use of the force due to the deformation (tension, compression) of a spring (i.e. an "elastic" force). Thus, one arrives for example at the arrangement sketched in Fig. 4.13: A body of mass m is held between two helical springs. D, the proportionality factor between the force of the springs and the deflection, is the "spring constant" which we have already encountered. In general, it is called the *constant of the restoring force*.

From Eq. (4.14) with $\omega = 2\pi\nu$, we find the frequency

$$\nu = \frac{1}{2\pi}\sqrt{\frac{D}{m}}. \tag{4.16}$$

This equation is not new to us. We have already encountered it for motion on circular orbits in the special case of a linear force law (Sect. 4.2). There, we found that the frequency is independent of the orbital radius; here, it is independent of the amplitude of the oscillations. The frequency is determined in both cases uniquely by the quotient (force constant D/mass m).

Even with qualitative experiments (comparing wooden and iron balls of the same diameter), one can see the decisive influence of the mass of the oscillating body on the frequency (or on its inverse, the oscillation period). The effect of a larger mass can be compensated by increasing the spring constant, etc. Equation (4.16) belongs among the most important formulas in physics. For that reason, measurements of the frequency ν for different values of m and D are among the most popular and useful exercises in beginning practical laboratory courses. – The experimental setup can take many forms. A massive object can simply be hung from a helical spring (Fig. 4.14a). In the rest position, the quotient of weight/spring extension yields the spring constant D. The weight, which is constant in magnitude and direction

C4.6. In order to compensate the weight of the pendulum body (ball), which is directed downwards, in a horizontal arrangement we could replace the ball by a nearly frictionless rider on an air-track; the air cushion would then take up the weight of the rider.

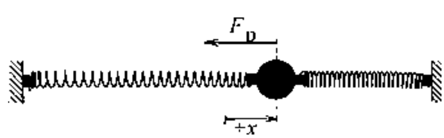


Figure 4.13 Implementation of a "linearly polarized" sinusoidal oscillation with a simple elastic oscillator or "ball and spring pendulum"[C4.6]

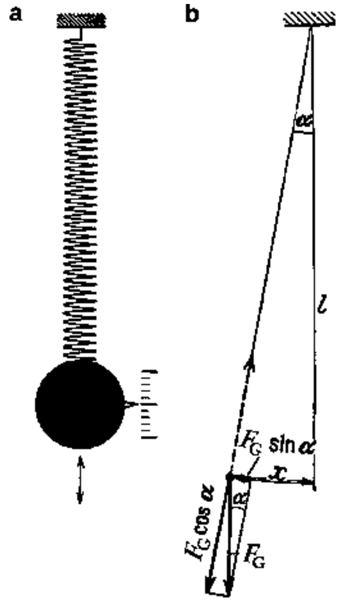


Figure 4.14 **a** A vertically oscillating spring pendulum for testing Eq. (4.16), **b** a gravity pendulum. Its angular displacement α is negative in this figure

and is thus simply an additional constant "background" force, has no influence on the frequency.

> The linear force law is only a special case. Nevertheless, it is of great importance; for in every oscillatory system, no matter how complicated, one can replace the true force law by a linear force law. In that case, the motion must be limited to sufficiently *small* amplitudes (at very small amplitudes, every force law is approximately linear).
> Mathematically, this means that every force law $F = -f(x)$ can be expanded as a series:
>
> $$f(x) = D_0 + D_1 x + D_2 x^2 + \ldots$$
>
> The constant D_0 must be zero, since the force must vanish for $x = 0$. With sufficiently small values of x (small amplitudes), the series can be terminated after the first term, so that one obtains $F = -D_1 x$.

Another example of this type of oscillator is the well-known *gravity pendulum* (also known as a *simple pendulum*). At small amplitudes, the construction sketched in Fig. 4.14b applies. It illustrates how the force acting on the pendulum ball, its weight F_G, can be decomposed into two components. One of them, $F_T = F_G \cos\alpha$, serves only to maintain tension on the cord. The other, $F_R = F_G \sin\alpha$ (the "restoring force"), accelerates the ball along its (circular) orbit.

At small angular displacements α, the orbit can be approximated as linear; furthermore, one can set $\sin\alpha \approx x/l$. Then for displacement angles of less than 4.5°, the relative errors of these approximations are less than 10^{-3}. (Note that α is negative in the figure. The restoring force F_R is always opposed to the momentary displacement x of the pendulum). We thus find $F_R = -F_G \sin\alpha \approx -F_G x/l$. That is, the restoring force is proportional to the displacement x. The proportionality factor F_G/l is the constant D of the restoring force. – The mass m of the pendulum ball, its weight F_G, and the acceleration of gravity $g = 9.81\,\text{m/s}^2$ are related by the equation $F_G = mg$. Therefore, $D = mg/l$. Inserting D into the general equation for the frequency of oscillatory motion (4.16) yields[1]

$$\frac{1}{\nu} = T = 2\pi\sqrt{\frac{l}{g}}\,. \tag{4.17}$$

Numerical example

$l = 1\,\text{m}$; $T = 2.006\,\text{s}$, i.e. half of an oscillation cycle in 1 s. This is the so-called "seconds pendulum". The longest gravity pendulum in the physics lecture hall in Göttingen has a length of $l = 11.4\,\text{m}$ and $T \approx 6.8\,\text{s}$[C4.7].

C4.7. The longest gravity pendulum used in the old physics lecture hall in Göttingen was the FOUCAULT pendulum, which hung from the roof ridgebeam above the hall (Sect. 7.7 and **Video 7.5**, as in Fig. 7.21, **"FOUCAULT's pendulum"**, http://tiny.cc/luqujy).

Frequency and period of the gravity pendulum are thus independent of the mass of the pendulum ball. The gravity pendulum accordingly occupies a special position. It should be treated as a special case and should not be the first example introduced in the study of sinusoidal oscillations.

Equation (4.16) is important for metrology (the science and technology of measurements). The *periodic repetition* of motion makes it possible to determine the oscillation period T of a pendulum very precisely. Thus, Eq. (4.17) is suitable for the task of obtaining reliable values of the *acceleration of gravity* (Sect. 2.4). The necessary precondition is the best possible approach to a "point mass" hanging from a "massless" cord (i.e. a *mathematical gravity pendulum* or simple pendulum).

4.4 Motions Around a Central Point

In the case of sinusoidal oscillations of a spring pendulum, the acceleration was seen to be *no longer constant*, but the path of motion was still a straight line. The additional velocity d$\boldsymbol{u}$ attained within the time interval dt remained always in the direction of the velocity at the beginning of the time interval, $\boldsymbol{u}$, either increasing it (Fig. 4.15a) or decreasing it (Fig. 4.15b). Only a *path* acceleration was present. In the general case, however, the vectors d$\boldsymbol{u}$ and $\boldsymbol{u}$ may differ in direction by an arbitrary angle α (Fig. 4.15c). Then both path and radial (transverse) accelerations are present at the same time. Both are components of the *total acceleration* $\boldsymbol{a}_g$ (Fig. 4.16). The path acceleration $\boldsymbol{a}$ changes the *magnitude* of the orbital velocity. The radial acceleration $\boldsymbol{a}_\varrho$ produces the *curvature* of the orbit. Its magnitude, from Eq. (2.12), is $a_\varrho = u^2/\varrho$. Here, ϱ is the *radius of curvature*, which

[1] The influence of the amplitude α_0 on the period T is small. One finds values of T for $\alpha_0 = 5°$ which are too large by 0.048 %, for $\alpha_0 = 10°$ by 0.19 %, for $\alpha_0 = 15°$ by 0.43 % and for $\alpha_0 = 20°$ by 0.76 %.

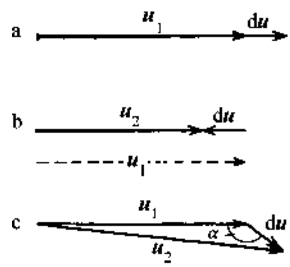


Figure 4.15 The definition of the total acceleration

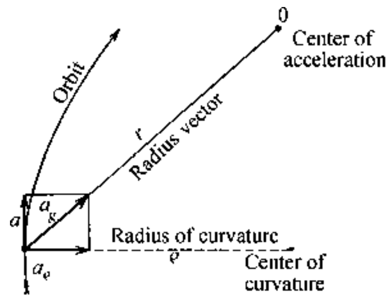


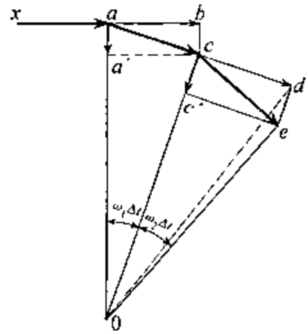
Figure 4.16 Decomposition of a central acceleration $\boldsymbol{a}_g$ into two components $\boldsymbol{a}$ and $\boldsymbol{a}_\varrho$

points from the momentary *center of curvature*. The latter is the center point of the circle which most closely approximates the short segment of the orbit that one is considering at a particular time. From the enormous variety of possible motions (just think of the possible motions of our fingers, arms and legs!), we choose a single group for closer examination, that of the *motions around a central point*.

A central motion is the motion of a body (a "point mass") on an arbitrary planar orbit in which an acceleration of varying magnitude and direction always points towards a single point, the center of acceleration. The line connecting the moving body with the center of acceleration is called the "radius vector". From this definition, motion around a circular orbit as well as linearly polarized pendulum oscillations are both limiting cases of central motion. In the case of a circular orbit with constant rotational frequency, there is no orbital acceleration; in the case of linearly-polarized pendulum oscillations, there is no radial acceleration. For a generalized central motion, two simple propositions hold. First: The motion is played out in a plane. Second: The radius vector sweeps out equal areas in equal times (the "area law"). – Both propositions belong to kinematics. They are geometric consequences of the premise that the acceleration is arbitrary but always directed towards the same central point.

The area law can be seen from Fig. 4.17. This figure is based on Fig. 2.12. Three curved segments of the orbit of a central motion, corresponding to three successive time intervals, are approximated by the three arrows *xa*, *ac*, and *ce*. The central acceleration increases in going from left to right. The thin arrows *ab* and *cd* show the continuation of the motion from the previous time interval with the same velocity along the tangent to the orbit.

Figure 4.17 The area law



The arrows aa' and cc' are the *accelerated* paths traversed in the same time intervals Δt towards the central point O. All the arrows are in the plane of the page, so that the orbits remain planar. The radius vectors Oa, Oc, Oe etc. sweep out equal areas in equal times Δt:

$$\text{Area} = \begin{cases} \Delta Oac = \Delta Ocd, & \text{since by construction } ac = cd, \\ \Delta Ocd = \Delta Oce, & \text{since the altitudes of the triangles are} \\ & cd = c'e \\ \hline \Delta Oce = \Delta Oac & \text{area law.} \end{cases}$$

Demonstration experiment: The cord of a catapult holding a stone which is moving on a circular orbit passes through a short, smooth tube in the left hand of the experimenter. The right hand is pulling on the cord to shorten the length of the radius vector r. The angular velocity ω increases proportionally to $1/r^2$.

4.5 Elliptical Orbits and Elliptically Polarized Oscillations

Central motions need not follow closed orbits; think for example of a spiral orbit. But one group of closed orbits among the central motions is particularly important: these are the *elliptical* orbits. We can distinguish two cases:

1. *Elliptically-polarized oscillations* ("polarized" refers to the "shape" of the oscillations). The center of acceleration of the orbiting body lies at the *midpoint* of the ellipse, at the intersection of its two principal axes.

2. KEPLER's *elliptical orbits.* The center of acceleration of the orbiting body is at one of the two *focal points* of the ellipse (Sect. 4.7).

In this section, we treat elliptically-polarized oscillations. Kinematically, they are produced by the superposition of two perpendicular linearly-polarized, sinusoidal oscillations of the same frequency. The shape of the ellipse is determined by the ratio of the amplitudes of the

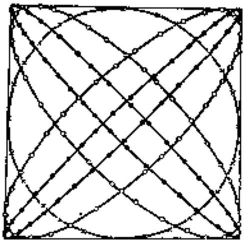


Figure 4.18 The envelope of elliptical oscillation orbits for equal amplitudes of the two perpendicular component oscillations **(Video 4.3)**

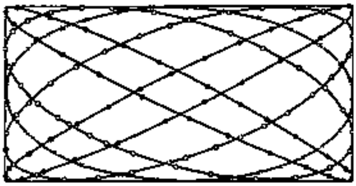


Figure 4.19 The envelope of elliptical oscillation orbits for unequal amplitudes of the two perpendicular component oscillations **(Video 4.3)**

two component oscillations and by their phase difference $\Delta\varphi$. The phase difference plays the more important role.

The set of all possible orbits has a square envelope (Fig. 4.18). When the two amplitudes are unequal, it degenerates into a rectangle (Fig. 4.19)[C4.8].

We summarize: For the kinematic demonstration of an elliptically-polarized oscillation of arbitrary shape, two perpendicular linearly-polarized sinusoidal oscillations of the same frequency but with adjustable phase difference are sufficient. When their phase difference is 0° or 180°, the ellipse degenerates into a line. At a phase difference of 90° or 270°, a circularly polarized oscillation, i.e. a circular orbit, can result. In that case, the two individual amplitudes must be equal.

4.6 LISSAJOUS Orbits

The results and the methods of the preceding section can also be applied without difficulties to treat the most general case of elastic oscillations. We limit ourselves here to giving a summary overview.

When the frequency difference between the two component oscillations becomes large, the change in phase difference is already noticeable after a single orbital passage. The ellipse becomes deformed. We see the characteristic picture of a planar LISSAJOUS figure. Figures 4.20 and 4.21 show several examples of LISSAJOUS orbits. Their shape depends on two factors:

1. the ratio of the frequencies of the two component oscillations;

2. the phase difference of the two component oscillations at their rest point at the beginning of the experiment.

Video 4.3:
"Circular oscillations"
http://tiny.cc/1pqujy
With a simple experiment which is described in detail in Vol. 2, Chap. 9, last section, the elliptical orbits in Figs. 4.18 and 4.19 can be demonstrated. Two leaf springs mounted perpendicular to each other carry disks with slits which can allow a light beam to pass in order to follow their motions by projection onto a screen. After demonstrating the excitation of the individual oscillations, both springs are excited simultaneously but with varying phase differences.

C4.8. For the experimental demonstration of elliptical oscillations, POHL describes some impressive mechanical arrangements. For reasons of space, they are not included in this edition. However, we mention the convenient possibilities of an electrical demonstration using frequency generators and an oscilloscope. A simple mechanical experiment is shown in **Video 4.3**.

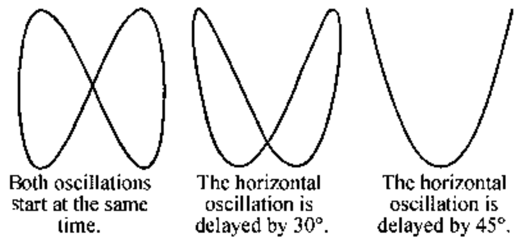


Figure 4.20 LISSAJOUS figures for a frequency ratio of 2:1 for the two perpendicular component oscillations. The vertical oscillation has the higher frequency. (J.A. LISSAJOUS, 1822–1880)[C4.9]

C4.9. J.A. Lissajous, *Comptes Rendus des Séances de l'Académie des Sciences* **41**, pp. 93 and 814 (1855).

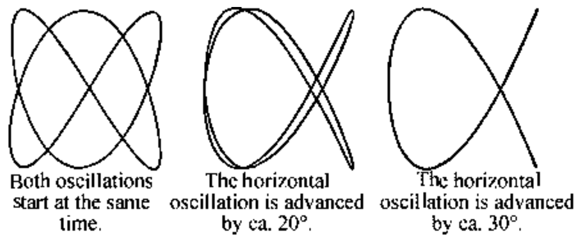
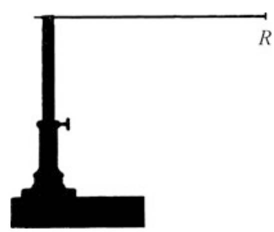


Figure 4.21 LISSAJOUS figures for a frequency ratio of 3:2 for the two perpendicular component oscillations. The vertical oscillation has the higher frequency.

Figure 4.22 The production of LISSAJOUS orbits using the bending oscillations of a rod of rectangular cross section (2 mm wide, 3 mm high). The small mirror R reflects the image of a pointlike light source onto the screen which is perpendicular to the rod. **(Video 4.4)**



Video 4.4:
"LISSAJOUS figures"
http://tiny.cc/6pqujy
First, the horizontally and the vertically polarized, damped component oscillations are shown individually. Then both are excited simultaneously. From the comparison with the LISSAJOUS figures shown in Fig. 4.21, we can see that the frequency of the vertical oscillation has a ratio of roughly 3:2 to that of the horizontal oscillation.

A mechanical method of producing LISSAJOUS orbits is shown in Fig. 4.22. The frequencies of the horizontal and the vertical component oscillations are in the ratio of about 2:3. One can readily observe them after giving a horizontal or a vertical excitation 'kick'. – The LISSAJOUS image sequence is the well-known one shown in Fig. 4.21.

4.7 KEPLER's Elliptical Orbits and the Law of Gravity

"For demonstration experiments, they are truly tortuous".

KEPLER's elliptical orbits have played a fundamental role twice in the history of physics: Once in the development of celestial mechanics, and a second time in BOHR's model of the atom. For demonstration experiments, they are truly tortuous. They cannot be demonstrated with simple and quickly understandable means (cf. **Video 4.5**).

In a KEPLER orbit, the center of acceleration is located at one of the two focal points of the ellipse. A KEPLER orbit results when the object has an initial velocity which does not point towards the center. The acceleration at every point is inversely proportional to the square of the distance (the length of the radius vector r), so that

$$a = \frac{\text{const}}{r^2}\,. \tag{4.18}$$

The derivation of this equation can be found in every textbook on theoretical physics.

How can we implement the acceleration required by Eq. (4.18) physically? The answer was first found on the basis of astronomical observations, by ISAAC NEWTON.

The moon orbits around the earth. Its orbit is very nearly circular. Its radius is – take note of this number – equal to 60 earth radii. Kinematically, we have already described the moon's orbit in Sect. 2.5: The moon has an orbital velocity of 1 km/s and experiences a radial acceleration of $a_\mathrm{r} = 2.7\,\mathrm{mm/s^2} = 2.7\cdot 10^{-3}\,\mathrm{m/s^2}$. Therefore, the ratio is

$$\frac{\text{Acceleration of gravity } g}{\text{Moon's radial acceleration}} = \frac{9.8\,\mathrm{m/s^2}}{2.7\cdot 10^{-3}\,\mathrm{m/s^2}} = 3600 = 60^2.$$

From these facts, NEWTON drew the conclusion that the same force acts on the moon as on every stone near the earth's surface. This force is directed towards the center of the earth and is called 'weight'. However, the weight of a body is, counter to all our everyday prejudices, not a *constant* force acting on a body. Instead, it varies depending on the distance r of the body from the center of the earth, and is proportional to r^{-2}. – Therefore, NEWTON set the weight of the moon not equal to $F = mg$, but rather to

$$F = \text{const}\,\frac{m}{r^2}\,. \tag{4.19}$$

And now the final conclusion was nearly inescapable: If the earth attracts the moon, then the converse must also be true; the moon must attract the earth. For an observer on the moon (a different frame of reference!), the earth has a weight. An observer supposed to be on the sun could apply the law of '*actio* = *reactio*' (another change of frame of reference!). For this observer, both forces or weights must be identical but oppositely directed. Then in general, in place of the weight, we must consider the mutual attraction of two bodies by the force

$$F = G\,\frac{mM}{r^2}\,. \tag{4.20}$$

Video 4.5:
"KEPLER's elliptical orbits"
http://tiny.cc/tpqujy
On a planar horizontal aluminum plate, an object can move with minimal friction. It is filled with liquid nitrogen, which evaporates and escapes as gas through a hole in the bottom of the object, so that it can glide suspended on the gas film. A cord attached to the object leads at a distance r to a small electric motor, which produces a constant torque M. The cord winds around a spiral spindle of variable radius ρ on the motor shaft, so that the force F transmitted by the cord depends on the momentary radius ρ, which is adjusted so that $\rho \sim r^2$, and thus also $F = M/\rho \sim 1/r^2$. After a catapult launch, the object therefore follows an elliptical KEPLER orbit.

Part I

C4.10. For the definition of the center of gravity, see Sect. 5.6.

(m and M are the masses of the two bodies and r is the distance between their centers of gravity[C4.10]. In the case of homogeneous spheres, this law holds for all values of r. For bodies of arbitrary shape, r must be large compared to the dimensions of the bodies.)

This is NEWTON's famous *law of gravity*. The proportionality factor G in this law is called the *universal gravitational constant*.

4.8 The Gravitational Constant

The constant G in the law of gravity cannot be determined from astronomical observations. It must be measured in the laboratory. – The principle of the measurement is as follows: We construct a small model of the astronomical situation. As the "earth", a large lead ball is employed (of mass M, several kg), and as the "moon" or a "stone", a smaller ball (of mass m) made of some arbitrary material. The large ball is fixed, and the smaller one is allowed to move freely as far as possible. One measures the acceleration a of the smaller ball and computes the gravitational constant G from the equation

$$a = G\frac{M}{r^2}\,. \tag{4.21}$$

Carrying out the experiment: We employ a symmetrical arrangement (Fig. 4.23). The two small balls are attached to the ends of a rod (dumbbell) which is hung at its center using a fine metal band (the torsion band, rotation axis), allowing it to rotate. Fig. 4.24 shows the silhouette of a well-tested apparatus (a torsion balance) without the large balls. To carry out the measurement, we rotate the large balls from their initial position as indicated in Fig. 4.23 as dark circles into their final position (shown as shaded circles). – Immediately after this roating the large balls, the small balls are accelerated and begin to move. A mirror and a long light pointer allow us to follow this motion along a distance s with roughly 1600-fold linear magnification. We observe it for around a minute with a stopwatch and compute the acceleration from $a = 2s/t^2$.

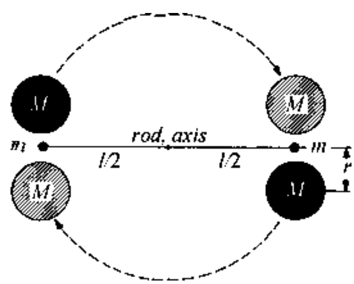


Figure 4.23 Measurement of the gravitational constant (HENRY CAVENDISH, chemist, 1798)

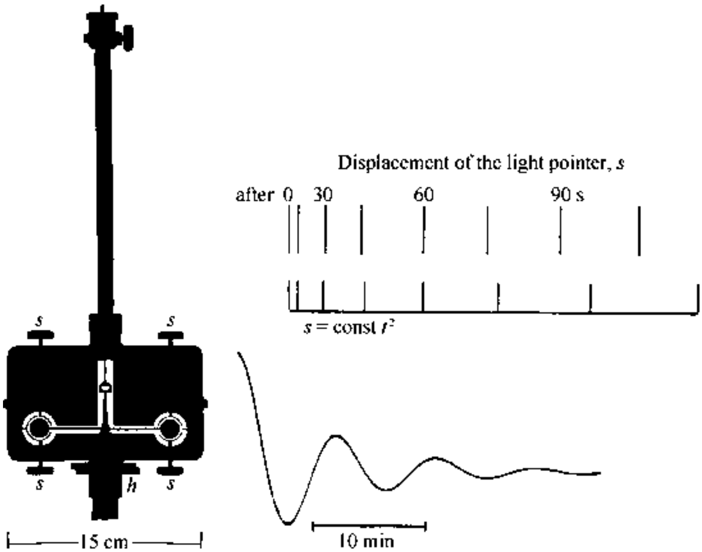


Figure 4.24 Silhouette of a torsion balance. The large balls have been removed from their movable support beam h. The dumbbell rod holding the small balls is held in a metal block closed at front and back by glass plates (good heat exchange). The screws s are used to arrest the torsional motion; they press four semicircular metal disks against the small balls. The torsion band has been highlighted in the figure so that it can be seen clearly. The mirror is mounted at one end of the band. The oscillation period T of the torsion balance is around 9 minutes. The lower part of the figure on the right shows the damped ringing-down of the oscillations. The upper-right inset of the figure illustrates the measurement of the acceleration a. The position of the light pointer after each 15 s was photographed at 108-fold magnification.

At the start of the experiment (immediately after rotating the large balls), the distance r of the centers of the balls and the twisting of the torsion band are practically unchanged and therefore, the acceleration is nearly constant. However, the torsion band was already twisted up to a maximum value at the beginning of the experiment: In the rest position, the attractive forces between the balls were in equilibrium with the torsional force due to the twisting of the band. As a result, after rotating the large balls, the resulting acceleration a of the small balls is exactly twice as large is if we had brought the large balls from far away in to the distance r.

Numerical Example

$M = 1.5\,\text{kg}$; $r = 4.75\,\text{cm}$; the length of the dumbbell rod is $l = 10\,\text{cm}$. The light pointer has a length of $L = 40\,\text{m}$, giving a linear magnification of the distance of $V = 2L/(l/2)$. The factor 2 is justified by the following argument: A rotation of the mirror by the angle α rotates the reflected light beam by the angle 2α. On the screen, the measured acceleration of the light pointer is $a = 1.28 \cdot 10^{-2}\,\text{cm/s}^2$. From this we obtain $G = 6 \cdot 10^{-11}\,\text{m}^3/(\text{kg}\,\text{s}^2)$.

C4.11. Compared with other physical constants, the universal gravitational constant is the least precisely known. Only four places after the decimal point are considered to be certain, mostly measured using the principle of the CAVENDISH balance, but to some extent also in experiments with falling bodies (see e.g. http://en.wikipedia.org/wiki/Gravitational_constant#History_of_measurement, and https://royalsociety.org/events/2014/gravitation/).

Precision measurements of the gravitational constant[C4.11] yield a value of

$$G = 6.6738 \cdot 10^{-11} \, \frac{\mathrm{m}^3}{\mathrm{kg\,s}^2} \,. \tag{4.22}$$

The experimental determination of the gravitational constant G represented a great step forward: using its value, the mass of the earth could be computed. – The earth's surface is at a distance of $r = 6400\,\mathrm{km} = 6.4 \cdot 10^6\,\mathrm{m}$ from the center of the earth. At the earth's surface, the acceleration due to the force of gravity has the value $g = 9.81\,\mathrm{m/s}^2$. Inserting these two values together with G into Eq. (4.21), we obtain

$$\text{Earth's mass } M = \frac{9.81\,\mathrm{m\,s}^{-2}(6.4 \cdot 10^6)^2\,\mathrm{m}^2}{6.67 \cdot 10^{-11}\,\mathrm{kg}^{-1}\,\mathrm{m}^3\,\mathrm{s}^{-2}} = 6 \cdot 10^{24}\,\mathrm{kg}.$$

The volume of the earth amounts to roughly $1.1 \cdot 10^{21}\,\mathrm{m}^3$. As a result, the average density of the earth is found to be $= \frac{6 \cdot 10^{24}\,\mathrm{kg}}{1.1 \cdot 10^{21}\,\mathrm{m}^3} = 5500\,\mathrm{kg/m}^3 = 5.5\,\mathrm{g/cm}^3$.

This is of course only a mean value. The density of the stones in the earth's crust is on the average $2.5\,\mathrm{g/cm}^3$. Therefore, in the interior of the earth, we have to assume materials with a higher density. Indications are that the earth's core has a high iron content.

4.9 The Law of Gravity and Celestial Mechanics

The discovery of a general attractive force between all massive bodies is rightly counted among the great achievements of the human mind. NEWTON's law of gravity not only correctly predicts the motion of the moon; it is the dominant principle throughout the whole of celestial mechanics, which describes the motions of the planets, the comets, binary stars, etc.

The observations of the motions of the planets were summarized by JOHANNES KEPLER (1571–1630) in the form of three laws. These "KEPLER's laws" state that:

1. Each planet moves on a plane around the sun. Its orbit is an ellipse with the sun at one of the two focal points.

2. The radius vector of each planet sweeps out equal areas on the orbital plane in equal times.

3. The squares of the orbital periods T (planetary 'years') are in the ratios of the cubes of the semimajor axes of their orbits.

> The deviations of the elliptical orbits from circular orbits for the eight planets in the Solar System are rather small. If one for example draws the orbit of Mars on a sheet of paper, giving its semimajor axis a length of 20 cm, the deviation of the orbital shape from a circle is everywhere less than 1 mm. Given these numbers, we can appreciate KEPLER's achievement in verifying the elliptical orbit of this planet.

The three laws formulated by his great predecessor could be explained in a unified way by NEWTON using his law of gravity[2]:

1. Every elliptical orbit requires a central (centripetal) acceleration. In the elliptical orbits observed by KEPLER, one focus was distinguished over the other. According to Sect. 4.7, the accelerations should be proportional to $1/r^2$. This is precisely the case from Eq. (4.20) for mutually attractive (gravitational) forces.

2. KEPLER's second law is just the law of areas, which holds for every central motion (cf. Sect. 4.4).

3. KEPLER's third law likewise follows from Eq. (4.20). This can be seen by considering a special case; we allow the KEPLER ellipse to become a circle. For a circular orbit, we have (cf. Eq. (4.2)):

$$\boldsymbol{F} = -4\pi^2 m v^2 \boldsymbol{r} = -\frac{4\pi^2 m\boldsymbol{r}}{T^2}\,. \tag{4.23}$$

For the magnitude of $\boldsymbol{F}$, we insert the value obtained from the law of gravity (Eq. (4.20)). Then we obtain

$$\text{const}\,\frac{m}{r^2} = \frac{4\pi^2 mr}{T^2}, \quad T^2 = \text{const}\,r^3\,. \tag{4.24}$$

In contrast to the planets, *comets* often have extremely long elliptical orbits. The semimajor axis of the ellipse can be as much as 100 times its semiminor axis. But even for the general case of these arbitrarily elongated elliptical orbits, KEPLER's third law can be derived as a result of NEWTON's law of gravity. That, to be sure, requires more extensive computations.

A simple example to conclude this section can serve to imprint the most important aspects of celestial mechanics in the readers' memories:

We imagine a projectile to be fired in a horizontal direction near the earth's surface. The atmosphere (and with it the resistance of the air) are assumed not to exist. How great must the velocity u of the projectile be so that it becomes a little moon, circling the earth at a constant distance from the surface?

A circular orbit with the orbital velocity u requires a radial acceleration of $a = u^2/r$ according to Eq. (2.12). This radial acceleration

[2] KEPLER himself never progressed beyond qualitative efforts at explaining his laws. For example, in 1605, he wrote, "If we were to suppose that the earth is at rest in some particular place, and were to set another, larger earth nearby, it would be attracted by the earth just as stones are attracted to its surface".

Part I

is provided by the force of the weight of the projectile. The weight gives the projectile an acceleration of $g = 9.81\ \mathrm{m/s^2}$ towards the center of the earth. On the other hand, the distance of the surface of the earth from its center is equal to the radius r of the earth, about $6.4 \cdot 10^6$ m. We thus obtain

$$9.8\,\frac{\mathrm{m}}{\mathrm{s}^2} = \frac{u^2}{6.4 \cdot 10^6\,\mathrm{m}},$$
$$u = 8000\,\mathrm{m/s} = 8\,\mathrm{km/s}.$$

At a muzzle velocity of 8 km/s in a horizontal direction, we therefore have the case shown in Fig. 4.25 on the left: the projectile circles the earth close to the surface like a little satellite. This case is realized by artificial earth satellites which are placed in orbit at an altitude of around 400 km, moving in a direction perpendicular to the earth's radius.

If this initial velocity is exceeded, we obtain an elliptical orbit of the type shown in the center of Fig. 4.25. For velocities $u > 8$ km/s, the projectile will orbit the earth like a planet or a comet, on an ellipse. The center of the earth is at that focus of the ellipse which is *closer* to the cannon that fired the projectile. For velocities > 11.2 km/s, known as the *escape velocity*, the elliptical orbit degenerates into a hyperbola. The projectile leaves the earth, never to return[3].

For velocities $u < 8$ km/s, there is also an ellipse, cf. Fig. 4.25, right-hand side. But only that part of it which is not dotted in the figure can be realized in practice. In this case, the center of the earth is at the focus of the ellipse which is *further* from the cannon (the attraction of the earth's gravity acts as though the earth had shrunk to a small body around its midpoint, while maintaining the same mass).

The lower the initial velocity u, the more elongated is the elliptical orbit. Finally, one arrives at the limiting case shown in Fig. 4.26.

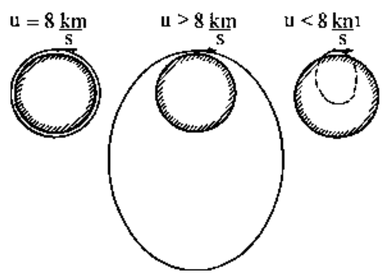


Figure 4.25 Elliptical orbits around the center of the earth for different initial velocities

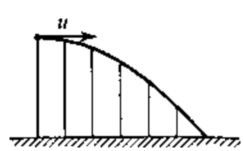


Figure 4.26 The parabolic trajectory of an object thrown horizontally

[3] For the sun, the corresponding escape velocity is 618 km/s, and for earth's moon, it is 2.3 km/s.

The center of acceleration, the midpoint of the earth, appears to be practically infinitely distant. The radius vectors which point there are nearly parallel. One finds that the portion of the elliptical orbit which remains above the earth's surface is to a good approximation a *parabola*. It is the well-known parabolic trajectory of a horizontally-thrown projectile. – These considerations are useful, although the air resistance in reality makes their practical verification impossible. Even at normal velocities of several 100 m/s, the braking action of air resistance is considerable. The parabola can be seen only as a very rough approximation to the true trajectory, the so-called ballistic curve.

Exercises

4.1 Find the amplitudes of the velocity u_0 and the acceleration a_0 of a body that is undergoing sinusoidal oscillations with a maximum displacement (displacement amplitude) of $x_0 = 10\,\mathrm{cm}$ and a period of $T = 1\,\mathrm{s}$. (Sect. 4.3)

4.2 The acceleration of gravity on the moon is 1/10 of its value g on the earth. How long is the period T_M of a simple pendulum on the moon, if its period on the earth is $T_E = 1\,\mathrm{s}$? What is the mass m_M of the moon, given its radius r_M? (Sects. 4.3 and 4.7)

4.3 A stone of mass m attached to a rope of length l is being whirled around on a circular horizontal orbit at the angular velocity ω. How does its angular velocity change if the length of the rope is suddenly decreased by l_1? Treat the stone as a "point mass". (Sects. 4.4 and 6.6)

4.4 The heat content of anthracite coal is about $3 \cdot 10^4\,\mathrm{kJ/kg}$ (cf. Sect. 19.8). At what depth z under the earth's surface could the coal lie if the work of raising it to the surface were just equal to its heat content? (See Sect. 13.3). (Sects. 4.8 and 5.2)

(In order to solve this exercise, one needs to know that the weight F_G of an object of mass m *decreases* in the interior of the earth! At the center of the earth, its weight would be zero. It then increases linearly – assuming that the matter of the earth has a uniform, constant density – with increasing distance r from the center, finally reaching the value mg at the surface of the earth ($r = R_E$). As an equation: $F_G = mg(r/R_E)$, where $R_E = 6378\,\mathrm{km}$.)

4.5 In Göttingen, there is a "Planetary Way": A sphere representing the sun is in the Goetheallee, then towards the center of town follow Mercury, Venus, Earth, etc.; all the linear dimensions are reduced by a factor of $2 \cdot 10^9$ compared to the real distances. If this model were set up in an otherwise empty space, free of external gravitational forces, could the planets circle the sun? And if so, how long would the "year" of the model Earth be? (The radius of the real orbit of the earth is $R = 150 \cdot 10^9$ m.) (Sect. 4.9)

For Sect. 4.3 see also Exercise 2.5.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_4) contains supplementary material, which is available to authorized users.

5 Three Useful Concepts: Work, Energy, and Momentum

5.1 Preliminary Remarks

Making use of NEWTON's equation of motion and the law of "*actio = reactio*", one can treat every motion quantitatively. Many motions are however extremely complex; one need think only of the motions of machine parts or the motions of our bodies and their extremities. In such cases, a description based on the equation of motion requires a great amount of computational effort. This effort can often be reduced considerably by using some additional, cleverly formulated concepts – these are *work, energy,* and *momentum*. *These concepts will not be introduced on the basis of experimental facts which we have neglected up to now, but rather by considering the equation of motion more closely.* We begin with the concept of 'work'.

5.2 Work and Power

Three things are agreed upon at the outset:

1. The product 'force in the direction of the displacement' *times* 'distance displaced' will be defined as the *work*.

2. $+Fx$ means: The force F and the displacement x are in the same direction. *The force F performs*[1] *work.*

3. $-Fx$ means: The force F and the displacement x point in opposite directions; *work is performed* against *the force* F.

"The force *F performs* work". One should avoid saying, "The force works, is working".

In general, the force is not parallel to the displacement, nor is it constant. In this more general case, we denote the components in the direction of the m-th displacement element Δx by $F_1, F_2, \ldots, F_m$; and we define the work W as the sum

$$F_1\Delta x_1 + F_2\Delta x_2 + \ldots + F_m\Delta x_m = \sum F_i\Delta x_i$$
$$(i = 1,\ 2,\ 3, \ldots,\ m)$$

[1] One should avoid saying "The force works, is working". The concept of *power* will be introduced at the end of this section.

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_5

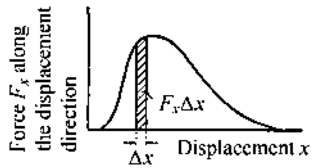


Figure 5.1 The definition of the work as a path integral over the force

or in the limit[C5.1]

$$W = \int F_x \mathrm{d}x\,. \tag{5.1}$$

C5.1. The general definition of work is given by the path integral $W = \int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{x}$ over the scalar product of the vector quantities 'force $\boldsymbol{F}$' and 'displacement element $\mathrm{d}\boldsymbol{x}$'. This also fixes the sign of the work. In all the examples discussed in this section, W is positive, i.e. work is performed on the object on which the force F acts. Minus signs occur only when the corresponding counterforces (reaction forces) are employed instead of $\boldsymbol{F}$. The sign of W does not change. Since often only the parallel components of $\boldsymbol{F}$ and $\boldsymbol{x}$ are used, whose direction is included in their magnitudes, sign errors can easily occur.

Figure 5.1 shows a graphical representation of a path integral over the force.

With this definition of the work, its units are also determined; these must consist of the product of a unit of force and a unit of distance. We call

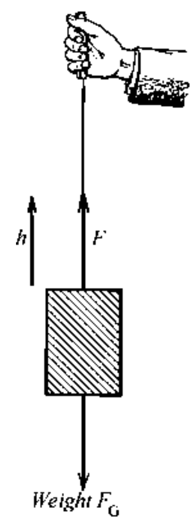
$$\begin{aligned} 1\ \text{newton} \cdot \text{meter (N m)} &= 1\ \text{Joule (J)} = 1\ \text{watt} \cdot \text{second (W s)} \\ &= 1\,\text{kg}\,\text{m}^2/\text{s}^2\,, \\ 1\ \text{kilowatt} \cdot \text{hour (kWh)} &= 3.6 \cdot 10^6\ \text{watt} \cdot \text{second (W s)}. \end{aligned}$$

We will calculate the work for three different cases:

1. *Lifting work.* In Fig. 5.2, a muscular force lifts a massive block straight up, *very slowly* with the force F. The force F performs work along the path $\mathrm{d}h$:

$$\mathrm{d}W = F\mathrm{d}h\,. \tag{5.2}$$

Figure 5.2 The definition of lifting work



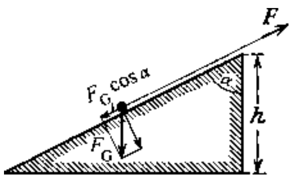


Figure 5.3 Lifting work up a ramp. The work is not performed directly against the weight F_G of the object, but rather only against its component in the direction parallel to the surface of the ramp, $F_G \cos\alpha$. However, the displacement x is greater than the perpendicular height h; it is $x = h/\cos\alpha$. The lifting work along the whole length of the ramp is thus $W = -F_G \cos\alpha \; h/\cos\alpha = -F_G h$. – Similar considerations apply to any ramp, no matter what its shape and curvature, or to other lifting machines, such as a block-and-tackle.

In a very slow lifting action, the velocity of the block lifted remains practically zero. As a result, to a very good approximation, $F = -F_G$. Then we find

$$dW = -F_G dh\,. \tag{5.3}$$

This work is performed *against* the weight of the block. The weight F_G is practically constant for all heights h in the neighborhood of the earth's surface. So the shaded region indicated in Fig. 5.1 is a *rectangle* with the area $F_G h$. We then obtain for the lifting work up to the height h against the force of weight F_G[C5.2]:

$$\text{lifting work} = -F_G h. \tag{5.4}$$

C5.2. Lifting work is also positive. Since F_G and h are directed opposite to one another, the sign of the corresponding scalar product must be taken into account: $W = -\boldsymbol{F}_G \cdot \boldsymbol{h} = mgh$.

All types of lifting machines, e.g. the simple ramp in Fig. 5.3, can do nothing to change the magnitude of the product $-F_G h$. It is always the *perpendicular height h* which determines the amount of work.

Numerical Example
A person with a mass of 70 kg climbs to the top of a 7000 m(!) mountain in one day. The force of his or her muscles perform lifting work of 70 kg · 9.81 m/s^2 · 7000 m ≈ 5 · 10^6 N m or around 1.5 kWh. This "day's work" has a market value of about 2 cents![C5.3] – When jumping, one need consider only the height h by which the center of gravity of the jumper is raised. When a person is standing, his/her center of gravity is about 1 m above the ground. In jumping over a crossbar at a height of 1.7 m (cf. Fig. 5.4), the center of gravity is raised to a height of around 2 m. The height lifted is thus only 2 m − 1 m = 1 m. So the muscular force of the jumper performs a lifting work of 70 kg · 9.81 m/s^2 · 1 m, or around 700 N m.

C5.3. This corresponds in today's currency to around 30 Eurocents per kWh for private users in Germany, and around 11 Ct. in the U.S. (**Exercise 4.4**).

2. *Work against a spring*. In Fig. 5.5, we see an object which is held by a spring. Muscular force is used to extend the spring *very slowly* in the direction x. The muscular force F performs work along the displacement element dx:

$$dW = F dx\,. \tag{5.5}$$

C5.4. High-jump techniques have been developed in order to keep the center of gravity of the jumper as low as possible. In 1968, the American high-jumper R. FOSBURY earned an Olympic gold medal with his backwards roll – the 'Fosbury flop' (sneered at by the skeptics).

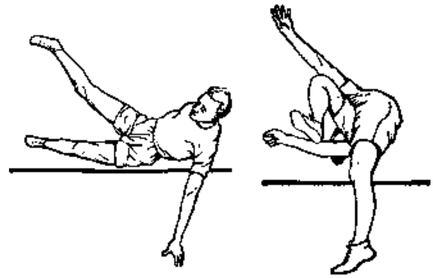


Figure 5.4 Proficient high-jumpers roll over the crossbar[C5.4]

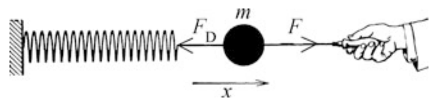


Figure 5.5 The definition of work against a spring, or "elastic work"

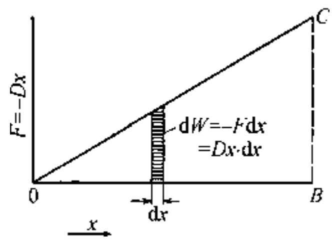


Figure 5.6 The calculation of work against a spring (elastic work): $\int dW$ = the sum of the shaded rectangular areas = area of the triangle *COB*

When the spring is stretched very slowly, the velocity of the object remains practically zero. As a result, to a very good approximation, the elastic force resulting from the extension of the spring is $F_D = -F$, and

$$dW = -F_D dx\,. \tag{5.6}$$

This work is performed *against* the elastic force of the spring. The elastic force is given by the linear force law (Hooke's law), (Fig. 5.6)

$$F_D = -Dx\,. \tag{4.15}$$

Inserting Eq. (4.15) into Eq. (5.6), we obtain

$$dW = Dx dx. \tag{5.7}$$

Along the displacement x, the path integral of the force is the *triangular region COB*, whose area is $\frac{1}{2}xDx$. Thus we find

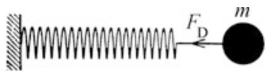
$$\text{Elastic work} = \tfrac{1}{2}Dx^2\,. \tag{5.8}$$

Numerical example

A bow is pulled back with a muscular force of $F = Dx = 200\,\text{N}$ through a distance of 0.4 m. The muscular force has to perform work amounting to $0.5 \cdot 200\,\text{N} \cdot 0.4\,\text{m} = 40\,\text{N m}$.

3. *Work of acceleration*. Fig. 5.7 is a continuation of Fig. 5.5: The object has just been released by the hand; the spring relaxes back

Figure 5.7 The definition of work of acceleration



to its resting position by pulling together. It thereby accelerates the object, which was previously at rest, to the left, and the elastic force F_D of the spring performs *work of acceleration*:

$$dW = F_D dx\,. \tag{5.9}$$

According to the equation of motion, Eq. (3.3), we have

$$F_D = m\frac{du}{dt} \tag{5.10}$$

and, from the definition of the velocity,

$$dx = u dt\,. \tag{5.11}$$

Equations (5.9) through (5.11) together yield

$$dW = mu du\,. \tag{5.12}$$

The integration (analogous to Fig. 5.6) gives the

$$\text{Work of acceleration} = \tfrac{1}{2}mu^2\,. \tag{5.13}$$

Numerical examples
A railroad train (locomotive + 8 cars), mass = $5.1 \cdot 10^5$ kg, velocity = 20 m/s, work of acceleration $\approx 10^8$ N m $\approx$ 28 kWh; a bullet from a pistol (cf. Sect. 2.2), mass = 3.26 g, velocity = 225 m/s, work of acceleration $\approx$ 82 N m.

The quotient work/time[C5.5] or – equivalently – the product of force times velocity are called the *power P*. The *unit* of power is

$$1\ \text{watt (W)} = 1\ \text{N m/s}\,. \tag{5.14}$$

C5.5. More precisely, the differential quotient $dW/dt = \dot{W}$.

When working continuously for several hours (turning a crank, walking on a treadmill etc.), a person can deliver a steady power output of around 0.1 kilowatt (kW). For a few seconds, the power output of a human can be greater than 1 kW; for example, a person can jump up a 6 m high staircase in 3 s. The power required is $70\,\text{kg} \cdot 9.81\,\text{m/s}^2 \cdot 6\,\text{m}/3\,\text{s} \approx 1.4\,\text{kW}$. More will be said about this topic in Sect. 5.11.

5.3 Energy and Its Conservation

In Sect. 5.2, we defined the integral $\int F dx$ and called it 'work'. We calculated the work for three different cases and gave numerical examples.

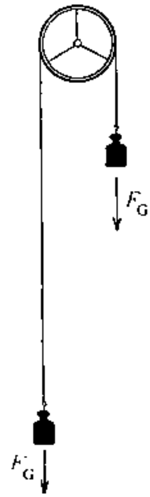


Figure 5.8 A body which has been raised up can perform work: it could lift another body of the same mass to the same height, without accelerating it.

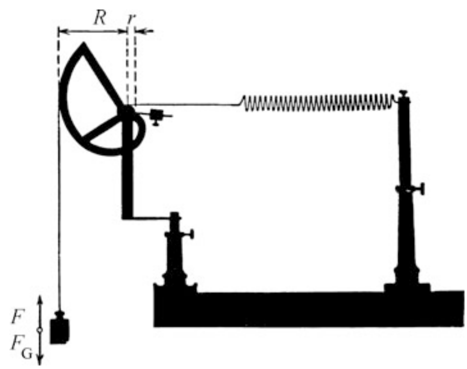


Figure 5.9 A spring which has been compressed or extended can lift an object and thereby perform lifting work, i.e. without work of acceleration. A continuously variable leverage keeps the lifting force F in equilibrium with the weight F_G at each moment; r is here the constant lever arm, while R is the lever arm which varies during the rotation.

In each of these three cases, the *ability to perform work* was increased; or, expressed differently, work was *converted* into the *ability* to perform work: A body lifted higher or a compressed spring can themselves perform work. They can for example lift an object (Figs. 5.8 and 5.9) or accelerate it (e.g. Fig. 5.7). This work which has been converted into 'the ability to perform work' is called

$$\left.\begin{array}{l}\text{lifting work } mgh\\ \text{elastic work } \frac{1}{2}Dx^2\end{array}\right\}\left\{\begin{array}{l}\text{the potential}\\ \textbf{energy } E_{\text{pot}}\end{array}\right\}\begin{array}{ll}\text{of the lifted body} & (5.4)\\ \text{of the spring.} & (5.8)\end{array}$$

Similarly, an object which has been accelerated has not only a velocity, but also the ability to perform work; it can for example deform another body in a collision and thereby perform elastic work. In this case, the work of acceleration which has been converted into the ability to perform work is called:

work of acceleration $\frac{1}{2}mu^2$ {the kinetic **energy** E_{kin}} of the accelerated body. (5.13)

In the above examples, the sum of the two forms of energy (potential and kinetic) is a constant quantity; thus

$$E_{\text{pot}} + E_{\text{kin}} = \text{const.} \tag{5.15}$$

This is a fundamental law of mechanics, the *conservation of energy* (briefly: *energy conservation*).

Explanation: In Fig. 5.7, assume that the spring relaxes through a displacement of $\mathrm{d}x$. Its elastic force F_{D} performs the work $\mathrm{d}W$ in the process. This work can be described in two ways: First of all, as work of acceleration which *increases* the kinetic energy E_{kin} of the object, thus

$$\mathrm{d}W = +\mathrm{d}E_{\text{kin}}\,. \tag{5.16}$$

Secondly, it can be considered to *reduce* the potential energy stored in the spring, so that

$$\mathrm{d}W = -\mathrm{d}E_{\text{pot}}\,. \tag{5.17}$$

Equations. (5.16) and (5.17) together yield

$$\mathrm{d}E_{\text{pot}} + \mathrm{d}E_{\text{kin}} = 0$$

or

$$E_{\text{pot}} + E_{\text{kin}} = \text{const}\,. \tag{5.15}$$

Similarly, in the case of the free fall of a body: Its weight F_{G} performs work of acceleration $\mathrm{d}W$ along the displacement $\mathrm{d}h$[C5.6], $\mathrm{d}W = +F_{\text{G}}\mathrm{d}h$. This is $= +\mathrm{d}E_{\text{kin}}$ and $= -\mathrm{d}E_{\text{pot}}$. Thus, here also, $\mathrm{d}E_{\text{pot}} + \mathrm{d}E_{\text{kin}} = 0$ and $E_{\text{pot}} + E_{\text{kin}} = \text{const}$.

C5.6. $\mathrm{d}h$ is directed *downwards* in this case.

We have thus far treated energy conservation in mechanics for only two types of forces, for elastic forces and for weight (the force of gravity). Such forces are called *conservative*. In their case, energy is "conserved". Frictional and muscular forces are *non-conservative*. For non-conservative forces, the mechanical energy conservation law, i.e. Eq. (5.15), does not apply. They will be included later by means of a brilliant extension of the energy conservation law.

Non-conservative forces: **"They will be included later by means of a brilliant extension of the energy conservation law".**

5.4 First Applications of the Conservation Law of Mechanical Energy

1. *Sinusoidal oscillations* (Sect. 4.3) consist of a periodic conversion of the two mechanical energy forms into each other. For every displacement x, it holds that

$$\tfrac{1}{2}Dx^2 + \tfrac{1}{2}mu^2 = \text{const}\,. \tag{5.18}$$

During passage through the rest position, all of the oscillation energy has been converted to kinetic energy; here we have

$$\tfrac{1}{2}mu_0^2 = \text{const} = E_{\text{kin}}\,. \tag{5.19}$$

At the points of reversal (maximum displacement x_0, or *amplitude*), all of the energy is potential, and we find

$$\tfrac{1}{2}Dx_0^2 = \text{const} = E_{\text{pot}}\,. \tag{5.20}$$

In words: *The energy of a sinusoidal oscillation is proportional to the square of its amplitude* x_0.

Setting Eq. (5.19) equal to Eq. (5.20), together with Eq. (4.5), leads to an important equation, which we have already encountered:

$$u_0 = \omega x_0\,. \tag{4.12}$$

2. *Oscillations whose frequency is strongly dependent on their amplitudes*. During free fall, the weight $F_G = mg$ of an object performs work of acceleration, $\frac{1}{2}mu^2 = F_G h = mgh$. Thus, the object's final velocity after falling from the perpendicular height h is

$$u = \sqrt{2gh}\,. \tag{5.21}$$

With the corresponding kinetic energy, the object may elastically deform itself and/or the surface on which it impacts (cf. Fig. 5.10) and thereby convert its kinetic energy into potential energy. This will be converted back to kinetic energy by relaxation of the deformed bodies. The object again rises, gaining potential energy, and so forth. This is the origin of the motion of the *dancing ball*: A good example of an oscillatory phenomenon with a *non-linear* force law. The deflection does not depend sinusoidally on time; its frequency increases strongly with decreasing amplitude (as in 'wobble oscillations', which are related to the dancing ball; see Sect. 11.16).

3. The definition of *elastic*. Deformations are called elastic when they fulfill the mechanical energy conservation law. Practically, this can be realized only as a limiting case. A fraction of the apparent mechanical energy is continually converted into the energy of molecular motions, i.e. into *heat*. The dancing ball never quite returns to its previous height.

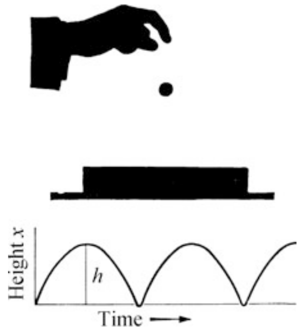


Figure 5.10 The conservation of energy. A steel ball is dancing on a steel plate. The latter could also be a glass plate covered with soot; then one could readily discern the flattening of the ball on impact. Below: The time evolution of the dancing ball's motion. **(Video 5.1)**

Video 5.1:
"Dancing steel ball"
http://tiny.cc/3qqujy
The steel ball is released magnetically, so that it falls straight down without rotation.

5.5 Impulse and Momentum

The *path* integral of the force, that is the work $\int \boldsymbol{F}\,\mathrm{d}\boldsymbol{x}$, led us to a fundamentally important concept, namely *energy*. The corresponding *time* integral $\int \boldsymbol{F}\,\mathrm{d}t$ of the force does the same[C5.7]: It is called the *impulse* and leads to the concept of *momentum*.

C5.7. In contrast to the path integral of the force, $W = \int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{x}$, the *time* integral of the force, $\int \boldsymbol{F}\,\mathrm{d}t$, and thus the *momentum*, is a *vector*.

Many motions occur with jerks or impacts; forces which change quickly in magnitude and direction are at work. Figure 5.11 is intended to illustrate the time variation of such a force. – Starting with processes of this kind, one arrives at the concept of 'impulse':

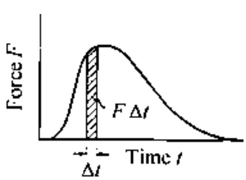
$$\text{Impulse} = \int \boldsymbol{F}\,\mathrm{d}t\,. \tag{5.22}$$

The unit of impulse[2] is for example the newton · second (N s).

Work performed upon an object increases its energy. What is the result of an impulse? The answer is provided by applying the equation of motion. Before the impulse, let the object have a velocity $\boldsymbol{u}_1$. During a time interval $\mathrm{d}t_\mathrm{i}$, the acceleration is given by $\boldsymbol{a}_\mathrm{i} = \boldsymbol{F}_\mathrm{i}/m$. Within the time interval $\mathrm{d}t_\mathrm{i}$, it gives rise to an increase in the velocity of

$$\mathrm{d}\boldsymbol{u}_\mathrm{i} = \boldsymbol{a}_\mathrm{i}\mathrm{d}t_\mathrm{i} = \frac{1}{m}\boldsymbol{F}_\mathrm{i}\,\mathrm{d}t_\mathrm{i} \tag{5.23}$$

Figure 5.11 The time integral of the force, or 'impulse'



[2] In electrodynamics, the corresponding quantity is the current impulse $\int I\mathrm{d}t$, measured in ampere · second (A s), and the voltage impulse $\int U\mathrm{d}t$, measured in volt · second (V s).

or

$$m\mathrm{d}\boldsymbol{u}_\mathrm{i} = \boldsymbol{F}_\mathrm{i}\,\mathrm{d}t_\mathrm{i};$$

and, after integration over time,

$$m(\boldsymbol{u}_2 - \boldsymbol{u}_1) = \int \boldsymbol{F}\,\mathrm{d}t\,. \tag{5.24}$$

"The product of mass times velocity was called by Newton the *quantity of motion*. In more recent times, this useful term has been supplanted by the word *momentum*, and we too must adopt this usage".

The product of mass times velocity, that is $m\boldsymbol{u}$, was called by Newton the *quantity of motion*. In recent times, this useful term has been supplanted by the word *momentum*, and we too must adopt this usage. Then, in words, Eq. (5.24) means: *An impulse* $\int \boldsymbol{F}\,\mathrm{d}t$ *changes the momentum of a body from its initial value* $m\boldsymbol{u}_1$ *to a final value* of $m\boldsymbol{u}_2$.

The symbol used for the momentum in most textbooks is $\boldsymbol{p}$:

$$\text{Momentum}\,\boldsymbol{p} = m \cdot \boldsymbol{u}\,. \tag{5.25}$$

Then Eq. (5.24) becomes

$$\Delta\boldsymbol{p} = \boldsymbol{p}_2 - \boldsymbol{p}_1 = \int \boldsymbol{F}\,\mathrm{d}t\,. \tag{5.26}$$

5.6 Momentum Conservation

The definitions given in Sect. 5.5 can be combined with the empirical rule "*actio* = *reactio*": Forces always occur in pairs; they act on bodies with the same magnitudes but in opposite directions. Figure 5.12 shows the simplest example: A compressed spring is placed between two small carts of masses M and m which are at rest. The overall momentum of this "system" is zero. A trigger mechanism then releases the spring. The two carts are each acted on by an impulse of the same magnitude but in opposite directions. As a result, each receives a momentum, also equal but opposite. As a formula, we can represent this by:

$$M\boldsymbol{u}_1 = -m\boldsymbol{u}_2; \quad M\boldsymbol{u}_1 + m\boldsymbol{u}_2 = 0\,. \tag{5.27}$$

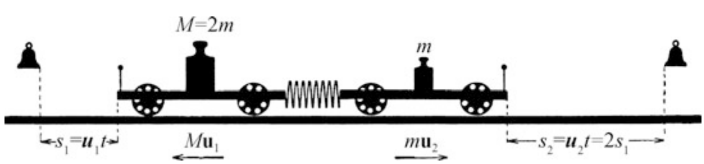


Figure 5.12 The conservation of momentum. Two carts with the masses $2m$ and m travel distances which are in the ratio 1:2 in a given time. Therefore, their velocities are in the same ratio, 1:2.

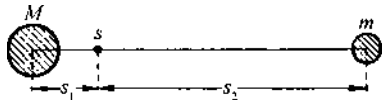


Figure 5.13 The definition of the center of gravity S

The sum of the two momenta remains zero. That is, in a descriptive generalization: *Without the effects of "external" forces, the sum of all the momenta in any system of arbitrarily moving bodies remains constant.* This is the law of conservation of momentum. This *momentum conservation* law is no less important than the law of conservation of energy.

"This momentum conservation law is no less important than the law of energy conservation."

Momentum conservation is often called the "law of conservation of the center of gravity". The reason for this can be seen in Fig. 5.12. The distances travelled during a given time interval obey $Ms_1 = ms_2$.
The same equation defines the center of gravity of two bodies at rest (Fig. 5.13).[C5.8]

C5.8. The center of gravity is defined here using momentum conservation, that is by means of the *inertial* mass. Its name however is derived from the more frequently applied definition based on the *gravitational* mass: it then corresponds to the point at which a body can be hung so that the sum of all the torques resulting from gravity is zero (i.e. the body hangs at rest without rotating; see Sect. 6.2.)

5.7 First Applications of Momentum Conservation

As with energy conservation, we illustrate momentum conservation by means of a few simple examples.

1. Let us start with a flat cart, about 2 m long. A man is standing on its right end (Fig. 5.14). The cart and the man form a system. The man begins to walk towards the left; he thus gains a momentum directed to the left. At the same time, the cart rolls to the right. According to momentum conservation, it has gained a momentum of the same magnitude but opposite direction to that of the man. – The man continues walking and walks off the left end of the cart. He takes his momentum with him. The cart now rolls with a constant velocity to the right, since it has the same momentum as the man, apart from its opposite sign.

2. To verify this quantitative statement, we let the empty, rolling cart meet up with a second man who is walking to the left (Fig. 5.15). The mass and velocity of this second man are chosen to be the same as

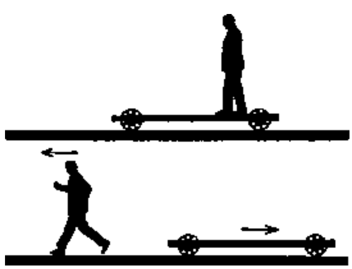


Figure 5.14 Momentum conservation. A man accelerates himself to the left on a cart and imparts a momentum in the opposite direction to the cart. **(Video 5.2)**

Video 5.2:
"Conservation of linear momentum"
http://tiny.cc/tqqujy

Part I

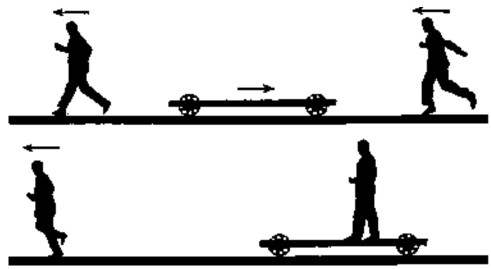


Figure 5.15 The momentum of the cart in Fig. 5.14 is equal to the momentum of the man

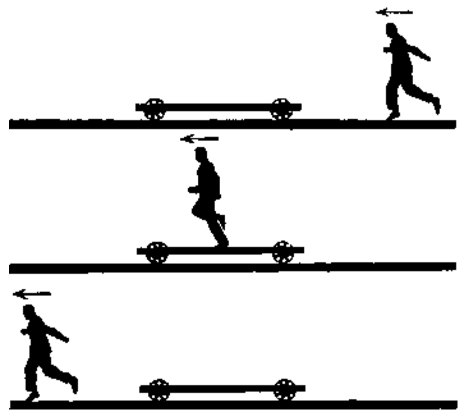


Figure 5.16 Momentum conservation. The walking man has kept the same momentum during and after walking over the cart. **(Video 5.2)**

Video 5.2: "Conservation of linear momentum"
http://tiny.cc/tqqujy

those of the first man. The second man steps onto the cart and stops walking. Immediately, the cart also stops rolling. The momentum transferred from the man to the cart was of the same magnitude but opposite direction to the momentum of the rolling cart.

3. The flat cart is standing at rest. From the right, a man comes walking rapidly with a constant velocity. He steps onto the cart on the right and leaves it on the left (Fig. 5.16). The cart remains standing at rest. The man brought his total momentum along with him and did not change it noticeably while he was on the cart. As a result, the momentum of the cart cannot change relative to its initial value of zero.

4. The flat cart has rubber tires. At right angles to its long axis, it resists being shoved. However, it can easily roll in the direction of its long axis. Therefore, it can be used to demonstrate the vector nature of the momentum $\boldsymbol{p}$:[C5.9]

Assume the man to be walking at an angle α towards the long axis of the cart when he steps onto it. Then only the component $p \cos \alpha$ of his momentum is along the cart's long axis. For $\alpha = 60°$, the cart reacts with only half the velocity ($\cos 60° = 0.5$); for $\alpha = 90°$, the cart remains at rest ($\cos 90° = 0$).

C5.9. These impressive demonstrations of momentum conservation have the advantage compared to experiments on an airtrack, which are frequently used today, of being able to demonstrate the *vector nature* of momentum.

5.8 Momentum and Energy Conservation During Elastic Collisions of Objects

Figure 5.17 shows two loaded carts with spring bumpers. They each have the same mass. The left-hand cart is moving towards the other with the velocity u. When they collide, they exchange their velocities; the right-hand cart moves away with the velocity u, while the left-hand cart is now at rest. It stops at the moment when the spring bumpers are completely relaxed. – to explain this experiment, one requires both momentum conservation *and* energy conservation. We will illustrate this using an example where the masses are not equal.

Momentum conservation requires

$$\begin{array}{ccccc} \text{Left-hand cart} & & \text{Left-hand cart} & & \text{Right-hand cart} \\ mu & = & mu_1 & + & Mu_2 \\ \text{before the collision} & & & \text{after the collision} & \end{array}$$

or

$$m(u - u_1) = Mu_2 \,. \tag{5.28}$$

Energy conservation requires

$$\tfrac{1}{2}mu^2 = \tfrac{1}{2}mu_1^2 + \tfrac{1}{2}Mu_2^2$$

or

$$m(u + u_1)(u - u_1) = Mu_2^2 \,. \tag{5.29}$$

Equations (5.28) and (5.29) together yield

$$u_2 = (u + u_1) \,. \tag{5.30}$$

Figure 5.17 Demonstration of a slow collision. The helical springs F are attached firmly to the cart at a and to the bumper at b. When the two bumpers meet, the springs will be extended; an elastic force to the left thus acts on the left-hand cart, and an equal force to the right acts on the right-hand cart. These two forces have the same magnitudes at every moment during the extension of the springs. They brake the motion of the left-hand cart and accelerate the right-hand cart. **(Video 5.3)**

Video 5.3:
"Elastic collisions in slow motion"
http://tiny.cc/xqqujy
The major advantage of this arrangement is that the collision takes place *slowly*. Forces and energy transfers can be clearly observed. The collision process is made slower by increasing the masses of the carts by adding the experimenters, while keeping the overall mass on each side equal.

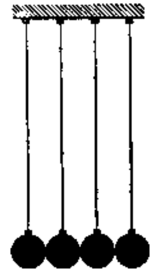


Figure 5.18 The demonstration of elastic collisions between two objects of the same mass. The balls are hung by two threads ("bifilar suspension"), so that they can move freely from left to right (but not forward or backward). **(Video 5.4)**

Video 5.4:
"Elastic collisions"
http://tiny.cc/xrqujy
Collisions between balls having the same and different masses are shown.
(**Exercise 5.5**)

Using Eq. (5.30), one can eliminate either u_1 or u_2 from Eq. (5.28). The resulting velocity is

$$\text{for the colliding object (of mass } m) \qquad u_1 = u\,\frac{m-M}{M+m}\,, \tag{5.31}$$

$$\text{for the target object (of mass } M) \qquad u_2 = u\,\frac{2m}{M+m}\,. \tag{5.32}$$

In the special case that $M = m$, it thus follows for the velocity after the collision that u_1 (colliding cart) $= 0$; u_2 (target cart) $= u$. For $M > m$, u_1 becomes negative, i.e. it is directed oppositely to the velocity u (the cart is 'reflected').

The demonstration sketched in Fig. 5.17 can be carried out with a larger number of carts. One can occasionally see this in a railroad switchyard. In the lecture hall, instead of carts one generally uses a series of steel balls which are hung as simple pendulums; cf. Fig. 5.18. The leftmost ball is pulled out and released, so that it collides with its neighboring ball. That ball receives the momentum and takes on the role of colliding ball (projectile), while its next neighbor becomes the target, and this proceeds in sequence through the whole line of balls within a very short time. The rightmost ball then flies up and out to the right with nearly the same velocity as the original leftmost ball just before the first collision.

5.9 Momentum Conservation in Inelastic Collisions and the Ballistic Pendulum

When both bodies are moving with the common velocity u_2 after a collision, one speaks of a (completely) *inelastic* collision. It can be demonstrated by replacing the elastic bumpers in Fig. 5.17 with lead or some other easily deformable material, so that the two bodies "stick together". In this case, the mechanical law of energy conservation does *not* hold. But momentum conservation alone allows us to

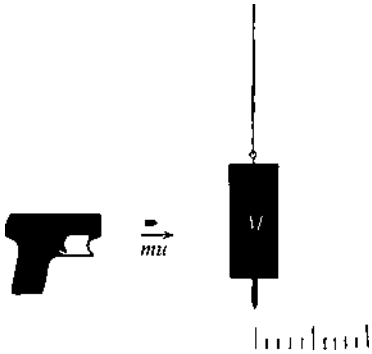


Figure 5.19 A ballistic pendulum as a force meter. Measuring the velocity of a bullet (suspension by two threads of length ca. 4.2 m, natural period of oscillation $T = 4.1$ s). **(Video 5.5)**

Video 5.5:
"Measuring the velocity of a bullet"
http://tiny.cc/qrqujy
For safety, the pendulum swings along two light metal rails. Its period is 2 s, and its mass is 2.2 kg. The mass of the bullet is 2.6 g. The pendulum exhibits an amplitude of around 12 cm (6 marks on the scale; compare the silhouette). From these data, we find the bullet's velocity to be 320 m/s.

make a prediction (since there is only one common velocity after the collision). Momentum conservation requires

$$\begin{array}{ccc} \text{Left-hand cart} & & \text{Both carts together} \\ mu & = & u_2(m+M) \\ \text{before the collision} & & \text{after the collision} \end{array} \tag{5.33}$$

or

$$u_2 = u\,\frac{m}{M+m}\,. \tag{5.34}$$

The velocity u_2 of the target body after the collision is thus only half as great as in an elastic collision for equal masses; cf. Eq. (5.32). This means that twice as much momentum is transferred in an elastic collision as in a completely inelastic collision.

An example of an application is the measurement of the muzzle velocity of the bullet from a pistol. In Fig. 5.19, we see the bullet impacting with its velocity u onto a block of mass M, where it lodges. Bullet and block then move together to the right with their common velocity u_2. This velocity is measured, and using it, the muzzle velocity u can be computed from Eq. (5.34).

The measurement of u_2 can be carried out with a simple stopwatch. For that purpose, we construct a *ballistic pendulum*. This means that we mount the block M in such a way that it can oscillate, e.g. between two springs or as the mass of a gravity pendulum (Fig. 5.19). In either case, we guarantee a linear restoring force law: the springs or the cord of the pendulum are made sufficiently long. For a linear force law, the important equation

$$u_0 = \omega x_0\,. \tag{4.12}$$

holds. In words: the velocity u_0, here u_2, with which a previously resting body leaves its rest position in an oscillator system is obtained in a simple manner from its amplitude x_0: One need only multiply the latter by the circular frequency $\omega = 2\pi/T$ (T is the period of oscillation of the pendulum).

What an enormous simplification we have achieved by using the concept of momentum! Before (in Sect. 2.2), we had to use a chart recorder with time marks, an electric motor, a rheostat, a tachometer, and a backstop. With momentum conservation at our disposal, for the same measurement we need only a cigar box full of sand, some twine, a balance and a stopwatch.

Beginners sometimes try to apply energy conservation for measuring the velocity of a bullet with the ballistic pendulum. They set the kinetic energy $\frac{1}{2}mu^2$ of the bullet equal to the kinetic energy $\frac{1}{2}M(\omega x_0)^2$ of the pendulum. This, however, is not permissible; the impact of the bullet is *not* elastic (Sect. 5.4). On the contrary, the kinetic energy of the bullet is converted on impact almost completely into heat; only about 0.16 % remains as kinetic energy.

5.10 Non-Central Collisions

If two bodies approach each other along a path which is *not* the line connecting their centers of gravity, then instead of a "central" collision, there will be a "non-central" one. It will produce angular deviations between 0° and 180°. This is demonstrated and discussed in **Video 5.6**: "Non-central collisions". On the glass plate of an overhead projector, which has been dusted with fine lycopodium powder, two round steel disks can glide nearly without friction. Their tracks in the powder are readily visible. The masses of the disks can be varied, as well as the velocity of the projectile (in magnitude and direction). Thus, with this model we can demonstrate collision phenomena, which play such an important role in atomic, molecular and nuclear physics.

Video 5.6:
"Non-central collisions"
http://tiny.cc/jqqujy
(See also **Exercise 5.8**).

5.11 Motions Against Dissipative Forces

Inelastic collisions are an exception among the motions we have otherwise considered, in that they exhibit *in principle* a "loss" of mechanical energy. "Energy loss" or "dissipation" describes the conversion of energy into heat, or more correctly, into internal energy. In all the other motions, such losses were practically negligible side effects; they were reduced to a minimum by careful design of the experiments, and were completely neglected in the computations and theoretical treatment of the measurements.

However, many motions with continual and unavoidable energy losses play an important role in mechanics. As a first example, Fig. 5.20 shows the velocity-time curve of a person falling from a great height. Initially, the motion is *accelerated*; after one second, the velocity has reached a value of nearly 9.8 m/s. Rather soon, however, it begins to increase much more slowly than in a vacuum, and finally it reaches a *constant value* of $u \approx 55$ m/s. – Explanation:

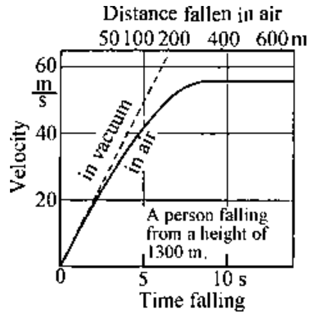


Figure 5.20 The influence of air resistance on a falling person. In order to maintain the constant terminal velocity of around 60 m/s for a person with a mass of 70 kg, his/her falling mass has to generate a power of about 40 kW(!), provided by its potential energy.

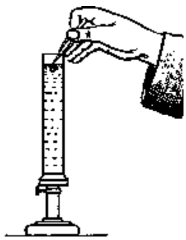


Figure 5.21 The constant sinking velocity of balls in viscous fluids

During the acceleration, a force F_2 acts in the direction opposing the motion. It increases with increasing velocity, until its maximum value $F_2 = -F_G$ is reached, i.e. it is equal to the force of gravity F_G, but opposite to it. Then the sum of the forces acting on the person, $F_2 + F_G$, is zero. As a result, no further acceleration occurs, and the velocity of the falling person has reached its constant *terminal or saturation value*: the person is no longer freely *falling*, but rather *sinking* with the constant *terminal velocity u*.

The essential point here, the increase of the velocity of falling up to a constant terminal velocity u, can be readily shown in a demonstration experiment, e.g. by letting small steel balls fall in a viscous fluid (Fig. 5.21). Details are given in Sect. 10.3.

As a further example, we can consider our various means of transport: automobiles, railroads, ships and aircraft. Even on a horizontal path, a vehicle requires a driving force not only to *accelerate*, but also to maintain a *constant* velocity! – In Fig. 5.22, we see a cart carrying about 50 kg of mass on the horizontal floor of the lecture hall. It is being pulled via a cord with a force $F_1 = 9.81$ N. After it has moved about 1 m, its velocity reaches a *constant* value of ca. 0.5 m/s, and its acceleration becomes zero. Therefore, there must be a second force acting during its acceleration, which is directed oppositely to its motion and which increases with velocity $\mathrm{d}x/\mathrm{d}t$, that is a force with a maximum value of $F_2 = -F_1$. – This force may be produced in a variety of ways, for example by friction in the bearings of the cart's wheels or by displacement and turbulence of the surrounding medium, usually air or water.

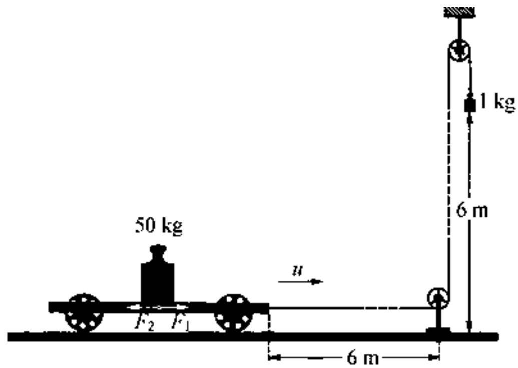


Figure 5.22 As the result of a "dissipative" or "energy consuming" force F_2, maintaining a constant velocity u requires a driving force F_1. The driving force F_1 performs work against the dissipative force (resistance) F_2. F_2 increases with increasing velocity.

This is the reason that we cannot define a generally valid relation between force and velocity. In the simplest case, for example that shown in Fig. 5.21, the resistance force increases proportionally to the momentary velocity. In the case of ships and aircraft, it increases to a rough approximation proportionally to the *squares* of their velocities.

Whatever produces the force F_2, the driving force F_1 must always perform work against the resistance force F_2 along the distance x travelled; this work is equal to $F_1 x$. The quotient (work W/time t) gives the corresponding *power* $\dot{W}$, i.e. $\dot{W} = F_1 x/t = F_1 u$. Therefore, some sort of motor or energy source must be available to maintain a constant 'cruising speed' u, and it must produce a power of

$$\dot{W} = F_1 u\,. \tag{5.35}$$

Examples

Automobile engines deliver between 10 and 200 kW, locomotives and aircraft engines usually several 10^3 kW, the engines of large ships up to more than 10^5 kW. The power required for the natural method of locomotion of humans, walking, is in comparison very modest. For a typical walking speed of 5 km/h = 1.4 m/s on a horizontal path, one requires around 60 W; a rapid walk at 7 km/h increases this to ca. 200 W. – The work performed in walking is composed of two parts: 1. A periodic raising of the body's center of gravity (walk along a wall holding a piece of chalk on the wall at hip height, and look at the resulting wavy line!); and 2. the work required to accelerate our legs. The inelastic impacts of our feet on the ground convert most of this energy into heat, which is thus 'lost' for further mechanical usage. – When riding a bicycle, the up-and-down motion of the center of gravity is minimized, and the work of moving the legs is also smaller. For a bicycle speed of 9 km/h, one consumes a power of about 30 W, and at 18 km/h, only about 120 W. – With numbers of this kind, we are in a better position to judge the power values quoted for various technical devices.

In the case of humans and work animals, locomotives and automobiles, the production of the driving force F_1 is possible due to sticking

Figure 5.23 Production of the driving force for water-craft (and aircraft)

friction[3] (Sect. 8.9). But how is the driving force for motorized air and water vehicles produced? Answer: The motor takes in some portion of the surrounding medium (water or air) with propellers, jets, paddle wheels or similar mechanisms and accelerates it backwards. This gives rise to a driving force F_1 in the forward direction (momentum conservation!).

As an example, Fig. 5.23 shows a boat moving at a constant speed u towards the right relative to the shoreline. The 'motor' is a man who uses a paddle to accelerate water backwards, giving it a velocity u' to the left. In a time t, he accelerates an amount of water with a mass M, thereby producing a momentum Mu' towards the left. At the same time, the man himself (and the boat in which he is sitting) receives an impulse to the right, i.e. in the direction of travel:

$$F_1 t = Mu' \,. \tag{5.36}$$

The force

$$F_1 = \frac{Mu'}{t} \tag{5.37}$$

is the driving force. It is required to maintain a constant speed of the boat against the resistance of the water. The force F_1 performs the work $W_1 = F_1 t\, u$ in a time t along the path $t\, u$, or, from Eq. (5.37),

$$W_1 = M\, u\, u' \,. \tag{5.38}$$

At the same time, the water which was accelerated to the left receives a kinetic energy $W_2 = \frac{1}{2}Mu'^2$. The motor has to deliver the sum of these two energies, W_1 and W_2, but only the fraction W_1 is available as useful work (for moving the boat)[C5.10]. Then we find for the propulsion *efficiency*

$$\eta = \frac{W_1}{W_1 + W_2} = \frac{Mu\,u'}{Mu\,u' + \frac{1}{2}Mu'^2} = \frac{1}{1 + \frac{1}{2}\frac{u'}{u}} \,. \tag{5.39}$$

C5.10. This computation is relevant not only to boatsmen and other water-sports enthusiasts, but also to the propulsion of all types of ships and aircraft. In most textbooks and lecture courses, this point is seldom discussed.

We can see that it is advantageous to reduce the velocity u' of the water which is accelerated backwards, in order to maximize the propulsion efficiency. Then, however, from Eq. (5.37), the mass M of the water accelerated must be large in order to obtain the required driving force F_1. With

[3] If we wish to apply momentum conservation to a walking person, we should consult Fig. 5.14 and imagine that the cart is replaced by the earth, with its enormous mass.

boats moved by propellers or paddle wheels, the water which is accelerated backwards can be clearly seen in the wake of the boat.

In the case of aircraft, the same considerations hold. In the early years of powered flight, the backwards-directed stream of air was produced only by propellers. Today, jet engines or turbojets of various designs are employed in many cases (Sect. 10.11). The latter of course eject not only the air they have taken in, but also combustion products from their burning fuel.

Rocket propulsion is not fundamentally different; however, none of the material which is accelerated backwards is taken from the surrounding medium, but rather it is part of the initial load (fuel and oxidant) of the vehicle. Even at a constant flight speed u, the load has momentum and energy as seen by an observer at rest. This has to be taken into account in a quantitative treatment. For the efficiency, one then finds

$$\eta = \frac{u(u + 2u')}{(u + u')^2} .$$

Space rockets, however, do not fly at constant velocities; they are accelerated. The forces of resistance are small in comparison to the forces required for acceleration. To compute the speed $\boldsymbol{u}$ of the rocket under these conditions, we start with momentum conservation (Eq. (5.27)). Then we have

$$M\boldsymbol{u} = (M + \Delta M)(\boldsymbol{u} + \Delta\boldsymbol{u}) - \Delta M(\boldsymbol{u} - \boldsymbol{u}_\mathrm{r}) .$$

At the left is the momentum at the time t, and on the right at the time $t + \Delta t$, when the rocket has ejected combustion products of mass ΔM (ΔM is negative, the mass of the load is decreasing) with the relative velocity $\boldsymbol{u}_\mathrm{r}$ (oppositely directed to $\boldsymbol{u}$). It follows that

$$(M + \Delta M)\Delta\boldsymbol{u} = \Delta M\boldsymbol{u}_\mathrm{r} ,$$

and for $\Delta t \to 0$

$$\mathrm{d}\boldsymbol{u} = \boldsymbol{u}_\mathrm{r}\frac{\mathrm{d}M}{M} .$$

Integration, beginning with the initial mass M_0 down to a mass M, then yields

$$\boldsymbol{u} = -\boldsymbol{u}_\mathrm{r} \ln \frac{M_0}{M} \tag{5.40}$$

for the final velocity $\boldsymbol{u}$ (after burning a mass $\Delta M = M_0 - M$ of fuel). This is called the *rocket equation*. At constant $\boldsymbol{u}_\mathrm{r}$, the final velocity $\boldsymbol{u}_\mathrm{f}$ is thus determined uniquely by the ratio M_0/M. $\boldsymbol{u}_\mathrm{f}$ is independent of the ejected mass current $\mathrm{d}M/\mathrm{d}t$; this is the surprisingly simple result!

5.12 The Production of Forces with and without Consuming Power

In Sect. 5.11, we have just considered the propulsion of vehicles along horizontal paths. The weight of the vehicle had to be supported in some manner by an *upwardly-directed* force. In the case of road and rail vehicles, this force arises from elastic deformations of the road surface or the rails; for ships and airships, it results from static buoyant forces (Sect. 9.4). For heavier-than-air craft, however, this upwardly-directed or lifting force must be produced aerodynamically using lifting surfaces or airfoils (see Sect. 10.10). This dynamic lift simply replaces the cables holding up the cabins in the case of an aerial tramway; its effect in the end is no different from that of a hook in the ceiling. However, a hook or a permanent magnet can produce a lifting force year in and year out without any external power source. This is in contrast to an airfoil, which requires a steady input of power. The production of a lifting force by an airfoil is thus fundamentally similar to the production of a force by an electromagnet or a muscle: An electromagnet draws power from its source of electric current, a muscle requires an input of chemical energy from the metabolism and tires even from "holding" (isometric force), i.e. even without performing work in the physical or technical sense. For *physical work requires not only a force, but also a displacement along which the force acts.* – All types of force which require input power have a common characteristic: They involve "losses" of mechanical, chemical or electrical energy; that is, some part of this energy is converted into heat (more precisely: into internal energy). Heating by electric currents and by muscular activity is well known. In the case of airfoils, heat is produced by various mechanisms; one of them is the formation of turbulence at the tips of the aircraft wings. Physicists are inclined to consider only those forces which perform work even in economic applications. That is of course not correct. Often enough, even the provision of forces which do not perform any work requires a considerable economic expenditure.

"Physicists are inclined to consider only those forces which perform work even in economic applications. That is of course not correct."

Example

A vehicle is required to use a very short towrope for towing another; otherwise, when the road is curvy, large force components perpendicular to the towing direction can occur. These do not perform work, but they use additional fuel and strain the suspensions of both vehicles.

5.13 Closing Remarks

The logical path of our considerations has taken us from the equation of motion (3.3) to the equation for momentum, (5.24). Of course, the reverse path would be equally valid (and it was indeed that followed by NEWTON)[C5.11]. We could start with the definition of the momentum, $m\boldsymbol{u}$, and require that *the time variation of the momentum is proportional to the net force acting*, or, mathematically,

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{u}) = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{F}\,. \tag{5.41}$$

C5.11. This reverse path, starting with momentum instead of forces, or, more generally, starting from the conservation laws instead of the equation of motion, is in fact seldom treated in physics courses. This is certainly due not only to didactic considerations, but may also be a result of the historically-determined preferences of the teachers.

For the limiting case of a constant mass m, one can then write for the acceleration

$$m\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \boldsymbol{F} \quad \text{or} \quad \boldsymbol{a} = \frac{\boldsymbol{F}}{m}; \tag{3.3}$$

and in the limit of a purely *radial* acceleration

$$m\boldsymbol{\omega}\times\boldsymbol{u} = \boldsymbol{F} \quad \text{or} \quad \boldsymbol{a}_{\mathrm{r}} = \frac{\boldsymbol{F}}{m} = \boldsymbol{\omega}\times\boldsymbol{u}\,. \tag{2.6}$$

For a *constant* mass, the two routes are equivalent. The path we took is better adapted to the needs of a course in *experimental* physics.

The assumption of a *constant* mass m is however only an *approximation*, albeit one that is justified within wide limits. Its applicability defines the boundaries of "classical mechanics". In general, the relativistic velocity dependence of the mass must be taken into account, and instead of m, one must write

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}} \,. \tag{5.42}$$

Here, m_0 is the mass at zero velocity (rest mass), and c is the velocity of light in vacuum, $c = 3 \cdot 10^8$ m/s. When this correction is taken into account, the momentum equation (5.41) remains valid, but not the equation of motion (3.3). In the region of extremely high velocities u, the oh-so-simple equation of motion reaches the limits of its validity.

Exercises

5.1 Two identical cylindrical containers whose bottoms (each with a surface area A) are at the same height, are filled with water (density ϱ) up to a height of h and H, respectively. What is the work performed by the force of gravity when the two containers are connected so that the water levels equalize? (Sect. 5.2)

5.2 An automobile of mass $m = 1$ t drives for one kilometer up an incline with a grade of 1:25. In addition to the force of gravity, it must overcome frictional forces; they amount to the equivalent of 1 % of the automobile's weight. How much work W does the engine perform? What is its power output, if the car is moving at a speed of 60 km/h? (Sect. 5.2)

5.3 An object of mass m slides down a frictionless inclined plane. It starts from a height h with a velocity of zero. At the height h_1, it receives an impulse $\Delta p = \int F \mathrm{d}t$ opposite to its direction of motion, which is so strong that it slides back up the incline. What height h_2 does it reach, and what is its velocity v when it finally arrives at the bottom of the inclined plane ($h = 0$)? (Sect. 5.5)

5.4 A stream of water with a mass current of 1.2 kg/min strikes a vertical plate horizontally. Calculate the velocity of the water stream if the force that it exerts on the plate is 0.2 N, for two cases: a) The water comes to a stillstand at the plate; and b) the water is deflected backwards from the plate with the same velocity with which it arrived. (Sect. 5.5)

5.5 In Video 5.4, "Elastic collisions", among other things collisions between a large ball (55 mm diameter) and a small ball (11 mm diameter) are shown. a) If the small ball (mass m) strikes the large ball (mass M, initially at rest) from the right with a velocity v_1, the large ball moves to the left with a velocity v_2, while the small ball moves to the right with a velocity $v_2 = v_1 - \Delta v$. Calculate v_2 and Δv. b) After the second collision, the large ball is again at rest. Try to make this plausible without any further calculations. c) If the large ball comes from the left with a velocity v_2 and strikes the small ball, which is initially at rest, the latter moves off to the right with a velocity v_1. Compare the two velocities by measuring the amplitudes of the swings, and check them against the calculation. (Sect. 5.8)

5.6 A ballistic pendulum consists of a sandbag of mass M hanging from a cord. It suffers an inelastic collision with a small object of mass m and velocity u. To what height will the sandbag be raised by the resulting pendulum swing? (Sect. 5.9)

5.7 A bullet of mass $m_1 = 30\,\text{g}$ is shot with a horizontal velocity of $u_1 = 300\,\text{m/s}$ into a large wooden block with a mass of $m_2 = 3\,\text{kg}$ which is suspended from thin cords. What is the velocity of the block after the impact (without rotation)? What fraction of the original kinetic energy of the bullet is converted to kinetic energy of the block plus bullet? (Sect. 5.9)

5.8 Show that in a non-central collision of two objects of the same mass, the angle between their paths following the collision is 90°, presuming that the collision is elastic and that rotations can be neglected. (Sect. 5.10)

5.9 A rocket takes off with a total mass of $M_o + M_p$. M_p is the mass of its propellant (fuel), which is ejected out of the nozzle of the rocket at a relative velocity of u_r when burned. Determine the mass M_p which would be necessary to accelerate the rocket to a velocity of u_r. (The acceleration of gravity can be neglected.) (Sect. 5.11)

5.10 A rocket is to be accelerated vertically from the ground with an acceleration of a_o. Its total mass on takeoff is M_t, and the combustion gases are ejected from its nozzle with a relative velocity u_r. Find the mass current $\mathrm{d}M/\mathrm{d}t$ which has to be ejected from the rocket in order to produce this acceleration on takeoff. (Sect. 5.11)

For Sect. 5.2, see also Exercise 4.4; for Sect. 5.3, see also Exercise 3.3.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_5) contains supplementary material, which is available to authorized users.

Rotational Motion of Rigid Bodies

6

Part I

6.1 Introductory Remarks

In general, for a body moving in an arbitrary way, we can distinguish two different forms of motion, namely *progressive* motions and *rotations*. Our treatment up to now has been limited to progressive motions. *Formally*, we have treated the moving bodies as point-like objects or point masses. *Experimentally*, we prevented rotational motions by using two tricks: In the case of motions along a straight-line path, we let the accelerating force act along a direction passing through the *center of gravity* (also called the "center of mass") of the body. For motions along a curved path, we chose all the dimensions of the body to be small compared to the radius of curvature of the orbit. Admittedly, for example a stone in a sling completes a full *rotation* around its center of gravity each time it circles before being released. But the kinetic energy of this rotational motion (Sect. 6.4) is small compared to that of its progressive motion around the circle. Therefore, we could neglect the rotational motion as compared to the progressive motion along the circular orbit. – In this chapter, we consider the other limiting case: A body is not moving progressively along a path; instead, its motion is limited to pure *rotations*. *The axis of rotation will initially be assumed to be fixed in space by bearings.*

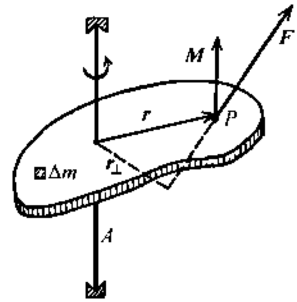
6.2 Definition of Torque

Figure 6.1 shows a flat, rigid object with a perpendicular axle *A* mounted on bearings. When the object is rotated around this axis, every small piece of it (its 'volume elements' or 'point masses'), each with a mass Δm, rotates within a plane perpendicular to the axis, called the *plane of rotation*. The object is assumed to be able to remain at rest in any arbitrary angular position around its axis of rotation. To this end, the influence of its weight has to be suppressed. We have to orient its axis of rotation exactly vertically; then the plane of rotation of each point mass in the object is horizontal.

Not just any arbitrary force is sufficient to produce a rotational motion. Instead, the force must generate an effective *torque* around the axis of rotation in question. This means that the force must contain

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_6

Figure 6.1 The definition of a torque $\boldsymbol{M}$, which is directed parallel to the axis of rotation



a component within the plane of rotation, and its direction must not pass through the axis of rotation.

Quantitatively, we define the torque $\boldsymbol{M}$ simply for a force $\boldsymbol{F}$ that is parallel to the plane of rotation by means of the equation[C6.1]:

C6.1. Equation (6.1) contains a *vector product*, whose definition we give briefly here: The result $\boldsymbol{c}$ of the vector product $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{c}$ is itself a vector which is perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$. $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ form a right-hand coordinate system; that is, when one looks along the direction of $\boldsymbol{c}$, a rotation to the right (clockwise) leads from the direction of $\boldsymbol{a}$ to the direction of $\boldsymbol{b}$. The magnitude of $\boldsymbol{c}$ is given by $c = ab \sin\varphi$, where φ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$. The vector product is not commutative; instead: $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$. (For further details, see mathematics textbooks.)

$$\boldsymbol{M} = \boldsymbol{r} \times \boldsymbol{F} \,. \tag{6.1}$$

Here, $\boldsymbol{r}$ is the position vector which points from the axis of rotation to the point where the force acts. The magnitude of $\boldsymbol{M}$ corresponds to the product $r_\perp F$, where $r_\perp$ is the perpendicular (i.e. shortest) distance from the line along which the force acts to the axis of rotation, also called the "lever arm" of the force. The unit of the torque is the newton meter (N m).

Let a torque $\boldsymbol{M}$ rotate a body (Fig. 6.1) through the angle $\mathrm{d}\beta$. Then it performs the work $\mathrm{d}W = F\mathrm{d}x = Fr_\perp \mathrm{d}\beta = M\mathrm{d}\beta$, or $M = \mathrm{d}W/\mathrm{d}\beta$. Its unit is therefore the unit of work/unit of angle, e.g. newton meter/rad. The unit 'rad' is equivalent to the number 1, and it is therefore often left off. For this reason, a torque has the same units as work.

The torque $\boldsymbol{M}$ is also a vector. It is perpendicular to both the force $\boldsymbol{F}$ and to the lever arm $\boldsymbol{r}$. Thus, in Fig. 6.1, it points parallel to the axis of rotation. Looking along the direction of the torque, we see the resulting rotation as clockwise.

Torques can also be produced by other forces which are not parallel to the plane of rotation. The direction of the resulting torque is then no longer parallel to the axis of rotation (see Eq. (6.1)). In this case, only the component of the torque parallel to the axis of rotation is effective in producing a rotation around that axis.

Usually, a number of different forces act on a body which is free to rotate, producing a variety of different torques. These combine vectorially to give a resultant overall torque. This is the case for example for an *electric motor*. In order to measure the total torque exerted by the motor even during its operation, we can make use of the equal but oppositely directed *reaction torque* which acts on the housing of the motor. Details are given in Fig. 6.2.

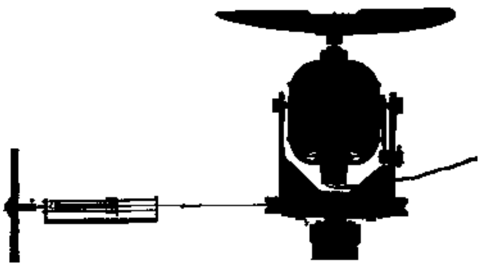
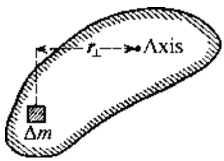


Figure 6.2 Measurement of the torque of an electric motor, *while* it is producing a vertical air current with its output power of $\dot{W} \approx 0.5\,\mathrm{kW}$. The force meter (cf. Fig. 3.9) is attached by a cord tangentially to the circumference of a round table ($2r = 0.25\,\mathrm{m}$), which is mounted on ball bearings so that it is free to rotate around a vertical axis. The mounting of the rotatable table is the same as that shown in Fig. 6.17. The product of the torque M and the angular velocity ω (Sect. 2.5) of the motor gives the power $\dot{W}$, e.g. in N m/s or watt.

Figure 6.3 The center of gravity



In Fig. 6.1, the axis of rotation was chosen to be vertical. In this limiting case, the weight of the object or of its individual point masses Δm cannot give rise to any torques *parallel to the axis*, which would thus be able to produce a rotation around it. In the second limiting case, that of a *horizontal* axis, the situation is different. Here, each individual point mass Δm contributes a torque as shown in Fig. 6.3, proportional to $r_\perp \Delta m$. In general, if the object is initially in some arbitrary position, it will begin to rotate; only in a special case will it stay at rest in any initial position. This special case arises when the axis of rotation passes through the center of gravity of the object. Thus, for an axis through the center of gravity, the resulting total torque and therefore also the sum $\sum r_\perp \Delta m$ must be equal to zero. This equation comprises a definition of the center of gravity.[C6.2] We will make use of it later. Otherwise, we will consider the center of gravity of a body and its determination to be well known, as we have done up to now. The center of gravity is usually treated in detail in school physics courses in connection with levers, balances and simple machines.

C6.2. As a general definition of the center of gravity, we find for the resulting torque $\boldsymbol{M} = \int \boldsymbol{r} \times \boldsymbol{g}\, \mathrm{d}m = 0$ ($\boldsymbol{g}$ is the vector of the acceleration of gravity).

If an axis of rotation is mounted on *fixed bearings*, the direction, magnitude and sense of rotation of a torque will be generally clear. In other cases (free axes), the beginner occasionally encounters difficulties. Among these for example is the child's trick of "obedient" and "disobedient" spools of yarn: A spool of yarn has fallen to the floor and rolled under the sofa. Someone is trying to retrieve it by pulling on the end of the yarn. Some spools come out obediently, others hide further back in their lairs. Figure 6.4 shows the explanation. One must consider not the symmetry axis of the spool,

Part I

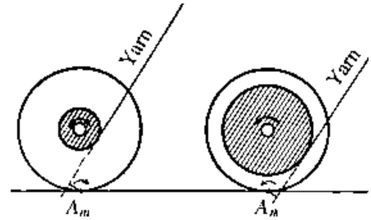


Figure 6.4 Torques acting on spools of yarn

but rather its line of contact with the floor as its axis of rotation. This is indicated in Fig. 6.4 as A_m (the "momentary axis" or "instantaneous axis"). If the yarn is held at a sufficiently "flat" angle to the floor, even the most obstinate spool can be forced to obey. As is often true in life, a little physics is of more use here than petulant outbursts of temper.

"As is often true in life, a little physics is of more use here than petulant outbursts of temper."

6.3 The Production of Known Torques, the Constant D^*, and the Angular Velocity $\boldsymbol{\omega}$

Forces of known magnitude and direction can be produced in an intuitively clear manner by using *helical springs* (Fig. 5.5). If the spring is properly dimensioned (sufficiently long), the elastic force that it produces is proportional to the displacement x (extension or compression) of the length of the spring. A linear force law applies (Hooke's law):

$$\boldsymbol{F} = -D\boldsymbol{x}\,. \tag{4.15}$$

The quotient of the magnitudes

$$\frac{\text{Force } F}{\text{Displacement } x} = D$$

is called the spring constant or *elastic coefficient* of the spring.

Analogously, *torques M* of known magnitude and direction can be produced in an intuitively clear manner by using a *spiral spring* attached to an axle. Figure 6.5 shows such a *torsion shaft*. When properly dimensioned (sufficient length of the spring), it produces a torque proportional to the angle of rotation. We again find a linear relation

$$M = -D^*\beta. \tag{6.2}$$

The quotient of the magnitudes

$$\frac{\text{Torque } M}{\text{Angle } \beta} = D^* \tag{6.3}$$

will be called the "torsion coefficient" of the spring.

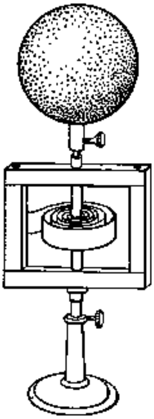


Figure 6.5 A small torsion shaft, mounted vertically and carrying a sphere, is the basis of a *torsional pendulum*. The torsion shaft makes use of the bending elasticity of a spiral spring. Its torsion coefficient is $D^* = 0.055$ N m/rad.

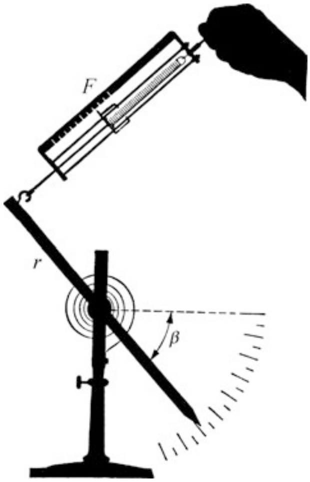


Figure 6.6 Calibration of the torsion shaft from Fig. 6.5 in a horizontal position. For example: $r = 0.1$ m, $\beta = 180° = \pi = 3.14$; $F = 1.71$ N; $rF = 0.171$ N m; $D^* = 0.055$ N m/rad.

Numerical examples

See the captions of Figs. 6.5 and 6.6. Radian (rad) is another name for the number 1, the unit of an angle (cf. Sect. 1.5).

In analogy to helical springs of known *spring constant D*, in the following we will frequently require a spiral spring (together with its axle) of known *torsion coefficient* D^*. Therefore, we *calibrate* the torsion shaft sketched in Fig. 6.5 using the straightforward scheme shown in Fig. 6.6. A numerical example is given in the figure caption. The axle and spiral spring are often replaced by a torsion wire or band; but a torsion shaft with a spiral spring is a particularly clear-cut arrangement.

Beginners often underestimate the ability to twist even thick steel rods. Figure 6.7 shows a steel rod of 1 cm diameter and only 10 cm long which is clamped in a vise. This apparently very rigid object can be twisted visibly with just our fingertips. We simply need to use a light pointer of around 10 m in length. It is reflected from the mirrors *a* and *b* (compare Fig. 3.1).

Video 6.1:
"Twisting a rod"
http://tiny.cc/csqujy

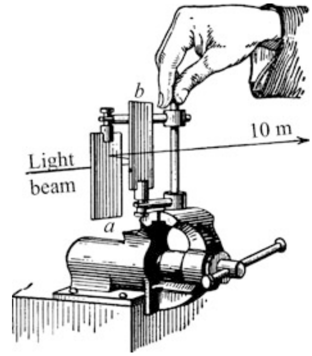


Figure 6.7 Two fingers twist a short, thick steel rod. **(Video 6.1)**

C6.3. In analogy to the orbital velocity $\boldsymbol{u}$, which describes a progressive motion along an orbit, the *angular velocity* $\boldsymbol{\omega}$ is useful for the description of rotational motion. Since during the rotation of a rigid body, all its point masses move along circular orbits around the axis of rotation, they all have the same angular velocity $\omega = u/r$ (Eq. (2.11)), which can therefore be attributed to the body as a whole.

The angular velocity[C6.3] has already been defined by the equation

$$\omega = \frac{\mathrm{d}\beta}{\mathrm{d}t}\,. \tag{2.5}$$

The orbital velocity $\boldsymbol{u}$ is fully defined only when both its magnitude *and* its direction are completely determined; it is a vector. The same is true of the angular velocity $\boldsymbol{\omega}$. Equation (2.5) gives its magnitude. The vector of the angular velocity is drawn along the direction of the axis of rotation. Figure 6.8 illustrates this. A point P circles the axis I with the angular velocity ω_1 and simultaneously the axis II with the angular velocity ω_2. Within a sufficiently short time interval Δt, the point traverses the nearly linear orbital segment $\Delta s = P\ldots3$. This segment Δs can be obtained as the resultant from the vector addition of the individual orbits

$$\Delta s_1 = \omega_1 r \Delta t \quad \text{and} \quad \Delta s_2 = \omega_2 r \Delta t\,.$$

A second route can also lead us to the orbit $P\ldots3$. We draw vectors with the magnitudes of the angular velocities ω_1 and ω_2 along the two axes I and II. These two vectors add vectorially to give the resultant

C6.4. From experience, it is found that the vector addition of angular velocities (or of axial vectors in general) often causes difficulties for beginners. The intuitive description given here is very helpful for overcoming them.

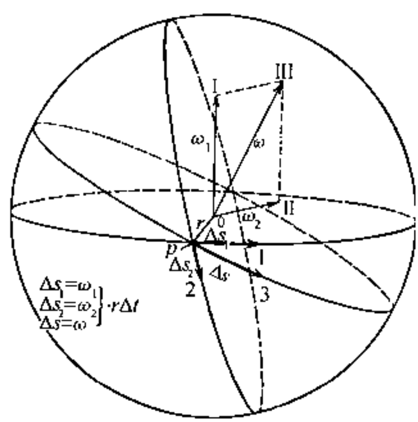


Figure 6.8 The angular velocity as a vector [C6.4]. Looking along the direction of the arrow, one sees a clockwise rotation. (See Fig. 6.31)

angular velocity $\boldsymbol{\omega}$. It determines a new axis *III*, and around this new axis, the body rotates with the angular velocity $\boldsymbol{\omega}$. It then traverses the orbit $\Delta s = \omega r \Delta t$ in the time interval Δt. The vector addition of two angular velocities can be understood readily in this example. One need only remember the similarity of the two triangles used in the construction.

6.4 The Moment of Inertia, Equation of Motion for Rotations, and Torsional Oscillations

Once we have understood the concepts of the torque $\boldsymbol{M}$ and the torsion coefficient D^*, making the transition from progressive motion to rotational motion is not difficult. We refer to Table 6.1. Its two upper

Table 6.1 Comparison of the corresponding quantities for progressive motions and rotational motions

Progressive motions			Rotational motions	
Velocity $u = \frac{\mathrm{d}x}{\mathrm{d}t}$	(2.2)	1	Angular velocity $\omega = \frac{\mathrm{d}\beta}{\mathrm{d}t}$	(2.5)
Acceleration $a = \frac{\mathrm{d}u}{\mathrm{d}t}$		2	Angular acceleration $\dot{\omega} = \frac{\mathrm{d}\omega}{\mathrm{d}t}$	
Mass m		3	Moment of inertia $\Theta = \int r^2 \mathrm{d}m$	(6.4)
Equation of motion $\boldsymbol{a} = \frac{\text{force } \boldsymbol{F}}{\text{mass } m}$	(3.3)	4	Equation of motion for rotations $\dot{\boldsymbol{\omega}} = \frac{\text{torque } \boldsymbol{M}}{\text{moment of inertia } \Theta}$	(6.7)
$\frac{\text{force } F}{\text{displacement } x} =$ spring constant D	(4.15)	5	$\frac{\text{torque } M}{\text{angle } \beta} =$ torsion coefficient D^*	(6.3)
Oscillation frequency $\nu = \frac{1}{2\pi}\sqrt{\frac{D}{m}}$	(4.16)	6	Oscillation frequency $\nu = \frac{1}{2\pi}\sqrt{\frac{D^*}{\Theta}}$	(6.13)
Work $W = \int F_x \mathrm{d}x$	(5.1)	7	Work $W = \int M \mathrm{d}\beta$	
Kinetic energy $E_{\mathrm{kin}} = \frac{1}{2} m u^2$	(5.13)	8	Kinetic energy $E_{\mathrm{kin}} = \frac{1}{2} \Theta \omega^2$	(6.5)
Momentum $\boldsymbol{p} = m\boldsymbol{u}$	(5.24)	9	Angular momentum $\boldsymbol{L} = \Theta \boldsymbol{\omega}$	(6.14)
Power $\dot{W} = Fu$	(5.35)	10	Power $\dot{W} = M\omega$	
Force $\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}$	(5.41)	11	Torque $\boldsymbol{M} = \frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}$	(6.15)

rows contain the two kinematic quantities velocity and acceleration (left) and the corresponding quantities for rotations (right): the angular velocity $\boldsymbol{\omega}$ and the angular acceleration $\dot{\boldsymbol{\omega}}$. In the following rows, we see in the *left* column the quantities for *progressive* motions with which we are already familiar, in the order in which they were introduced.

We can then calculate the corresponding quantities for rotational motions, starting with the *kinetic energy* of a body which is rotating around its axis. This energy must consist of the sum of all the kinetic energies of each of the point masses which constitute the body, each with mass Δm. Any arbitrarily chosen point mass moves at a distance r_i from the axis of rotation with the orbital velocity u_i. Then the kinetic energy of this point mass or 'volume element' is:

$$\Delta(E_\mathrm{kin})_\mathrm{i} = \tfrac{1}{2}\Delta m_\mathrm{i} u_\mathrm{i}^2 \,.$$

From Eq. (2.11), we use the same angular velocity $\omega = u/r$ for all the point masses and obtain:

$$\Delta(E_\mathrm{kin})_\mathrm{i} = \tfrac{1}{2}\Delta m_\mathrm{i} r_\mathrm{i}^2 \omega^2 \,.$$

Taking the sum over all the point masses yields the kinetic energy of the whole rotating body, i.e.

$$E_\mathrm{kin} = \tfrac{1}{2}\sum(\Delta m_\mathrm{i} r_\mathrm{i}^2)\omega^2 \,.$$

The summation which stands to the left of ω^2 has a special name, namely[C6.5]

C6.5. r is the perpendicular distance of the point mass dm from the axis of rotation. Therefore, one often writes alternatively $\Theta = \int r_\perp^2 \,\mathrm{d}m$.

$$\text{the *moment of inertia*} \quad \Theta = \sum(\Delta m_\mathrm{i} r_\mathrm{i}^2) = \int \mathrm{d}m\, r^2 \,. \quad (6.4)$$

Using this definition, the *kinetic energy of a body rotating with the angular velocity ω* becomes:

$$E_\mathrm{kin} = \tfrac{1}{2}\Theta\omega^2 \,. \tag{6.5}$$

Now refer in Table 6.1 to the eighth row on the right; the corresponding equation for *progressive* motions, on the left, is

$$E_\mathrm{kin} = \tfrac{1}{2} m u^2 . \tag{5.13}$$

In words: For rotational motions, we replace the orbital velocity u by the angular velocity ω, and the mass m by the moment of inertia Θ. This is indicated in the third line of Table 6.1 on the right.

The equation of motion (row 4 in Table 6.1) can also be transformed in a corresponding manner. To see this, we imagine a rigid, rotating body to be composed of many point masses dm, each at a perpendicular distance r from the axis of rotation. Then in the case of an accelerated rotation, the *orbital* accelerations a of the individual mass points all have *different magnitudes*; however, the *angular* accelerations $\dot{\omega} = a/r$ are the same for all points. According to the equation of motion, the orbital acceleration a of each mass point requires a force which acts in the direction of the *momentary orbit*,

$$\mathrm{d}F = a\,\mathrm{d}m = \dot{\omega} r\,\mathrm{d}m .$$

Multiplication by r yields

$$r\,\mathrm{d}F = \dot{\omega} r^2\,\mathrm{d}m ,$$

or, as a vector equation,

$$\boldsymbol{r} \times \mathrm{d}\boldsymbol{F} = \dot{\boldsymbol{\omega}} r^2 \mathrm{d}m$$

and, after integration over all the mass points,

$$\int \boldsymbol{r} \times \mathrm{d}\boldsymbol{F} = \boldsymbol{M} = \dot{\boldsymbol{\omega}} \int r^2\,\mathrm{d}m = \dot{\boldsymbol{\omega}}\,\Theta . \tag{6.6}$$

The two integrals define the two new derived quantities, the torque $\boldsymbol{M}$ and the moment of inertia Θ. To produce a torque experimentally, as on the left side of the equation, we need only a single force $\boldsymbol{F}$, which acts on the rigid body at a known distance $\boldsymbol{r}$ from its axis of rotation (Eq. (6.1)). The resulting angular acceleration is directly proportional to the magnitude of the torque $\boldsymbol{M}$ and proportional to the inverse of the moment of inertia Θ; thus, the equation of motion for rotational motion (row 4 in Table 6.1, right) is

$$\dot{\boldsymbol{\omega}} = \frac{\boldsymbol{M}}{\Theta} . \tag{6.7}$$

C6.6. *Homogeneous* means that the mass is uniformly distributed within the body, i.e. $\varrho = \text{const}$.

For geometrically simple rigid bodies, the *computation* of the moment of inertia presents no difficulties. The necessary integration can usually be carried out in a few steps. Some examples of the results:

1. A *homogeneous*[C6.6] *circular ring*. Mass m, radii R and r, thickness d, density ϱ, axis of rotation through its center point[C6.7]:

perpendicular to the end surface: $$\Theta = \frac{\pi}{2}\varrho d(R^4 - r^4) , \tag{6.8}$$

along a diameter: $$\Theta = \frac{m}{12}d^2 + \frac{\pi}{4}d\varrho(R^4 - r^4) . \tag{6.9}$$

C6.7. We add that the expressions (6.8) and (6.9) hold quite generally for a homogeneous, hollow cylinder. Equation (6.9) can be simplified for the two limiting cases $d \ll R$ (a flat disk with a hole in its center) to
$\Theta = \frac{\pi}{4}d\varrho(R^4 - r^4)$
$= \frac{1}{4}m(R^2 + r^2)$,
i.e. just half of expression (6.8), and for $d \gg R$ (a long rod) to Eq. (6.11).

2. A *homogeneous sphere*, with the axis of rotation through its center:

$$\Theta = \frac{8}{15}\pi\varrho R^5 = \frac{2}{5}mR^2 . \tag{6.10}$$

3. A *long, homogeneous rod* of length l and cross-sectional area A. The axis of rotation passes through the center of gravity and is perpendicular to the long axis of the rod:

$$\Theta = \frac{1}{12}\varrho A l^3 = \frac{1}{12} m l^2 \,. \tag{6.11}$$

4. STEINER's *law*. Assuming that we know the moment of inertia Θ_S of an arbitrary body of mass m for an axis of rotation which passes through its *center of gravity* S: How large is its moment of inertia Θ_A for any other arbitrary axis which is parallel to the first axis and passes through the point A at a distance a from the first axis? Answer:

$$\Theta_A = \Theta_S + m a^2 \,. \tag{6.12}$$

Derivation

Rotating around the S-axis, the body will have the kinetic energy $\frac{1}{2}\Theta_S\omega^2$. – Figure 6.9 shows a rotation around the A-axis. A small arrow is drawn on the body, which starts at the center of gravity S. If the body makes a full rotation around the A-axis, then this arrow, and thus the entire body, makes a full rotation around the S-axis. As a result, the energy given above is conserved, $\frac{1}{2}\Theta_S\omega^2$. At the same time, the center of gravity S moves along the dashed *circular orbit*. We can think of the mass m of the body as localized at its center of gravity, and we then find for the kinetic energy of this circular motion $\frac{1}{2}mu^2 = \frac{1}{2}m(\omega a)^2$. This second energy (from the motion of the center of gravity) adds to the first (from the rotation around the A-axis). Thus the rotation of the body around the A-axis has a total kinetic energy of:

$$\tfrac{1}{2}\Theta_A\omega^2 = \tfrac{1}{2}\Theta_S\omega^2 + \tfrac{1}{2}m\omega^2 a^2 \,.$$

Dividing by $\frac{1}{2}\omega^2$ yields Eq. (6.12).

However, much more important than the computation of moments of inertia is their measurement. For bodies with a complex shape, the integration would be unnecessarily difficult.

For the measurement of moments of inertia, one in general makes use of torsional oscillations. We need only replace the mass m in row 6 of Table 6.1 by the moment of inertia Θ and the spring constant D of a helical spring by the torsion coefficient D^* of a spiral spring. The *torsion shaft* as shown in Fig. 6.5 will define a known value of D^*. At the upper end of this shaft, we attach the body that we wish to characterize (cf. Fig. 6.5). The axis of rotation of this body must fall along the extension of the torsion shaft. We rotate the body by about 90° from its rest position, release it, and observe the period of the

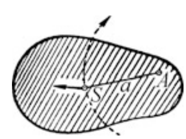


Figure 6.9 The straightforward derivation of STEINER's law; S = center of gravity

resulting torsional oscillations, T, with a stopwatch. We have[C6.8]

$$\Theta = \frac{T^2}{4\pi^2} D^* \,. \tag{6.13}$$

C6.8. Analogously to a linear oscillation (Fig. 4.13), we obtain an expression for the period of oscillation of a torsional motion (6.13) by inserting the "linear restoring force law" (6.2) into the equation of motion for rotational motion, (6.7) (cf. Comment C4.5.).

The torsion coefficient D^* of our small torsion shaft was already found to be $5.5 \cdot 10^{-2}$ N m. Thus, we obtain

$$\Theta = 1.4 \cdot 10^{-3} \frac{T^2}{\mathrm{s}^2} \mathrm{kg\,m}^2 \,.$$

Numerical examples

1. *Verifying a calculated moment of inertia*. Taking a circular wooden disk with $m = 0.8$ kg and 0.2 m radius, we calculate from Eq. (6.8) with $r = 0$ that the moment of inertia Θ_S should be $1.6 \cdot 10^{-2}$ kg m^2 for rotation around an axis passing perpendicularly through the center point of the disk. We observe $T = 3.37$ s, and thus $\Theta_S = 1.58 \cdot 10^{-2}$ kg m^2.
2. *A disk and a sphere of equal moments of inertia*. Figure 6.10 shows a disk and a sphere made of the same material, to the same scale. Their masses are in the ratio 1 : 2.9. Their moments of inertia should be equal, according to Eqns. (6.8) and (6.10). Indeed, they both exhibit the same oscillation period when attached to our small torsion shaft.
3. *Moments of inertia of hollow and full cylinders of the same mass*. Figure 6.11 shows a hollow metal cylinder and a full wooden cylinder, both with the same masses m, the same diameters and the same lengths. However, for rotations around their cylinder axes, the hollow cylinder is found to have a considerably larger moment of inertia Θ than the full cylinder.
This explains an often surprising experimental result: We set the two cylinders side by side on a ramp, for example a board propped up to give it a slope. The axes of the two cylinders are collinear. Then we release them both at the same time. The full wooden cylinder reaches the bottom of the ramp much sooner than the hollow metal cylinder. – Explanation: In rolling down the ramp, both cylinders are accelerated by the same torque, rF_G (Fig. 6.12), since their masses and radii are the same. As a result, the hollow cylinder, with its larger moment of inertia, experiences a smaller

Figure 6.10 A disk and a sphere with the same moment of inertia

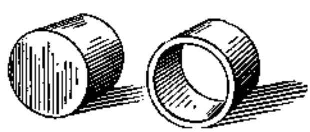


Figure 6.11 A full and a hollow cylinder of the same mass (wood and metal), with different moments of inertia

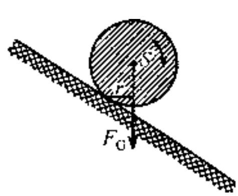


Figure 6.12 The torque $M = rF_G$ acting on a cylinder on a ramp

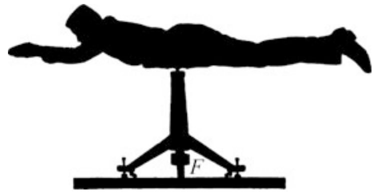


Figure 6.13 A large torsion shaft for measuring the moment of inertia of a human body in various positions. F is a strong spiral spring. Its torsion coefficient D^* is ca. 2.5 N m/rad. The moment of inertia Θ of a man lying horizontally is about 17 kg m^2. **(Video 6.2)**

Video 6.2:
"Moments of inertia"
http://tiny.cc/gsqujy
In the picture: R. HILSCH
(Dr. rer. nat. 1927)

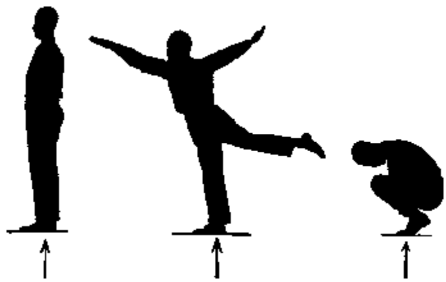


Figure 6.14 The moment of inertia of a human body in three different positions. The arrows indicate the direction of the axis of rotation. The moments of inertia (from the left to the right) are 1.2 kg m^2, 8 kg m^2, and 2.3 kg m^2. **(Video 6.2)**

angular acceleration $\dot{\omega}$ and angular velocity ω (row 4 in Table 6.1). – (Why must STEINER's law be taken into account here[C6.9]?)

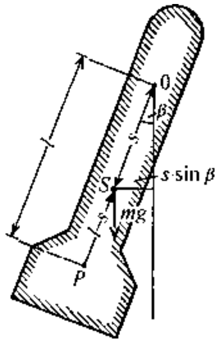
4. *The moment of inertia of the human body.* We determine the moment of inertia for rotation of a human body with various positions and axes of rotation. For these observations, we make use of a large torsion shaft as shown in Fig. 6.13. Some of the results of the measurements are collected in Fig. 6.14. We will make good use of them later.

C6.9. Those who investigate this question will also find the explanation of the similar, often surprising observation that the mass plays no role at all! (**Exercise 6.2**)

6.5 The Physical Pendulum and the Beam Balance

The gravity pendulum (simple pendulum) treated in Sect. 4.3 is also called a "mathematical pendulum". It represents the ideal case of a point-like body of mass m hung from a massless, frictionless cord. A real or "physical" pendulum often deviates strongly from this ideal form. For every physical pendulum, a "reduced" length can be defined: It is the length l of a mathematical pendulum which would have the same period of oscillation as the given physical pendulum.

Figure 6.15 The physical gravity pendulum (Axes through O or P, perpendicular to the plane of the page)



As an example, Fig. 6.15 shows a board of arbitrary shape, hung from a bearing as a gravity pendulum. O is the axis around which it can swing, S is its center of gravity, and s is the distance between the two. For the period of oscillation of this physical pendulum, the formula which holds for every torsional oscillation can be applied:

$$T = 2\pi\sqrt{\frac{\Theta_0}{D^*}}\,. \tag{6.13}$$

Θ_0 is the moment of inertia around the axis O. D^* is again the torsion coefficient, i.e. $D^* = M/\beta$. The magnitude of the torque M can be read off in Fig. 6.15:

$$M = mgs\sin\beta\,. \tag{6.1}$$

For small angles β, we can again set $\sin\beta = \beta$. We then obtain

$$D^* = M/\beta = mgs \tag{6.3}$$

and, from Eq. (6.13),

$$T = 2\pi\sqrt{\frac{\Theta_0}{mgs}}\,.$$

For a "mathematical" gravity pendulum, i.e. a point-like body hanging from a massless cord, we found the period previously to be

$$T = 2\pi\sqrt{\frac{l_0}{g}}\,. \tag{4.17}$$

For a *physical* pendulum, instead of the length of the cord l_0 of the mathematical pendulum, we insert the quantity Θ_0/ms. This is the reduced pendulum length. It is drawn in as a length $l = \Theta_0/ms$ in Fig. 6.15[C6.10]. Its lower end is called the midpoint of the oscillation P. We can think of the total mass m of the body as being concentrated at the point P, without changing the period of oscillation of the pendulum.
The period of oscillation of an arbitrary pendulum remains unchanged when one takes its axis of oscillation to pass through the midpoint of oscillation P[C6.11]. This forms the basis for a common experimental method of measuring the reduced pendulum length (reversion pendulum).

C6.10. l is longer than s: Making use of STEINER's law (Eq. (6.12)), it follows that
$l = \dfrac{\Theta_0}{ms} = \dfrac{\Theta_S + ms^2}{ms}$,
$l = s + \dfrac{\Theta_S}{ms} > s\,.$
(Θ_S = moment of inertia around the center of gravity.)

C6.11. To convince oneself of the correctness of this statement, use the torsion coefficients for s and $(l-s)$, $D_0^* = mgs$ and $D_P^* = mg(l-s)$, and, applying STEINER's law, the corresponding moments of inertia $\Theta_0 = \Theta_S + ms^2$ and $\Theta_P = \Theta_S + m(l-s)^2$, in Eq. (6.13) and solve for the reduced pendulum length l.

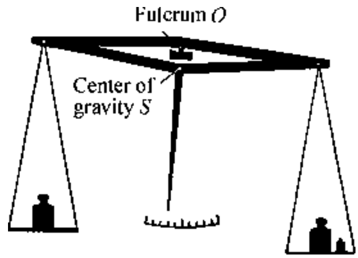


Figure 6.16 Schematic of a beam balance as a physical pendulum. In the interest of clarity, the distance from the center of gravity S to the axis of oscillation O (fulcrum) is greatly exaggerated in the drawing. Up to angular deflections of several degrees, the deflection is proportional to the difference in the weights on the balance pans.

C6.12. The mechanical beam balance as a measuring instrument (see e.g. http://en.wikipedia.org/wiki/Analytical_balance) has largely disappeared from modern physics and chemistry laboratories and has been replaced by sensitive electronic analytical balances. However, it remains an instructive example of a torsional pendulum (**Exercise 6.7**).

C6.13. The angular momentum introduced here depends, like the moment of inertia, on the axis of rotation chosen. If the axis passes through the center of gravity, we refer to the "proper angular momentum" or "spin" of the body. In other cases, one can insert Eq. (6.12) into Eq. (6.14) and obtain $\boldsymbol{L} = \Theta_s\,\boldsymbol{\omega} + ma^2\,\boldsymbol{\omega}$. The first term is again the proper angular momentum. The second term is called the "orbital angular momentum", which the body possesses due to its motion along the circular orbit of radius a (see Fig. 6.9). In general, the orbital angular momentum is defined by the vector equation $\boldsymbol{L} = m\boldsymbol{r} \times \boldsymbol{v}$, where $\boldsymbol{r}$ is the position vector and $\boldsymbol{v}$ is the orbital velocity (see also **Exercise 6.8**).

One of our most important measuring instruments, the *beam balance*[C6.12], is a physical pendulum. – We first neglect the two balance pans. Then the schematic in Fig. 6.16 differs only in its external details from that in Fig. 6.15.

The period of oscillation of a precision analytical balance without pans might be e.g. 12 s, corresponding to a reduced pendulum length of 36 m.

Adding the balance pans and their weights increases the effective moment of inertia of the balance beam. This also increases the oscillation period, e.g. to 18 s. An additional mass of 100 g on both sides increases it even further up to about 24 s.

6.6 Angular Momentum

For progressive motions, the momentum was defined as $\boldsymbol{p} = m\boldsymbol{u}$. Momentum is a vector, and for the momentum of a "closed system", a conservation law holds.

In rotational motions, the mass m must be replaced by the moment of inertia Θ, and instead of the orbital velocity $\boldsymbol{u}$, we use the angular velocity $\boldsymbol{\omega}$. These two quantities define the corresponding momentum (row 9 in Table 6.1), the *angular momentum*[C6.13]

$$\boldsymbol{L} = \Theta\,\boldsymbol{\omega}\,. \qquad (6.14)$$

The angular momentum is also a vector. It specifies the direction and the sign or sense of the rotation. Looking along the direction of the (positive) angular momentum vector, the sign of rotation is clockwise. *In the text, the sense of rotations applies to an observer who is looking down from above.*

Figure 6.17 The conservation of angular momentum. The arrows represent the angular momenta of the wheel and the experimenter. (At *small* angular velocities, the motion is disturbed by the residual friction.) **(Video 6.3)**

Video 6.3:
"Conservation of angular momentum using a rotating chair"
http://tiny.cc/4rqujy
In the picture: R. HILSCH (Dr. rer. nat. 1927)

Angular momentum also obeys a conservation law, i.e. as long as no external torques act, the angular momentum of a system (magnitude and direction) remains constant ($\boldsymbol{M} = 0$ in the equation of motion for rotations (6.7)). Just as in our treatment of progressive motions, we give some experimental examples. As an aid to visualization and for experiments, instead of the flat *carts* that we used to illustrate progressive (linear) motions (Fig. 5.14), for rotations we will employ a *swivel chair* (Fig. 6.17). It can rotate with little friction (ball bearings) around a vertical axis. It thus reacts only to vertically-directed angular momenta. If the angular momentum vector is slanted relative to the vertical, the chair takes on only its *vertical component*.

Examples

1. A man is sitting in the swivel chair, at rest. In his left hand, at about eye level, he is holding a top (a weighted bicycle wheel), also at rest, with its axis of rotation nearly vertical. The angular momentum is initially zero. The man grasps the spokes of the wheel from below with his right hand and starts the wheel spinning. The spinning top thus obtains an angular momentum of $\Theta_1 \boldsymbol{\omega}_1$ in a counter-clockwise sense. Owing to the conservation of angular momentum, the man receives a (counter-) angular momentum $\Theta_2 \boldsymbol{\omega}_2$ of the same magnitude but opposite rotation sense. Indeed, he begins to rotate with the chair in a clockwise sense. His angular velocity $\boldsymbol{\omega}_2$ is much smaller than that of the top (wheel), because his moment of inertia is much larger than that of the wheel.

2. The man presses the rim of the rotating wheel against his chest and brings it to a stop. The rotation of the wheel and of the man and chair both stop simultaneously. Both angular momenta are again zero.

3. The man sits motionless on the chair holding the motionless top with its axis horizontal. He gives a spin to the top (wheel); its resulting angular momentum vector is horizontal. The chair and the man remain at rest, since they cannot react to an angular momentum with no vertical component (the counter-component of the angular momentum is taken up by the earth through the man's arm and the shaft of the chair).

Part I

Video 6.3:
"Conservation of angular momentum using a rotating chair"
http://tiny.cc/4rqujy

Figure 6.18 A polo mallet can be used to generate angular momentum around various axes **(Video 6.3)**

4. The initially motionless top is held with its axis inclined at 60° to the vertical and then is given a spin. The man and the chair begin to rotate, but with only a small angular velocity. They receive a counter-angular momentum which is equal only to the vertical component of the top's angular momentum.

5. We hand a (clockwise) spinning top to the man, who is initially not moving. He remains at rest; we gave him the top which already had an angular momentum. Then the man rotates the axis of the spinning top by 180°. He turns its lower end upwards. He thereby changes its angular momentum from $+L$ to $-L$, thus all together by $-2L$. The man himself begins to rotate with an angular momentum of $2L$ in a clockwise sense. He then tips the top back around to its starting position and gives it back to us. The chair and the man are again motionless. – Thus, one can play with a "borrowed" angular momentum for a time and then give it back.

6. The man is sitting on the swivel chair, at rest. In his hand he is holding a polo mallet (Fig. 6.18). He will try to circle the mallet around his head in the horizontal plane, that is around a vertical axis of rotation. – While he is swinging the mallet, he begins to rotate, although with a smaller angular velocity than his arm and the mallet. The mallet and his arm can be circled through only around 180°. As soon as the motion of the mallet stops, so does the rotation of the man and chair; the man and the mallet can have angular momenta only at the same time. For a second swing, he has to return the mallet to its original position; he can do this along the same path, but then he loses the angle of rotation that he had gained with the first swing of the mallet. If he wants to maintain his angular position, he must return the mallet along a *different* path before beginning another swing; he moves it upwards in the *vertical* plane and then back down in another *vertical* plane back to its original position. The angular momenta of these motions in vertical planes (i.e. around horizontal axes of rotation) cannot be taken on by the chair with its vertical axis of rotation (they are transferred to the earth). From the mallet's original position, he can repeat the circular motion in the horizontal plane and thereby double his angular gain from his initial position. These three individual motions can of course be combined into a single motion, so that the man's arm and the mallet move along the mantle of a cone whose axis is inclined by for example 45° from the vertical.

Figure 6.19 The vector nature of angular momentum (α can be varied between 30° and 100°) **(Video 6.4)**

Video 6.4:
"Angular momentum as a vector"
http://tiny.cc/ltqujy

7. The vector nature of angular momentum can be demonstrated effectively with a fan which can rotate around a vertical axis (Fig. 6.19). The propeller and the air stream receive an angular momentum $\boldsymbol{L}$; the fan takes on the counter-angular momentum of equal magnitude but opposite rotational sense. The vertical component, i.e. $L \cos\alpha$, causes the fan itself to rotate around the vertical axis (see also Fig. 6.2.)

Examples

The propeller blades of the fan are supposed to be turning in a clockwise sense when we look through them towards the housing of the fan's motor. – First, the fan is blowing in a horizontal direction, that is $\alpha = 90°$. The fan remains at rest on the vertical axis of its stand, since $\cos 90°$ is zero.
Now the fan is blowing at an angle upwards, e.g. with $\alpha = 45°$. The fan rotates on its holder; seen from above, it rotates slowly in a counter-clockwise sense. The reason: We have $\cos 45° \approx 0.7$; $\cos\alpha$ thus has a positive, nonzero value. (The frictional losses in the bearings of the holder present no problem, because the motor and propeller are continually transferring more angular momentum to the air stream.)
When the power to the motor is shut off, this rotation initially stops, and then it begins again in the same sense of rotation as the propeller. The reason: The rotor of the motor and the propeller are slowed down by *friction* in the motor bearings. They give up their remaining angular momentum to the housing of the motor, and we can observe its vertical component.

8. We replace the swivel chair by the large *torsion shaft* shown in Fig. 6.13. A man is lying stretched out on it, holding on to two handles (Fig. 6.20). He is given an impulse and begins to exhibit torsional oscillations with a small amplitude. *Problem*: We want the man to increase his oscillation amplitude up to full circular oscillations of 360° without external help. *Solution*: He has to change his moment of inertia around the vertical axis periodically. When he passes through the rest position, he pulls in his legs and raises his torso. This reduces his moment of inertia Θ and increases his angular velocity ω[1]. At the extremes of his oscillations, the reversal points, he stretches out again and returns to a maximum moment of inertia. At the next passage through the rest position, he repeats this action. In a short time, he will be exhibiting torsional oscillations

[1] The gain in energy comes from the work performed by the man's muscles against inertial forces (Sect. 7.3).

Part I

Video 6.5:
"The physics of swinging"
http://tiny.cc/ssqujy

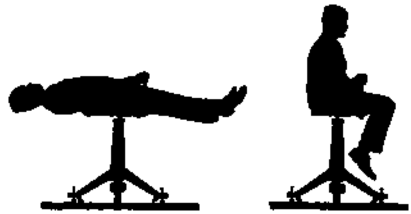


Figure 6.20 Bar exercises with swings, simulated on a swivel chair **(Video 6.5)**

with an amplitude of ±180°. – This experiment demonstrates in an excellent manner the entire *technique of exercising on the bars*. Here, the horizontal axis of the bars is replaced by the vertical torsion shaft and the torque is produced not by the weight of the gymnast, as on the bars, but instead by the spiral spring attached to the torsion shaft. This has the advantage that the motions are slower, so that we can more easily observe them. The experiment just described is, in the language of gymnastics, a *giant swing* (see http://www.ncbi.nlm.nih.gov/pubmed/10433423).

The gymnast on the bars knows how to reduce his or her moment of inertia in various ways at the right moment. For example, in the giant swing, by bending in the arms or legs or spreading the legs[C6.14].

C6.14. A playground swing offers another good opportunity to study this technique, which children call "pumping up" and physicists call "parametric excitation of oscillation", in which a parameter of the oscillator is varied, thus exciting vibrations (compare also Fig. 11.20).

9. The law of areas for motions around a central point (Sects. 4.4 and 4.5) is simply a special case of the conservation of angular momentum. In Fig. 4.17, at all points along the orbit, the triangular area $Oac = Oce = r^2\omega\Delta t/2 = L\Delta t/2m$ (Eq. (6.12)) remains constant.

6.7 Free Axes

In all of the rotational motions that we have considered so far, the axis of rotation of the rotating body was *fixed* as a real shaft mounted on bearings. We now drop this limitation. This will lead us to rotational motions of bodies around *free axes*. In order to explain this term, we will give several experimental examples.

a) Figure 6.21 shows a well-known circus trick: A flat plate is rotating at the tip of a bamboo stick. Its axis of symmetry (perpendicular to the plate) serves as a free axis of rotation.

b) A flat plate, if skilfully set in motion, can also rotate around an axis parallel to its diameter (in the plane of the plate) (Fig. 6.22).

c) We now demonstrate a small variation on these two experiments: we hang a cylindrical rod from one end onto the rapidly rotating shaft of an electric motor. Either its cylinder axis or – as in Fig. 6.23 – a transverse axis can serve as a free axis of rotation.

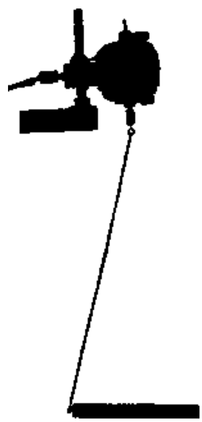
d) For technical applications, free axes can be used in the form of "supple" shafts. In Fig. 6.24, a grinding wheel is set in rotation by an electric motor ($\nu \approx 50\,\text{Hz}$). The wheel is mounted at the end

Figure 6.21 The figure axis of a plate as a free axis of rotation[C6.15]

C6.15. The author as a circus performer (the photo was taken by one of the students in his lecture course).

Figure 6.22 A diameter of the plate as a free axis of rotation

Figure 6.23 A rod rotating around the axis of its largest moment of inertia as a free axis of rotation **(Video 6.6)**



Video 6.6:
"Rotation around free axes"
http://tiny.cc/itqujy

Figure 6.24 The supple shaft of a grinding wheel (this shaft is too thin for practical applications!) **(Video 6.7)**

Video 6.7:
"Supple shaft as stable axis of rotation"
http://tiny.cc/5squjy

of a ca. 20 cm long wire with a diameter of only a few millimeters. It rotates in a stable manner around the axis of its largest moment of inertia and makes a springy contact with a workpiece which is pressed against it [C6.16].

C6.16. A practical application:
The unavoidable differences in direction between the geometric axes and the angular momentum vectors, which lead to "hammering" of a rapidly-rotating gyroscope or top, can be absorbed by supple shafts.

All of these examples have two points in common:

1. The bodies used have *rotational symmetry*. All of them could have been fabricated on a *lathe*, in principle. In each case, the body is characterized by a *figure axis or body axis*.

2. One of the free axes was a figure axis, while the other was perpendicular to it.

In the following demonstrations, the rotating bodies lack rotational symmetry. As an example, we take a flat cigar box (Fig. 6.25). Its three pairs of sides are each painted in a different color.

e) Eyelets are mounted at the center of each side. The box is then hung by a wire from the shaft of a motor, like the cylindrical rod in Fig. 6.23. The experiment demonstrates the following: The mid-lines *A* and *C* can serve as "free" axes, and the object can exhibit stable rotation around them. These two free axes are perpendicular to each other. – The third mid-line *B*, which is perpendicular to *A* and *C* and likewise passes through the center of gravity of the object, shows a different behavior. It cannot serve as a free axis of rotation. The object always returns to a stable rotation around one of the other two axes [C6.17].

C6.17. The most stable rotation always occurs around the axis with the largest moment of inertia (here, axis *A*).

f) We repeat this last experiment with a variation; we throw the box into the air, giving it a spin by a suitable positioning of our fingers on releasing it (Fig. 6.26). Again, *A* and *C* can serve as free axes of rotation. The same side of the box remains facing the observer, as can be seen from its color. Attempts to spin the box around axis *B* always lead to wobbling motions, so that the observer sees rapidly changing colors.

With these and other, similar experiments, one can arrive at the conclusion that *the axes of a body with the largest or the smallest moments of inertia can serve as free axes of rotation*.

In the examples a) to f), which were chosen for their simplicity, the axes are in each case readily visible from the geometry of the objects. In other cases, one can always make use of a torsion shaft (Figs. 6.5 and 6.13) to *measure* the moments of inertia of axes in different di-

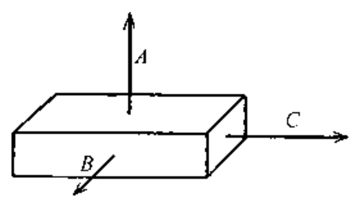


Figure 6.25 The axis *A* with the largest moment of inertia, axis *B* with an intermediate moment of inertia, and axis *C* with the smallest moment of inertia of a box

$$\left.\begin{aligned}\Theta_A &= 6.5\\ \Theta_B &= 5.6\\ \Theta_C &= 1.4\end{aligned}\right\} \cdot 10^{-3}\,\mathrm{kg\,m^2}$$

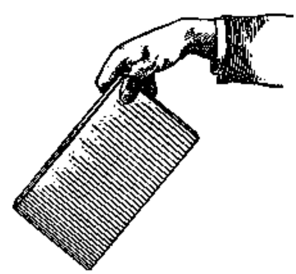


Figure 6.26 Throwing a box with a rotation around its free axis *A* (with the largest moment of inertia) **(Video 6.8)**

Video 6.8:
"Free rotation of a rectangular parallelepiped"
http://tiny.cc/1squjy

rections. For completeness, we carried out measurements of this kind on our flat cigar box; the results are set out in the caption of Fig. 6.25.

6.8 Free Axes of Humans and Animals

Free axes by no means require rotational symmetry of a body. This is shown by the experiments with the painted cigar box. Still more convincing are the uses of free axes by humans and animals.

Examples

a) *A jumper performs a somersault.* Leaning slightly forwards, usually with raised hands, he gives himself an angular momentum. The corresponding axis is indicated in Fig. 6.27a as a white spot. It is very nearly the free axis with the *largest* moment of inertia. The angular velocity is still small. A moment later, the jumper pulls his body together into a crouching position as in Fig. 6.27b. The axis of rotation remains that of the *largest* moment of inertia, but that moment is about three times smaller than before. As a result, the angular velocity increases by a factor of three, from conservation of angular momentum. This high angular velocity permits one or two, sometimes even three complete rotations while the jumper remains in the air. He then stretches out his body at the appropriate moment to restore the large moment of inertia, so that he can land on the floor with a small angular velocity. The jumping techniques of good circus performers are physically very instructive. The first requirement for jumping, however, is courage. Jumping is a matter of good nerves. The law of conservation of angular momentum automatically guarantees the necessary rotations.

"Jumping is a matter of good nerves. The law of conservation of angular momentum automatically guarantees the necessary rotations".

b) *A ballet dancer performs a pirouette* on tiptoe. She rotates around the long axis of her body. She uses the axis of her *smallest* moment of inertia as a *free* axis of rotation. She spins around this axis with a high angular velocity ω and the angular momentum $\Theta\omega$. To stop, she changes her body position at the appropriate moment to that shown in Fig. 6.14, center, and thereby increases the magnitude of her moment of inertia. This new moment of inertia is around seven times larger

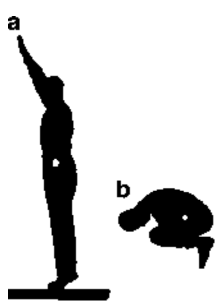


Figure 6.27 Changing the magnitude of the moment of inertia when performing somersaults[C6.18]

C6.18. In the picture: R. HILSCH (Dr. rer. nat. 1927) than before; as a result, her angular velocity is reduced by a factor of seven. She sets the sole of her foot onto the floor, braking the spin and lowering the point of support.

c) A cat which is held by its feet and then dropped always falls on its feet. The animal rotates around the axis of its smallest moment of inertia; it uses this axis as the substitute for the shaft of the swivel chair in Fig. 6.18, which is mounted in bearings. Instead of the *polo mallet*, the *hind quarters and the tail* are swung around to produce a rotation. Humans can readily imitate this cat trick in *their own* way. They can also induce rotations during a jump around their axis of smallest moment of inertia, i.e. the body's long axis.

6.9 Definition of the Spinning Top and Its Three Axes

In the rotations that we first considered, the axis of rotation was fixed with respect to the object, and in addition it was mounted on bearings outside the rotating object. We then treated rotations around free axes, where the axis of rotation was still fixed relative to the object, but was no longer held by bearings. The most general case of rotation occurs when the axis of rotation is not fixed either by external bearings or with respect to the rotating object. The axis always passes through the center of gravity of the object, but can continually change its direction within the object. This last, most general form of rotation is called *gyration*. Rotations around free axes or around shafts with bearings are special cases of this more general motion, gyration.

In their most general form, gyrations represent the most difficult problems in all of mechanics. Even with a great mathematical effort, one can obtain only approximate solutions. But all of the essentials of gyration can be understood using the special case of the *symmetric top*. This case is defined in Fig. 6.28. In the examples shown there, the symmetry axis is always the axis of the *largest* moment of inertia (it is also called the figure axis or body axis; in the following sections, we will refer to it as the "figure axis"). In the physical sense, these are *flattened* tops, gyroscopes, or simply *spinning tops* in everyday language.

A decisive point for describing and understanding gyrations is to discriminate strictly between three different axes, each of which passes through the center of gravity of the body. These are:

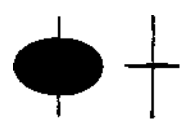


Figure 6.28 Two "flattened" tops. The symmetry or figure axis is the axis with the largest moment of inertia.

1. The figure axis (symmetry axis), and thus in our tops the axis of the largest moment of inertia of the rotating body;

2. the momentary or *instantaneous* axis of rotation, the axis around which the body is rotating at a particular, given moment; and

3. the angular-momentum axis. It lies between the axis of symmetry and the axis of rotation, and in the plane defined by those two axes. For brevity, we will refer to it in the following as the *spin axis*, since the proper angular momentum of a body is often called its "spin", just as one often refers to a "spinning top".

The figure axis is easily recognizable in each of our tops; but it requires a certain artifice to make the other two visible. The top shown in Fig. 6.29 is particularly well suited to this task. It is supported at its center of gravity by a point and socket, nearly without friction and without external forces, so that it remains in equilibrium in any position of the figure axis. Its center of gravity can be adjusted by moving the metal ring A. – The figure axis carries a disk at its upper end; paper stickers with various patterns can be placed on it. To start with, we want to demonstrate the instantaneous axis of rotation. We use a paper sticker with a printed text, start the top spinning, and give the figure axis a tap from the side. This sets the top into a wobbling motion, and we can observe the following: The text on the sticker merges into a uniform grey, owing to the rotation of the top. Only at one small spot, which moves around over the sticker, is the text briefly at rest and recognizable as printed letters. The midpoint of this spot is the *instantaneous axis of rotation*. This axis, as well as the figure axis, both move over the mantle of a cone, each with the same angular velocity $\boldsymbol{\omega}_{\mathrm{N}}$, and these two cones have a common axis which is fixed in space. This latter axis, which is at this point invisible, is the *angular momentum axis*.

In order to make the angular momentum axis visible, we start with a preliminary experiment. We attach a sticker with concentric circles

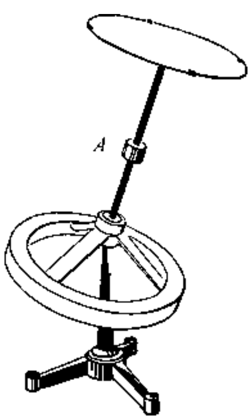


Figure 6.29 A top for demonstrating the three axes. To start it spinning, we clamp the figure axis between the palms of our hands and move them in opposite directions. **(Video 6.9)**

Video 6.9:
"The three axes of a top"
http://tiny.cc/2tqujy

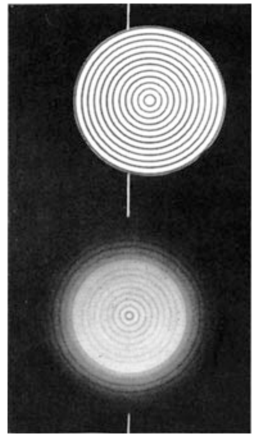


Figure 6.30 Visualizing the spin axis. About 1/10 of actual size (R. Hagedorn, Z. Phys. **125**, 542 (1949))

onto a rotating disk with the center of the circles on the axis of rotation. The rotating sticker looks just the same as when it is at rest; cf. Fig. 6.30, upper part. – Then we shift the center of the circles to the side, away from the axis of rotation of the disk; thus the center of the circles itself rotates around the axis of rotation of the disk. This yields Fig. 6.30, lower part; again a system of concentric circles. The spacing between the individual circles is the same as before, but their contours are fuzzy and washed out. The common center of these paler circles lies on the axis of rotation and shows us where it is. – So much for the preliminary experiment.

For the main experiment, we put the sticker with concentric circles onto the disk at the end of our top's figure axis (Fig. 6.29), with the common center point on the figure axis. By a tap to the side of the spinning top, we again separate the three axes from each other; the figure axis, and thus the center of the concentric circles, now rotates around the angular momentum axis. This makes the angular momentum axis visible in just the same way as in our preliminary experiment, Fig. 6.30.

This whole phenomenon, the common rotation of the figure axis and the instantaneous axis of rotation around the angular momentum axis, is called *nutation* (its angular velocity is $\boldsymbol{\omega}_{\mathrm{N}}$).

We give some more details on nutation in the following section. – Here, we simply establish from the experiment a fact which will later prove to be useful: The nutations decay after a certain time. This is a result of the unavoidable friction in the support point of the top; in our example, between the point and the socket on which the top rests.

6.10 The Nutation of a Force-Free Top and Its Fixed Spin Axis

The *nutation which we have just observed is an immediate result of the conservation of angular momentum.* Imagine that in Fig. 6.31, the plane of the paper contains the figure axis A of the top and its instantaneous axis of rotation Ω. The top is spinning around this instantaneous axis with its angular velocity $\boldsymbol{\omega}$, represented by the vector arrow along the axis of rotation, Ω. This angular velocity $\boldsymbol{\omega}$ can be decomposed into two components ω_1 and ω_2. ω_1 is the angular velocity around the axis A with the largest moment of inertia Θ_A, while ω_2 is the angular velocity around a perpendicular axis C with the moment of inertia Θ_C. The overall angular momentum is then given by $L_\mathrm{A} = \Theta_\mathrm{A}\omega_1$ in the direction of the figure axis A and $L_\mathrm{C} = \Theta_\mathrm{C}\omega_2$ in the direction of the perpendicular axis C.

These two angular momenta are shown in the figure as vector arrows with thick arrowheads. Their resultant is the total angular momentum $\boldsymbol{L}$. Its direction thus lies between the figure axis A and the instantaneous axis of rotation Ω, and within their common plane.

Now the top is by assumption "force-free". It is supported at its center of gravity by the point and socket. No torques of any kind act upon it. As a result, its angular momentum, magnitude and direction, must remain constant (see Sect. 6.6). Its spin axis must continually maintain the same direction in space. Both the figure axis A and the instantaneous axis of rotation Ω have to rotate around this fixed spin axis with the nutation frequency (ω_N). To make this clearer, in Fig. 6.32 we have indicated the three axes with stiff wires and let them rotate around the center wire, i.e. the angular momentum axis (spin axis). Then we see how two cones are traced out around this axis. One of them is traced by the wire representing the figure axis; this is the cone

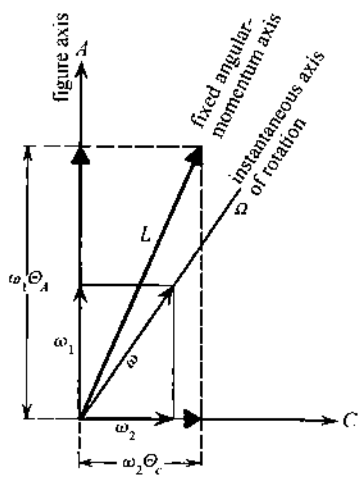


Figure 6.31 The three axes of a top

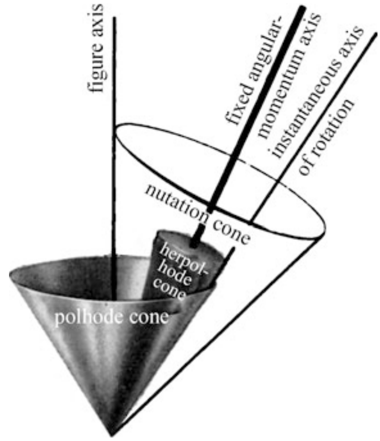


Figure 6.32 The cone of nutation

of nutation which we have already seen. The other cone is traced out by the wire representing the instantaneous axis of rotation; it is called the herpolhode cone. The relationship between these first two cones can be represented by a third cone as shown in Fig. 6.32, the polhode cone. It is attached to the figure axis, surrounds the fixed herpolhode cone and rolls along its surface "pericycloidally". The line of contact between these two cones with a common apex indicates the direction of the instantaneous axis of rotation Ω.

"Understanding the contents of this section requires a certain amount of study. But it is worth the effort."

Understanding the contents of this section requires a certain amount of study. But it is worth the effort. The word *nutation* occurs rather frequently in articles on physics and technology. One should have some understanding of what it means.

In certain cases, the angular momentum axis of a top can be collinear with its figure axis; the flat top degenerates to a spherical top, or the axis of rotation of a flat top coincides with its figure axis. – This latter case can be realized in various ways, e.g. with the top shown in Fig. 6.29. One spins the top carefully with its point support in the center of gravity, avoiding any transverse forces on the figure axis. Then its axis remains fixed in space. This is a phenomenon which is well known to many non-physicists. – Variations:

a) We throw a discus, using the well-known hand flip to start it spinning like a top. The direction of its figure axis coincides with the angular momentum axis and remains fixed (Fig. 6.33). The discus flies along the *falling branch* of its orbit through the air like the airfoil of an aircraft with a fixed angle of attack α. It experiences a lifting force just like a wing (Sect. 10.10). It sinks to the ground more slowly than a stone would, and therefore flies further than the free-fall parabola would predict (dotted curve in the figure). – Naturally, the word "force-free" can be considered in this case to apply only as an approximation. The airflow around the discus produces not only a lifting force, but also a small torque on the spinning discus; both are neglected here.
b) The diabolo top as shown in Fig. 6.34: Even when thrown to a considerable height, its figure axis maintains a fixed direction in space.

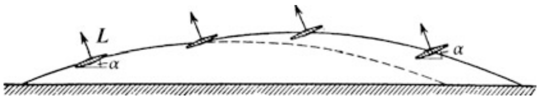


Figure 6.33 The flight path of a discus

Figure 6.34 A diabolo top

6.11 Tops Acted on by Torques. Precession of the Angular-Momentum Axis

After introducing the momentum $\boldsymbol{p} = m\boldsymbol{u}$, we were able to cast the equation of (progressive) motion in the following form:

$$\boldsymbol{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{u}) = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}. \qquad (5.41)$$

Furthermore, for progressive motions, we had to distinguish between two limiting cases. In the first case, the direction of the force $\boldsymbol{F}$ was parallel to the momentum $\boldsymbol{p}$ already present: Only the *magnitude*, but not the direction of the momentum was changed by the force (linear path). – In the second limiting case, the direction of the force was perpendicular to the previously-present momentum at every instant: Only the *direction* of the momentum changed (circular orbit).

In a corresponding manner, we will now treat the effect of a torque $\boldsymbol{M}$ on a *spinning top*. We thus write the equation of motion for rotation (6.7) in the form (cf. row 11 in Table 6.1):

$$\boldsymbol{M} = \frac{\mathrm{d}}{\mathrm{d}t}(\Theta\boldsymbol{\omega}) = \frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} \qquad (6.15)$$

and can again distinguish two limiting cases. In the first, the direction of the torque vector is *parallel* to the direction of the angular momentum axis: Then the top experiences an angular acceleration $\dot{\boldsymbol{\omega}}$. Only the *magnitude* of its angular momentum $\boldsymbol{L}$ is changed, but not its direction.

An arrangement which can be used for measurements is seen in Fig. 4.4 (MAXWELL's wheel). The effective torque M is equal to F_G times the radius of the shaft of the wheel.

In the second limiting case, the vector of the torque $\boldsymbol{M}$ is *perpendicular* to the direction of the angular momentum $\boldsymbol{L}$ already present in the top's motion. Then the magnitude of the angular momentum remains constant, only its *direction* changes.

A *torque* which acts perpendicular to the angular momentum axis gives rise to a *precessional motion of the angular momentum axis*. The angular momentum axis is no longer fixed in space. *It begins itself to rotate along a precession cone which is fixed in space.* The angular momentum axis however remains the center line of the nutation cone. The spinning top is now characterized by *three* angular frequencies or velocities:

1. The angular velocity $\boldsymbol{\omega}$ around its figure axis (spinning);

2. the angular velocity $\boldsymbol{\omega}_{\mathrm{N}}$ of the figure axis around the spin axis on the *nutation cone*; and

3. the angular velocity $\boldsymbol{\omega}_{\mathrm{p}}$ of the angular momentum axis in its motion around the fixed *precession cone*.

The motions of a top with simultaneous nutation and precession are rather complicated. For purposes of demonstration, one should therefore separate nutation and precession as clearly as possible. For this purpose, we usually start with a top which is *free of nutation*, as far as possible. We thus choose a top whose axes of symmetry and of angular momentum practically coincide – an exceptional case.

The top shown in Fig. 6.29 suffices for this purpose. One need only shift its center of gravity above or below the fulcrum point by sliding the weight A. However, it is clearer to use the arrangement shown in Fig. 6.35. It consists of a top with a horizontal axis ($\alpha = 90°$), a *gyroscope*. The support of the gyroscope is at the center of gravity of the whole system, in the form of a point and socket. In order to provide a torque $Fl \sin\alpha$ perpendicular to the angular momentum L, we hang a small weight on the shaft of the gyroscope . This torque has its largest magnitude, $M = Fl$, for $\alpha = 90°$. It has two effects: First, a small nutation; and second, a very noticeable *precession*. The angular momentum axis circles around a (here very squat) *precession cone* whose central axis is vertical.

The weak nutations can be ignored; we explain first how the precession comes about: The constant torque $\boldsymbol{M}$ produces in every time interval dt an additional component of angular momentum d$\boldsymbol{L}$ (Fig. 6.35). This component is perpendicular to the original angular momentum $\boldsymbol{L}$ and combines with it to give a resultant angular momentum in the direction R. The angular momentum axis moves through an angle dβ in the time dt, within the plane defined by the vectors $\boldsymbol{M}$ and $\boldsymbol{L}$. We find:

$$\boldsymbol{M} = \frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}, \tag{6.15}$$

and, from Fig. 6.35, $\mathrm{d}L = L\mathrm{d}\beta$. We thus obtain

$$M = L\frac{\mathrm{d}\beta}{\mathrm{d}t}, \quad M = \omega_{\mathrm{p}} L, \quad \omega_{\mathrm{p}} = \frac{M}{\Theta\omega}$$

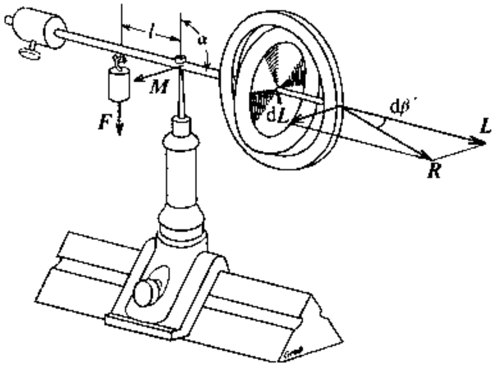


Figure 6.35 Precession of a spinning top (gyroscope) under the action of a constant torque (**Videos 6.10** and **6.11**) (see **Exercises 6.10** and **6.11**)

or, taking the directions into account by using vector notation,

$$\boldsymbol{M} = \boldsymbol{\omega}_{\mathrm{p}} \times \boldsymbol{L}\,. \tag{6.16}$$

These equations are confirmed by experiment: increasing the torque in Fig. 6.35 (heavier weight) increases the angular velocity ω_{p} of the precession.

When $\alpha < 90°$, the torque $Fl \sin\alpha = M \sin\alpha$ acts on the horizontal component $\Theta\omega \sin\alpha$ of the angular momentum. It follows that the angular velocity of the precession $\boldsymbol{\omega}_{\mathrm{p}}$ is independent of the angle α !

This primitive demonstration of precession has left nutation out of consideration, as we emphasized. It is, however, sufficient for understanding many practical applications of precession. We give three examples here:

1. *Riding a bicycle "no hands"*. Figure 6.36 shows the front wheel of a bicycle. The rider leans a bit to the right. This produces a torque on the axle of the front wheel around the horizontal direction of motion, B. At the same time, the front wheel, acting as a top, begins to precess around the vertical axis C and leads into a right-hand curve. The line connecting the ground contact points of the front and rear wheels is again under the center of gravity of the rider; the point of support is again under the center of gravity. – The signs of all the rotations and angular momenta are shown in Fig. 6.36.

A demonstration experiment which is especially graphic is provided by a small model bicycle. We give its wheels a spin with a high angular velocity by pressing them briefly against a rapidly rotating disk (Fig. 6.37) and then hold the long axis of the model suspended horizontally. A slight, cautious tilting to the right around this horizontal axis immediately brings the bicycle into position for a right-hand curve. If we set it on the ground, the bicycle moves off reliably on a straight-line path.

Video 6.10:
"Precession of a spinning top"
http://tiny.cc/wtqujy

Video 6.11:
"Precession of a rotating wheel"
http://tiny.cc/otqujy
An analogous experiment with a bicycle wheel hung from a cord is shown. A simple variant of this experiment is a toy top which dances around without falling over.

Part I

Figure 6.36 Riding a bicycle "no hands"

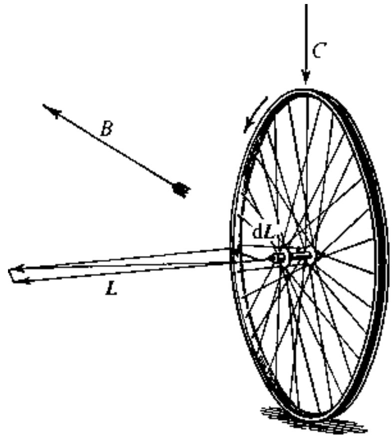
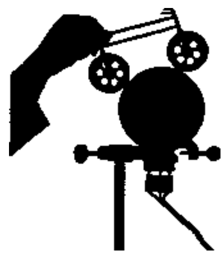


Figure 6.37 A model bicycle is "spun up" by pressing its wheels onto a rapidly rotating disk on the shaft of an electric motor **(Video 6.12)**



Video 6.12:
"Physics of riding a bicycle with no hands"
http://tiny.cc/vsqujy

The rider is completely dispensable. His or her role in riding the bicycle is very modest: he or she has simply to learn not to interfere with the automatic precessional motions of the front wheel[C6.19]. – The hoops which were once popular toys clearly make use of the same physical phenomena.

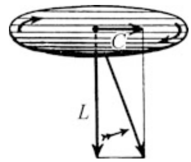
C6.19. The details of the physics of bicycle riding are in fact somewhat more complicated; see e.g. F. Klein and A. Sommerfeld: "*On the Theory of the Top*", German original published by Teubner 1910, p. 863. For the influence of the structure of the bicycle, see e.g. J.D.G. Kooijman *et al.*, Science **332**, 339 (2011).

2. *The beer coaster as a discus.* Throw a beer coaster with your right hand in a horizontal direction, tilting it slightly upwards. Then it will at first fly like a good discus as an "airfoil" (Fig. 6.33). Soon, however, the angle of attack will increase; the flight path, initially flat with a slight climb, will curve up steeply. At the same time, the coaster rises up on the right side, flies to the left and loses all its speed. From its maximum height, it will fall abruptly to the ground.

Explanation: The angular momentum of the beer coaster is much smaller than that of the heavier discus with its large moment of inertia. The torque produced by the airflow around the disk causes a strong precession of the figure axis, causing the angle of attack to increase and twist.

A non-rotating disk would tip up its leading edge (cf. Fig. 10.16). The airflow gives the disk an angular momentum along the *C*-axis, perpendicular to its flight path (Fig. 6.38). In the case of the rotating beer coaster, an angular *L* momentum is already present. The two angular momenta add vectorially, and the figure axis of the beer coaster carries out a precessional motion as indicated by the curved, feathered arrow in the figure.

Figure 6.38 A beer coaster as a discus, thrown flat with the right hand



3. The *boomerang*. The moment of inertia of the beer coaster can be increased and the unwanted precession thereby reduced without changing its overall weight and lifting force. We need only to thicken the rim of the coaster while thinning (or cutting out) the middle part.

Take a cardboard ring of about 20 cm diameter and 4×20 mm profile, and glue a sheet of writing paper onto its surface.

A 'prepared' beer coaster of this type with a larger moment of inertia experiences only a small precession, according to Eq. (6.16). It will also climb with an increasing angle of attack and lose speed, but at the top of its flight curve, it still has a reasonable angle of attack. This allows it to glide, still rotating, back to where it was thrown. It exhibits the typical behavior of the sport article known as a *boomerang*. *The usual curved-hook shape of the boomerang is not at all essential for its return flight.*

However, a circular disk is not a good airfoil. An elongated oval disk with a slight curvature is a much better airfoil and a fairly good boomerang (and at the same time a non-rotationally-symmetric top). For demonstrations, one can take a piece of cardboard of about 5×12 cm in size and 0.5 mm thickness, and fold it slightly in the middle along its long axis.

Small boomerangs are best not thrown free hand; they can be laid on a tilted book with an overhang, and struck with a rod held parallel to the edge of the book. A slight tilting of this 'flight ramp' to the side produces a flight path that curves to the left or the right as desired, or with the return path in the same vertical plane as the outward flight. We can make the projectile swing back and forth around the vertical from its starting point, etc. Using the curved hook form and twisting the two blades like those of a propeller, we can produce still more exotic flight paths (e.g. a "corkscrew" flight) and display a number of amusing tricks.

6.12 Precession Cone with Nutation

Under suitable experimental conditions, the precession of the angular momentum axis of a top produced by the action of an external torque can lead to a well-formed precession cone. We give two examples:

1. *The pendulum-top*. A top is hung in a stable position, but so that it can be turned in any direction (around a "universal joint") as shown in Fig. 6.39. It is made from a bicycle wheel (possibly with a lead-weighted rim). When it is not hanging straight down, a torque $\boldsymbol{M}$

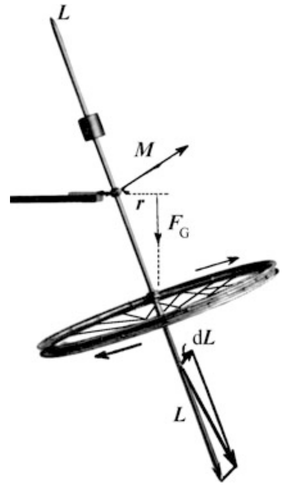


Figure 6.39 A top hung as a pendulum with three degrees of freedom (at its upper end, a small light bulb L allows it to be tracked in the photos shown in Fig. 6.40)

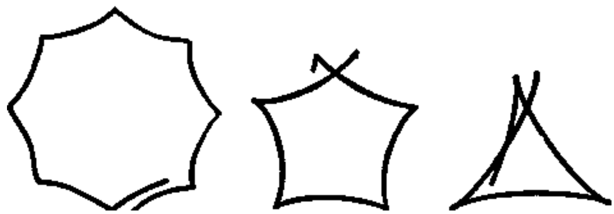


Figure 6.40 *Left*: The weak nutation of a suspended top, the approach to pseudo-regular precession; *center and right*: increasing nutation as the angular momentum of the top decreases (photo negatives)

acts, due to its weight $\boldsymbol{F}_G$ which pulls on the lever arm $\boldsymbol{r}$. The torque vector is shown in the figure, as is the additional angular momentum d$\boldsymbol{L}$ produced by the torque. In the position shown, the top begins to rotate along a well-defined precession cone with a small angular velocity. At the same time, it exhibits a weak nutation: The lower end of the top's figure axis (the axle of the wheel) does not move along a smooth circle, but rather along a circle with a wavy contour (Fig. 6.40, left). The greater the angular momentum of the top, the weaker is this nutation. The nutation can be made practically unnoticeable. In that case, one refers to the precession as *pseudo-regular*. The opposite of a pseudo-regular precession is true regular precession. In that case, the weak nutation caused by the external torque is suppressed. This is achieved through certain initial conditions. At the moment when it is released, one gives the top a nutation which is opposite but equal to the nutation which would be caused by the external torque; the top is given a small 'kick' when released. This kick has to be in the direction of the vector d$\boldsymbol{L}$. Its required strength can be found by trial and error (compare **Video 6.11**). A calculation would be out of place here.

Instead, we want to show how the nutation becomes stronger and stronger as the angular momentum of the top decreases, i.e. as the angular velocity around its figure axis (spinning) 'runs down'. The tip of the figure axis (axle of the bicycle wheel) follows the orbits that are photographed in Fig. 6.40. – With suitable initial conditions, the precession can even be completely suppressed. Then in spite of the external torque, only nutations survive. Again, going into detail here would take us too far afield.

2. *The earth as a top.* A famous example of precessional motion is provided by the earth itself. The earth is not a perfect sphere, but rather is somewhat flattened at the poles (and, precisely speaking, slightly pear-shaped). Its diameter at the equator is about 0.34 % greater than the figure axis of the earth, i.e. the line connecting the north and south poles. One can imagine that a waistline bulge has been added to a strictly spherical earth. The gravitational attraction of this bulge by the sun and the moon produces an external torque on the earth-top. Its figure axis moves along a precession cone with an opening half-angle of $23\frac{1}{2}^\circ$. A complete rotation around this cone takes about 26 000 years. At the same time, the torque produces extremely weak nutations. As a result, the axis of rotation deviates at each moment slightly from the earth's figure axis. But the points at which the two axes pass through the surface of the earth are only about 10 m apart.

6.13 A Top with Only Two Degrees of Freedom[2]

Torques $\boldsymbol{M}$ acting perpendicular to an angular momentum axis change the direction of the angular momentum vector (precession, Eq. (6.16)). Conversely, a change in direction (rotation) of an angular momentum produces torques $\boldsymbol{M}_\mathrm{p}$ perpendicular to the angular

[2] A *degree of freedom* refers to a spatial dimension in which the motion of a body can take place. Examples: A point-like body ("point mass") can in general experience a linear motion in any arbitrary direction in space. Its velocity can be decomposed into three components in a Cartesian coordinate system. The point mass thus has three degrees of freedom. – A point mass constrained to remain in a plane has only two degrees of freedom, and a point mass held to a linear track has only one. – A body of finite extension can experience rotations as well as progressive motions. Its angular velocity can in the most general case point in an arbitrary direction in space; it can then be decomposed into three perpendicular components. In addition to the three degrees of freedom of its progressive motions (translation), it has three degrees of freedom of rotation. If its axis of rotation is confined to a plane, only two degrees of freedom of rotation are present. A flywheel held by bearings has only a single rotational degree of freedom. – A body which is moving linearly and rotating can in addition exhibit oscillations of its individual parts relative to each other. A dumbbell-shaped object for example can oscillate along the axis (spring) connecting the two weights, and can simultaneously move progressively and rotate. Then in addition to the six degrees of freedom of translation and rotation, there is a seventh for the oscillation, etc.

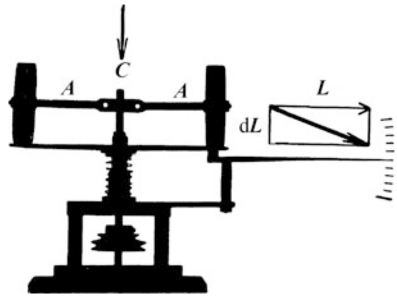


Figure 6.41 Demonstration model of a pan mill. The arrow over C indicates the angular velocity $\boldsymbol{\omega}_\mathrm{p}$ of the forced precession, and d$\boldsymbol{L}$ is the additional angular momentum that it produces within the time dt. Without the constraint provided by the grinding table, the axis A would adjust to the direction of the thick arrow. Thus, a torque which points away from the viewer and is perpendicular to the plane of the page must arise.

momentum axis and to the direction in which the axis is rotated. $\boldsymbol{M}_\mathrm{p}$ and $\boldsymbol{M}$ differ only in sign, so that we find

$$\boldsymbol{M}_\mathrm{p} = \boldsymbol{L} \times \boldsymbol{\omega}_\mathrm{p} \,. \tag{6.17}$$

The torques produced by forced precession play a major role in technology. As a first example of a top which is constrained to a plane, we mention the pan mill, a form of grain mill which was already known to the Romans (Fig. 6.41). During a rotation of the millstones, they form a top with a forced precession. The resulting torque is in this case in the *same* direction as the torque produced by the weight. It presses the millstones onto the grinding table and thus increases the grinding pressure. In a model, this can be made visible using a spring below the grinding table with a pointer attached. The following example is also impressive.

Figure 6.42 shows a gymnastic bar mounted in ball bearings KK. It carries a motorized top (gyroscope) and a seat. The gyroscope can swing along the axis of the bar within a U-shaped frame (in the plane of the page). The bearings of its frame are marked with a small white circle in the figure, and the outer frame is fixed to the bar (at 'R' in Fig. 6.42). A man sits down on the seat. The center of gravity of the whole system (bar, gyroscope, man) is well above the bar, so that the system is unstable (labile). It could for example tip to the man's right. This tipping produces a torque on the axis of the gyroscope; it answers with a precession. Assume that it is rotating in a counter-clockwise direction as seen from above. In this case, the upper end of the gyroscope moves away from the man. Now we come to the essential point: The man pushes the upper end of the gyroscope still further away from himself. He feels practically no greater resistance than with a gyroscope at rest. Nevertheless, due to this forced precession, a strong torque is produced. It acts on the

Figure 6.42 Stabilization by negative-damped precessional gyroscope oscillations (monorail). There is a protective metal fender between the person and the gyroscope, and at the right on the gyroscope frame R is a balance weight. **(Video 6.13)**

Video 6.13:
"Stabilization using a spinning top"
http://tiny.cc/5tqujy
The U-shaped frame R attached by bearings to the bar can be clearly seen at the beginning of the video. In the picture: the author.

bearings of the gyroscope and thus on the bar. The bar moves back to its original position. If it were initially tipped to the left, everything occurs similarly but with reversed signs of rotation. The upper axis of the gyroscope approaches the man; he pulls it a bit closer, etc. In this manner, one can balance with very little effort. The gyroscope swings with small amplitudes within the plane defined by its bearings. The man need only provide a *negative damping* or *amplification* of these precessional oscillations. That is, he must increase the amplitude already present in each case.

Our brains can learn how to perform this "negative damping" as a pure reflex in a surprisingly short time. When the dimensions of the gyroscope are suitably chosen, there is no time to think before reacting. But our nerve and muscle coordination very quickly registers the physical situation. After a few minutes, one feels just as secure on this unbalanced gymnastic bar as an experienced rider on his bicycle.

Chinese tightrope artists have long since empirically discovered this helpful device of negative-damped gyroscope oscillations. They use an umbrella which is rapidly twirled between the fingers as the gyroscope; they hold the shaft of the umbrella nearly parallel to the rope and balance by tipping its axis slightly. – For the most part, however, they use the umbrellas without spinning as parachutes.

Tops or gyroscopes with only one degree of freedom can be most effectively understood by using the methods of the following chapter.

Exercises

6.1 An object of mass $m_1 = 15\,\mathrm{kg}$ is at the end of a uniformly-shaped beam with a mass of 5 kg. a) The beam is suspended from a rotatable bearing at a distance of 1/5 of its length from the object m_1. What mass m_2 be added at its other end to keep it in equilibrium? b) Where must the bearing be located if a mass $m_2 = 25\,\mathrm{kg}$ is hung from the other end of the beam and it is to remain in equilibrium? What is then the total force F acting on the bearing? (Sect. 6.2)

6.2 Two full cylinders with radii r_1 and $r_2 > r_1$ and densities ϱ_1 and $\varrho_2 > \varrho_1$ are rolled down a ramp starting from the same initial conditions; the ramp has an inclination angle α. Which of the two cylinders will arrive first at the bottom of the ramp? (Sect. 6.4)

6.3 How much work W must be performed in order to set a brass sphere 1 m in diameter in rotation at 20 revolutions/min? (The density of brass is $8.5\,\mathrm{g/cm^3}$.) (Sect. 6.4)

6.4 A homogeneous thin rod of length a and mass m is oscillating around its suspension point at one of its ends. Find its reduced pendulum length l. (Sects. 6.4 and 6.5)

6.5 At what point along its length would the rod in the previous exercise have to be suspended a) to obtain the minimum oscillation period; and b) so that the oscillation period would be the same as when it is hung from one end? (Sects. 6.4 and 6.5)

6.6 A physical pendulum of mass $m = 0.5\,\mathrm{kg}$ is vibrating around a horizontal axis with a small amplitude under the influence of the earth's gravity; the axis is located 20 cm from its center of gravity (see Fig. 6.15). Its reduced pendulum length is $l = 50\,\mathrm{cm}$. Calculate its moment of inertia Θ_0 around the axis and Θ_S around a parallel axis which passes through its center of gravity. (Sect. 6.5)

6.7 Figure 6.16 shows the schematic of a beam balance as a physical pendulum. a) Calculate the moment of inertia Θ of the balance beam assuming that its oscillation period increases from $T_0 = 18\,\mathrm{s}$ to $T = 24\,\mathrm{s}$ when 100 g of additional weights are added to each balance pan, as described in the text. (The points from which the balance pans are hung lie on a straight line with the fulcrum point O, each at a distance of $l = 15\,\mathrm{cm}$.) b) By what distance d would the suspension points have to be lowered so that the oscillation period with the additional 100 g weights would remain equal to T_0? (The acceleration of gravity is $g = 9.81\,\mathrm{m/s^2}$.) (Sect. 6.5)

6.8 A person of mass m is standing on a symmetric, rotatable plate whose axis of rotation is vertical and passes through the center point of the plate. It has a moment of inertia Θ and is at rest. The person first moves from the center radially outwards by a distance r, then walks around a full circle of radius r on the plate and finally returns along a radial path to the center. Through what angle φ has the plate turned? (Sect. 6.6; see also Exercise 4.3)

6.9 A pencil of length $l = 20\,\mathrm{cm}$ falls freely in a horizontal position. After it has fallen a distance $h = 1\,\mathrm{m}$, one end strikes the edge of a table. How does it continue to fall and what is its angular velocity? (Sect. 6.6)

6.10 Referring to Video 6.10, "Precession of a spinning top" (Fig. 6.35): Determine the angular velocities of precession with the small torque and the torque which is three times larger (produced by attaching the small (15 g) and the large (45 g) weights, respectively), and explain the cause of the different angular velocities. (Sect. 6.11)

6.11 A spinning gyroscope whose rotation axis is horizontal is suspended on a sharp point at its center of gravity (Fig. 6.35). Under the influence of a torque $\boldsymbol{M}$, produced by a weight hung on the axle, the gyroscope precesses at an angular velocity $\boldsymbol{\omega}_P$. If we consider this experiment in a frame of reference which is rotating at the same angular velocity, adding the weight in this frame of reference also produces the torque $\boldsymbol{M}$, but the axle does not tip downwards. Why? A qualitative answer will suffice. (Sects. 6.11 and 7.3)

For Sect. 6.6, see also Exercise 4.3.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_6) contains supplementary material, which is available to authorized users.

Accelerated Frames of Reference

7

Part I

7.1 Preliminary Remarks. Inertial Forces

Up to now, we have considered physical processes from the point of view of the fixed surface of the earth or the floor of our lecture hall. Our frame of reference was the earth, treated as rigid and unmoving. The occasional exceptions were clearly indicated.

Making the transition to a different frame of reference can be trivial in some special cases. In these cases, the new frame of reference is moving at a *constant* velocity relative to the earth. Its velocity must not change, neither in magnitude nor in direction. Experimentally, we can sometimes meet this condition in a vehicle which is moving very "smoothly", e.g. on a ship or in a railroad car. In such cases, in the interior of the vehicle, we cannot "feel" the motion of our frame of reference. All phenomena take place within the vehicle just as they would in the lecture hall which is at rest. But these are rare and exceptional situations[C7.1].

C7.1. We will deal extensively with such cases when we come to the Special Theory of Relativity (in Vol. 2).

In general, vehicles of all types are *accelerated* frames of reference: Their velocity is changing in magnitude and direction. This acceleration of the frame of reference leads to noticeable modifications in our physical observations. The point of view of an observer within an accelerated frame requires new concepts to allow us to give a straightforward account of the observed physical phenomena. New forces appear to an accelerated observer. Their collective name is "inertial forces"[1]. Some of them also have been given special names (centrifugal force, CORIOLIS force). The description of these inertial forces is the subject of this chapter.

In our treatment of acceleration, we have consistently distinguished between two limiting cases: a pure *path* acceleration and a purely *radial* acceleration, i.e. a change of the velocity only in terms of its *magnitude* or only in terms of its *direction*. In an analogous fashion, we now want to treat frames of reference with a pure path acceleration, and separately frames of reference with a purely radial acceleration, as two limiting cases.

Frames of reference with a pure *path acceleration* often occur. Vehicles of all kinds when they *start moving* or *brake* along a straight path

[1] This choice of name presupposes that the observer is aware of his own acceleration. A less dramatic name or a neologism which corresponds to the word "weight" would be more appropriate.

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_7

are examples. But such accelerations usually last only a short time and the magnitude of the acceleration is constant at most for a few seconds. We can treat this limiting case rather briefly. We do this in Sect. 7.2.

A frame of reference with a purely *radial acceleration* is quite a different matter. Every carousel which is turning with a constant angular velocity $\boldsymbol{\omega}$ maintains its radial acceleration constant for an arbitrarily long time. The earth, in particular, is just such a carousel. Thus, we will need to study the *carousel system* rather thoroughly. We do this in the remaining sections of this chapter.

To make the treatment clearer, we will make use of a special device in the following: In each case we will first treat an observer "at rest" in the rest frame of the earth or the floor of the lecture hall (boldface heading!). Then, we describe the same phenomena from the point of view of an accelerated observer, moving with the accelerated frame. *Both observers start with the equation of motion* $\boldsymbol{a} = \boldsymbol{F}/m$ and consider forces to be the causes of observed accelerations.

7.2 Frames of Reference with Only Path Acceleration

First, some examples:

1. An observer is sitting at rest on a moving cart, in front of a frictionless table top on which a ball of mass m is lying (Fig. 7.1). The table top compensates the weight of the ball. The table and the chair are attached rigidly to the cart. The cart is being accelerated to the left along its long axis (by a kick!). This causes the ball and the man on the cart to approach each other. – Now we find two distinct ways of describing this process:

Observer at rest

The ball remains at rest. No force is acting on it, since it lies on the table without friction. In contrast, the cart and the man sitting on it are accelerated to the left. The man approaches the resting ball.

Figure 7.1 Experiments in an accelerated frame of reference

Figure 7.2 Measurement of the inertial force

Accelerated observer

The ball is accelerated to the right. Thus, a force $\boldsymbol{F} = -m\boldsymbol{a}$ directed to the right must be acting on the ball. This force is called the *inertial force*[C7.2].

C7.2. In the case of the inertial force $\boldsymbol{F} = -m\boldsymbol{a}$, $\boldsymbol{a}$ is the acceleration of the *frame of reference*.

2. The observer on the cart is holding the ball using a force meter (Fig. 7.2). The cart is again accelerated to the left. During the acceleration, the observer on the cart feels a force on his hand and arm muscles. The force meter indicates the magnitude of the force $\boldsymbol{F}$.

Observer at rest

The ball is accelerated to the left. A force $\boldsymbol{F}$ directed towards the left acts on it. For the acceleration, we have $\boldsymbol{a} = \boldsymbol{F}/m$.

Accelerated observer

The ball remains at rest. It is not accelerated. Then the sum of the two forces acting on it is zero. The inertial force $\boldsymbol{F} = -m\boldsymbol{a}$, acting to the right, and the muscular force acting to the left are equal and opposite. Their magnitude can be read off the force meter.

3. In Fig. 7.3, a cart is being accelerated to the left. The observer standing on the cart has to lean into the acceleration while the cart is starting up; otherwise, he would fall over. – Again we give two distinct descriptions:

Observer at rest

The center of gravity S of the man must move with the same *acceleration* (magnitude and direction) as the cart. The force $\boldsymbol{F}$ (Fig. 7.4) required to accelerate the center of gravity, acting to the left, is produced by the man through his weight $\boldsymbol{F}_G$ and an elastic deformation of the cart (force $\boldsymbol{F}_3$). He leans forward to achieve this.

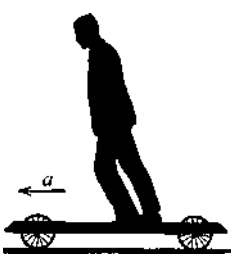


Figure 7.3 A man on a cart which is accelerating to the left

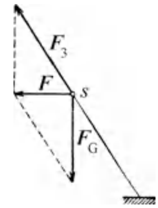


Figure 7.4 Forces as observed in the system at rest

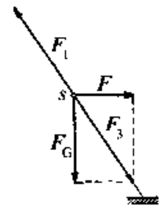


Figure 7.5 Forces in the accelerated system

Accelerated observer

The center of gravity S of the man remains at *rest*. The sum of the forces (Fig. 7.5) acting on him is zero. His weight $\boldsymbol{F}_{\mathrm{G}}$ pulls downwards, while the inertial force $\boldsymbol{F} = -m\boldsymbol{a}$ acts to the rear, that is to the right in the figure. The two combine to give the resultant force $\boldsymbol{F}_3$. It deforms the cart below the man's feet and thus produces the force $\boldsymbol{F}_1$ which is opposite but equal to $\boldsymbol{F}_3$.

4. An observer is in an elevator. In front of him is a table with a spring balance holding an object of mass m. The deformation of the spring indicates a force which is equal but opposite to the weight $\boldsymbol{F}_{\mathrm{G}}$. Then the elevator begins to accelerate *downwards*. The balance now indicates a smaller reading, $\boldsymbol{F}_1$.

Observer at rest

The object is accelerated downwards. Two unequal, oppositely-directed forces act on it. The weight $\boldsymbol{F}_{\mathrm{G}}$ pulls the object downwards, while the smaller force of the spring $\boldsymbol{F}_1$ pushes it upwards. The resultant difference force produces its downwards acceleration $\boldsymbol{a} = (\boldsymbol{F}_{\mathrm{G}} + \boldsymbol{F}_1)/m$, with the magnitude $a = (F_{\mathrm{G}} - F_1)/m$.

Accelerated observer

The object is at rest, the sum of the forces acting upon it is zero. The upwards-directed spring force $\boldsymbol{F}_1$ of the balance is smaller than the weight $\boldsymbol{F}_{\mathrm{G}}$ of the object. Therefore, a second upwardly-directed force is present, namely the inertial force $\boldsymbol{F} = -m\boldsymbol{a}$, so that $\boldsymbol{F}_{\mathrm{G}} + \boldsymbol{F}_1 + \boldsymbol{F} = 0$.

5. An observer jumps from a high table to the ground with the spring balance in his hand. On the balance pan is a weight of mass m. Immediately after he jumps, the indication of the balance goes from the value F_{G} to zero (Fig 7.6). The balance thus indicates no weight.

Figure 7.6 A freely-falling frame of reference **(Video 7.1)**

Video 7.1:
"Freely falling frame of reference"
http://tiny.cc/1uqujy
In the first part, S. KÖSTER (doctorate 2007) jumps from a table holding the balance; filmed by T. BECKER (doctorate 2004). The second part shows a similar jump by H. GRÜNDIG (1929–2003, doctorate in Göttingen, 1959), filmed in slow motion. This part was filmed in 1960.

Unfortunately, this demonstration is over within a fraction of a second. This disadvantage is not present in the four accelerations dealt with in Sect. 7.3[C7.3].

Observer at rest

The object is accelerated. It falls like the man with the acceleration of gravity $\boldsymbol{g}$. The only force acting on the object is its weight $\boldsymbol{F}_\mathrm{G}$, which is pulling it downwards. Neither a spring force nor a muscular force is pushing it upwards.

Accelerated observer

The object is at rest, the sum of the forces acting on it is zero. Its weight $\boldsymbol{F}_\mathrm{G}$ which is pulling it downwards is cancelled by the inertial force $\boldsymbol{F}$ acting upwards. Their magnitudes mg are equal. The object is "weightless".

These examples should suffice to elucidate the meaning of the words 'inertial force'. *Inertial forces exist only for an accelerated observer.* The observer must – at least in thought! – participate in the acceleration of his frame of reference. A hand which is accelerating a bowling ball is an accelerated frame of reference. Therefore, the hand feels an inertial force.

C7.3. POHL mentions here experiments with "zero gravity" in space stations. He also points out the microgravity experiments in the *drop tower at the University of Bremen*. In an evacuated tube, 120 m long, an experiment capsule (0.8 m in diameter, 1.6 to 2.4 m long) can fall freely for nearly 5 s before it is braked to a stop by a plastic foam cushion (see e.g. http://en.wikipedia.org/wiki/Fallturm_Bremen or https://www.zarm.uni-bremen.de/drop-tower.html).

7.3 Frames of Reference with Radial Acceleration. Centrifugal and CORIOLIS Forces

1. An observer is sitting on a rotating swivel chair with a vertical axis of rotation and a large moment of inertia (Fig. 7.7, cf. also Fig. 7.16). In front of him, attached to the chair, is a horizontal, smooth tabletop. The observer sitting in the chair places a ball on this tabletop. It flies off the table and falls to the ground.

Observer at rest

The ball is not accelerated. No forces are acting on it (its weight is compensated by the elastic resistance of the tabletop). As a result, C7.4. In the picture: The author.

Figure 7.7 Experimenting in a rotating frame of reference[C7.4]

it cannot move along a circular path. It flies off tangentially with the constant velocity $u = \omega r$ (ω = angular velocity of the rotating swivel chair, r = distance from the ball to the axis of rotation at the moment when it is placed on the tabletop).

Accelerated observer

The ball is accelerated from its resting position. It moves away from the center of rotation of the tabletop. Thus, an *inertial force* acts on the ball, which is initially *at rest*. This force is given the special name *'centrifugal force'*. Its formula is $\boldsymbol{F} = m\omega^2\boldsymbol{r}$.

2. The observer on the swivel chair holds a force meter between his hand and the ball. The horizontal axis of this force meter is directed towards the axis of rotation of the chair. The force meter indicates a force of magnitude $\boldsymbol{F} = m\omega^2\boldsymbol{r}$ while the chair is rotating.

Observer at rest

The ball follows a circular path of radius r (radius vector $\boldsymbol{r}$); it is *accelerated*. This requires the presence of a radial force $\boldsymbol{F} = -m\omega^2\boldsymbol{r}$, directed towards the axis of rotation, which acts on the ball ("centripetal force"), Eq. (4.1).

Accelerated observer

The ball remains at rest. It is not accelerated. Therefore, the sum of the two forces acting on it is zero. The centrifugal force which is pulling radially outwards and the muscular force pulling radially inwards (centripetal force) are equal and opposite. The magnitude of these forces is $m\omega^2 r$.

3. The observer on his swivel chair hangs a gravity pendulum on a stand on the table in front of him, for example a ball hung from a cord. This pendulum will not hang vertically (Fig. 7.8). It swings outward through an angle α within the plane defined by the radius and the axis of rotation. The angle α increases with increasing rotational frequency of the chair.

Observer at rest

The ball of the pendulum is moving along a circular orbit of radius r; *it is accelerated*. This requires the force $\boldsymbol{F} = -m\omega^2\boldsymbol{r}$ which acts

Figure 7.8 The effect of centrifugal force on a gravity pendulum

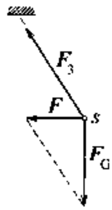


Figure 7.9 Forces in the system at rest

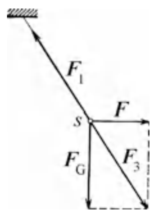


Figure 7.10 Forces in the rotating system

horizontally towards the axis of rotation (Fig. 7.9). It is produced by the weight of the ball $\boldsymbol{F}_\mathrm{G}$ and the elastic tension of the cord (force $\boldsymbol{F}_3$).

Accelerated observer

The ball of the pendulum is at rest, the sum of the forces which act at its center of gravity S is zero (Fig. 7.10). The weight of the ball pulls downwards with the force $\boldsymbol{F}_\mathrm{G}$, the centrifugal force $\boldsymbol{F} = m\omega^2\boldsymbol{r}$ pulls outwards. The two combine to give a resultant force $\boldsymbol{F}_3$. This force applies a tension to the cord and thereby produces the force $\boldsymbol{F}_1$, which is equal and opposite to $\boldsymbol{F}_3$.

4. An artificial satellite (a space station) circles the earth at an altitude of several 100 km (Fig. 4.25). Inside it, all the phenomena which normally result from the force that we call 'weight' are absent. Here (in contrast to the falling path acceleration in Fig. 7.6), we have as much time as we wish to observe the lack of such forces. If, for example, we put a metal block onto a spring balance, the balance will indicate no weight.

Observer at rest

Like the satellite itself, all objects within it are *accelerated* continually towards the center of the earth, i.e. in the direction of the radius of the satellite's orbit. Otherwise, they could not take part in the circular orbital motion. The weight $\boldsymbol{F}_\mathrm{G} = m\boldsymbol{g}$ of the metal block produces this acceleration. It is the only force which acts on the block and need not be compensated by the oppositely-directed force of a compressed spring.

Accelerated observer

The metal block is *at rest*. As a result, the sum of the forces acting on the block is zero. The force that we call weight, which pulls the block towards the center of the earth, is compensated by the centrifugal force which pulls it away from the earth's center (the center of the satellite's circular orbit). The two forces are equal and opposite; they have the same magnitude mg. (g at an altitude of 300 km $= 8.9\,\mathrm{m/s^2}$.)

5. In all the experiments up to now, we have observed a body which was *at rest* in the rotating system. We simply asked the question as to whether the body would be accelerated out of this resting position or not. Now, we observe a body which is *moving* relative to the rotating swivel chair. Here, we restrict our considerations to a limiting case, namely a body with a high velocity, a bullet. Then we can neglect the centrifugal force which is relatively small.

On the tabletop attached to the chair, we set up a small gun pointing horizontally outwards (Fig. 7.11). The direction of its barrel can make an arbitrary angle α with the radius connecting it to the axis of rotation (Fig. 7.12). The gun points at a target located a distance A from its muzzle and is aimed at the point a on the target disk. The target rotates with the chair, held by rods at the distance A. First, we fire a bullet from the gun while the chair is at rest, and determine where it hits the target, i.e. the point a. Then the chair is set in rotation with the angular velocity ω. *The chair is supposed always to rotate in a counter-clockwise sense as seen from above*. Now, a second shot is fired. The point b where it hits the target is deflected to the right by a distance s relative to the targeted point, which has in the meantime moved to a'.

Numerical Example
The chair rotates once every 2 seconds. The velocity of the bullet is $u = 60\,\mathrm{m/s}$ (air gun). The distance to the target is $A = 1.2\,\mathrm{m}$, the deflection to the right $s = 0.075\,\mathrm{m} = 7.5\,\mathrm{cm}$ (cf. Fig. 7.12).

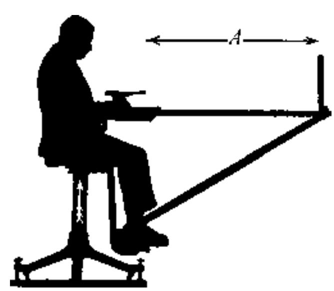


Figure 7.11 The CORIOLIS force[C7.5]

C7.5. In the picture: The author.

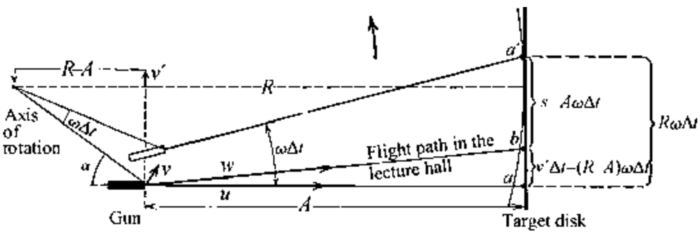


Figure 7.12 The experiment of Fig. 7.11 seen from above. $v' = v\cos\alpha = (R - A)\omega$ is the component of the bullet's velocity w which is parallel to the target disk, while v is the velocity of the muzzle of the gun. For clarity, the angle $\omega\Delta t$ is exaggerated in the drawing. This causes a minor error: the sight line from the gun to a' appears not to meet the target disk at right angles.

Observer at rest

When the chair is at rest, the bullet strikes the point a that it was aimed at. If the chair is stopped immediately after firing the bullet, the point b where it strikes the target is to the left of the targeted point. In this case, the velocity v of the muzzle of the gun adds to the velocity u of the bullet. As a result, the bullet flies in the direction w through the lecture hall.

In the experiment as in fact demonstrated, the chair continues to rotate following the firing of the bullet. The bullet, in contrast, flies on a force-free, linear path after leaving the muzzle of the gun, in the direction w through the lecture hall. *Therefore, the sighting line rotates relative to the path of the bullet.* At the end of the bullet's flight time Δt, the point aimed at is at a'. Thus the point b where the bullet strikes the target disk is deflected to the right by a distance s relative to the aiming point. From Fig. 7.12, we read off the relation

$$s = A\omega\Delta t.$$

For both flight paths (that is in the directions of u and w), the flight time of the bullet to the target is the same, namely

$$\Delta t = A/u.$$

Therefore, we find

$$s = u\omega(\Delta t)^2.$$

Accelerated observer

During the flight of the bullet, it is accelerated in a transverse direction. Its path is thus curved to the right. Within its flight time Δt, the bullet is deflected by the distance $s = \frac{1}{2}a(\Delta t)^2$ to the right. s, according to the result of the observer at rest, is $s = u\omega(\Delta t)^2$. From this, we find for the observed acceleration $a = 2u\omega$. It is named for its discoverer, CORIOLIS *acceleration*[C7.6]. There can be no acceleration

C7.6. For a simple derivation of the CORIOLIS acceleration, let the muzzle of the gun be on the axis of rotation and the moment of firing be $t_0 = 0$. Then the distance $s = R\,\omega\,t$ (Fig. 7.12). With $R = u\,t$, we have
$s = u\,\omega\,t^2$,
from which we obtain by differentiating twice with respect to time:
$a = \frac{d^2 s}{dt^2} = 2u\,\omega$.

Part I

a without a force $F = ma$; thus, the CORIOLIS force $F = 2mu\omega$ acts on the bullet transversely to its flight path, or, more generally,

$$\boldsymbol{F} = 2\, m\boldsymbol{u} \times \boldsymbol{\omega}\,. \tag{7.1}$$

The CORIOLIS force is an *inertial force* which acts on an object in *motion*. It acts perpendicularly to the vectors of the angular velocity and the path velocity. *Let a frame of reference be rotating, e.g. with the angular velocity ω. Within this system, a body is moving with an path velocity u perpendicular to the axis of rotation. Then the* CORIOLIS *force $F = 2mu\omega$* acts on the moving body perpendicular to its trajectory.

The equation $s = A\omega\Delta t = u\omega\Delta t^2 = A^2\omega/u$ which is accepted by both observers provides a very simple method for measuring the velocity u of a bullet.

6. The previous example demonstrates the transverse deflection of a *moving* body within an accelerated frame of reference, but only for a single initial direction of its motion. The magnitude of its deflection should be independent of the chosen initial direction (the aiming of the gun). But we intentionally did not demonstrate this; it can be shown much more quickly and simply by introducing a small change in the experiment: We replace the gun by the bob of a gravity pendulum. The pendulum is hung above the tabletop attached to the swivel chair in the usual manner (Fig. 7.13). To make it easier to observe, the moving pendulum bob itself is used to record its own path. To this end, a small ink pot is built into the pendulum bob, with a fine ink jet underneath. A sheet of white paper is pinned to the tabletop below the pendulum and the swivel chair is set in rotation with the angular velocity ω. The observer in the chair is initially holding the pendulum bob and plugging the ink jet. The cord of the pendulum is tipped out of its rest position within an arbitrary vertical plane. When the pendulum bob is released, it swings with a gradually decreasing amplitude around its *non*-vertical(!) rest position (Fig. 7.8). It draws a continuous curve on the paper with the ink jet, tracing out the rosette orbit shown at the left in Fig. 7.14. In a second experiment, the pendulum is kicked out of its rest position, giving the rosette orbit shown on the right in Fig. 7.14.

C7.7. In the picture: Master mechanic W. SPERBER

Video 7.2:
"Simple pendulum in a rotating reference frame"
http://tiny.cc/euqujy

Figure 7.13 A simple pendulum in a rotating frame of reference[C7.7] **(Video 7.2)**

Observer at rest

The pendulum swings around its rest position. In the case of Fig. 7.14, left, the orbit is an ellipse, since on releasing the pendulum bob from its deflected position, a tangential velocity component is also present. It swings with an "elliptical polarization" on an orbit fixed in space. The plane of the paper rotates under the swinging pendulum, producing the rosette, whose center remains free of ink dots.

In the second case (Fig. 7.14, right), there is no tangential velocity component at the moment the pendulum starts to swing. It swings with a "linear polarization" continually parallel within a vertical plane fixed in space.

The fact that the rest position of the pendulum is not vertical was explained above under Point 3.

Accelerated observer

During the motion, the pendulum bob is deflected by a CORIOLIS force continually in a direction transverse to the plane in which it is swinging. All the individual curves of the rosette have the same form in spite of their different orientations in the rotating chair. Therefore, *the direction of motion within the accelerated system is unimportant for the magnitude of the* CORIOLIS *force*.

The *deflection* of the rest position of the pendulum from the vertical results from a centrifugal force (cf. Point 3!). *Both a* CORIOLIS *force and a centrifugal force act on a moving body in a rotating frame of reference.*

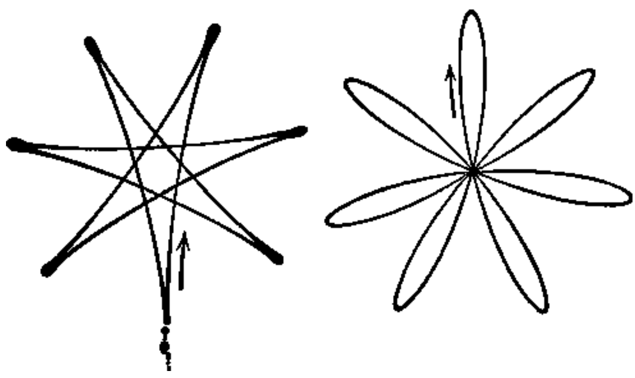


Figure 7.14 The rosette orbits of a pendulum on a carousel. In the left image, the pendulum was released from its maximum deflection above the ink dots at bottom, and it initially swung to the right (arrow). The endpoint of the rosette is coincidentally the same as its starting point. In the right image, the pendulum was given a kick out of its initial rest position **(Video 7.2)**.

(In this demonstration, the swivel chair was given only a small angular velocity ω. Otherwise, we could not have followed the position of the plane in which the pendulum bob is swinging.)

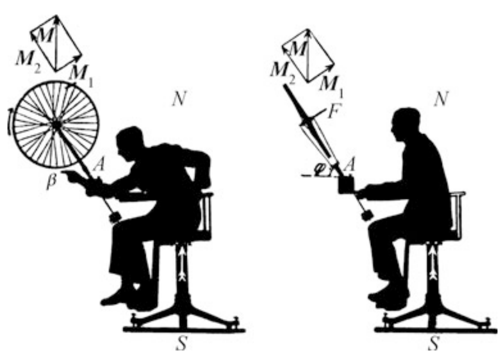
7. A spinning top in a rotating frame of reference (model of a gyrocompass on a rotating globe): Figure 7.15 shows a top mounted within a frame (gyroscope) on the swivel chair. For brevity, we will refer to the chair as the *globe*. Seen from above, it is assumed to rotate in a counter-clockwise sense. The frame of the top can itself rotate around an axis A which is perpendicular to the symmetry axis (figure axis) F of the top. The axis A lies in a *meridian plane* of the globe. Furthermore, the axis A can be adjusted to different latitudes. It can thus make an arbitrary angle φ ('latitude angle') with the plane of rotation of the globe, between 0° (equator) and 90° (poles). The horizon of the position of the top can be imagined to be perpendicular to the A axis. The observer in the rotating chair starts the top spinning by pulling a few times on its spokes (Fig. 7.15, left). Then he leaves the top to itself: The figure axis of the top takes up a fixed position in the meridian plane after a few initial swings around the axis A, like a compass needle (Fig. 7.15, right).

Now we let the two observers speak their piece:

Both observers, for simplicity, assume the same initial position of the figure axis F of the top: it lies parallel to a *circle of latitude*.

Observer at rest

The rotation around the chair's (or globe's) axis causes a torque $\boldsymbol{M}$ to act on the figure axis of the top. It has a component M_1 perpendicular to the A axis. This torque M_1 causes a precession of the top's figure axis F around the axis A of the frame. The figure axis F at first swings out of the meridian plane; but friction in the bearings of the A axis quickly damps these pendulum oscillations. The axis of the top comes to rest in the meridian plane. Then the M_1 component of the torque lies along the figure axis F, so that it can no longer produce further precessional motions.



Video 7.3: "Gyrocompass" http://tiny.cc/ququjy

Figure 7.15 Model of a gyrocompass (**Video 7.3**)

Accelerated observer

CORIOLIS forces deflect the parts of the top's rim at β to the right along their orbits. The half of the top on the right as seen by the reader moves out of the plane of the page towards the reader. This moves the figure axis of the top into the meridian plane. Thereafter, CORIOLIS forces indeed continue to act on the moving rim, but they no longer exert a torque around the A-axis.

So much for experiments to demonstrate the definitions of the concepts of *centrifugal force and* CORIOLIS *force. Both forces exist only for a radially-accelerated observer*. The observer must participate in the rotation of his frame of reference, at least in thought. He or she can still continue to use the equation $\boldsymbol{a} = \boldsymbol{F}/m$ within the rotating frame by taking these new forces into account.

The appearance or disappearance of *inertial forces* is thus determined by the choice of the frame of reference. For observers who move with the frame of reference, they are no less "real" than "genuine" forces are for the observer *at rest*. The term "fictitious forces" should therefore be avoided.

What is the situation for the accelerated observer with regard to the principle of "*actio* = *reactio*"? – Answer: He or she experiences the same situation as an observer on the earth regarding the reaction force to the *weight*. The observer in an accelerated frame of reference cannot detect the corresponding counter-forces to the inertial forces during the *free motion* of bodies within that frame.

7.4 Vehicles as Accelerated Frames of Reference

The choice between an accelerated and a non-accelerated frame of reference is simply a matter of taste in some cases, for example with circular motions of bodies around fixed axes. The important thing then is just to give a clear description of the frame of reference used (cf. Sect. 4.2, beginning). – In other cases, however, the accelerated frame of reference is preferable. These include most of the physics of the vehicles used in our technological society. The acceleration of these frames of reference is often rather complicated, since path accelerations (starting up and braking) and transverse accelerations (driving around curves) occur together.

Our everyday experience with inertial forces in vehicles was already treated in the examples given in Sects. 7.2 and 7.3. For example:

a) Leaning in a train when it is starting up or braking, and when it is travelling around curves. This prevents falling over.

b) Leaning into curves by bicycle and rider, rider and horse, aircraft and pilot.

c) The sideways deflection by the CORIOLIS force on the deck of a ship which is changing course. Only by tacking "sideways" can one walk a straight line in the intended direction.

d) One can "feel" the CORIOLIS forces especially clearly on a swivel chair with a large moment of inertia and therefore a steady, constant angular velocity. Try to move a weighted block (e.g. with 2 kg mass) *quickly* along an arbitrary straight-line path (Fig. 7.16). The result is startling. It feels as though your arm has gotten into a strong current within a viscous liquid. This is a particularly important experiment[C7.8]. (**Exercise 7.1**)

C7.8. This "particularly important experiment" can unfortunately not be shown as a demonstration, since only the experimenter him- or herself can "feel" the CORIOLIS force. We however want to point out the numerous possibilities of experiencing this phenomenon yourself on various carnival or amusement-park rides.

Numerical Example

Rotation rate: one revolution every 2 seconds, so that $\nu = 0.5\,\mathrm{s}^{-1}$, $\omega = 2\pi\nu = 3.14\,\mathrm{s}^{-1}$; mass of the metal block $m = 2\,\mathrm{kg}$; velocity $u = 2\,\mathrm{m/s}$; CORIOLIS force $= 2mu\omega = 2\cdot 2\,\mathrm{kg}\cdot 2\,\mathrm{m/s}\cdot 3.14\,\mathrm{s}^{-1} = 25\,\mathrm{kg\,m/s^2} = 25\,\mathrm{N}$; this force is greater than the weight of the moving metal block ($F_G = 2\,\mathrm{kg}\cdot 9.81\,\mathrm{m/s^2} \approx 20\,\mathrm{N}$)!

The list of qualitative examples is very long. More instructive than examining them all in detail, however, is a quantitative treatment of a particular case which at first seems weird. It involves a *horizontal torsion pendulum on a carousel*. Figure 7.17 shows a carousel in a side view. On it is a torsion pendulum with a rod-shaped mass which can be placed at any desired distance from the center axis of the carousel. Under what conditions will the torsion pendulum continually point its long axis towards the center of the carousel, independently of all accelerations?

At a constant angular velocity ω_1 of the carousel, the torsion pendulum remains at rest. This is because the purely *radial* acceleration

Figure 7.16 A swivel chair with a large moment of inertia for demonstrating CORIOLIS forces. The heavy weights hanging at the sides can also be conveniently used for the experiments shown in Figs. 7.7, 7.11, 7.13 and 7.15[C7.9]

C7.9. In the picture: Master mechanic W. NABEL.

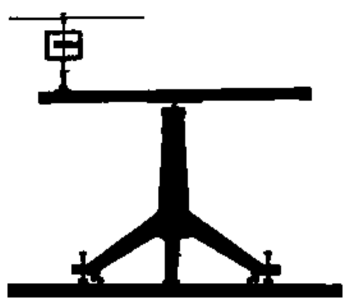


Figure 7.17 A torsion pendulum on a carousel (rotating table). The pendulum consists of a wooden rod (pendulum mass) on the small torsion axle that we have already seen in Fig. 6.5 **(Video 7.4)**

Video 7.4:
"Torsional pendulum on a rotating table"
http://tiny.cc/cuqujy

(centrifugal acceleration) of the rotating carousel points precisely along the *long axis* of the pendulum mass. Such an acceleration can never produce a torque.

To verify this, one could mount the pendulum on a rail so that it can be slid in a radial direction, orienting its long axis parallel to the rail. The pendulum will then not react to any accelerations along the rail.

However, any change in the angular velocity ω_1, that is any *angular* acceleration $\dot{\omega}_1$ of the carousel, will shift the pendulum out of its rest position. It will immediately begin to oscillate with a considerable amplitude, since now the acceleration a is *perpendicular* to the long axis of the pendulum mass. The task outlined above may at first seem hopeless; but on the contrary, it is in fact quite simple. We can make the torsion pendulum *completely insensitive* to *any angular acceleration* $\dot{\omega}_1$ simply through a *suitable choice of its moment of inertia* Θ_0! We require that (derivation follows immediately):

$$\Theta_0 = msR \tag{7.2}$$

or, using STEINER's law (Eq. (6.12)), we obtain an expression which is more convenient for calculations:

$$\Theta_\mathrm{S} = m(sR - s^2)\,. \tag{7.3}$$

Θ_0 = moment of inertia of the pendulum referred to its axis of rotation; Θ_S = moment of inertia referred to its center of gravity; m = mass of the pendulum (rod); s = distance from the pendulum's center of gravity to its axis of rotation; R = distance from the pendulum's axis to the axis of rotation of the carousel.

The torsion coefficient D^* of the spiral spring of the torsion pendulum is completely unimportant. It does not enter into the calculation at all. The silhouette (Fig. 7.17) shows a pendulum mass in the form of a rod which corresponds to this calculation (its dimensions are given below). This pendulum indeed remains at rest no matter how strong the angular acceleration of the carousel. This demonstration is quite surprising. Small variations of R or s restore the original sensitivity of the pendulum to angular accelerations.

To *derive* Eq. (7.3), we consider the situation from the viewpoint of a rest frame, see Fig. 7.18. When the carousel is accelerated, a force acts on the *axis of rotation O* of the torsion pendulum; it has a magnitude F (direction 1). We add to it two forces of the same magnitude which act oppositely at the *center of gravity S* (directions 2 and 3). The force in the direction 3 accelerates the center of gravity S. We have

$$F = ma = m(R - s)\,\dot{\omega}_1 \tag{7.4}$$

($\dot{\omega}_1$ = angular acceleration of the carousel).

At the same time, the forces in the directions 1 and 2 produce a *torque* $F \cdot s$. It causes a rotation around the *center of gravity S*, just as any

Part I

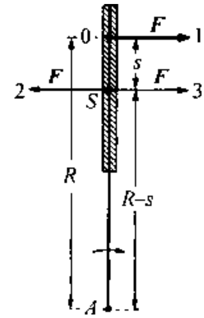


Figure 7.18 The insensitivity of a torsion pendulum towards angular accelerations of its center of rotation O

torque does when it acts on an otherwise free object. Quantitatively, we find

$$F \cdot s = \Theta_S \, \dot{\omega}_2 \tag{7.5}$$

($\dot{\omega}_2$ = angular acceleration of the object).

We require that $\dot{\omega}_1 = \dot{\omega}_2$. This condition can be fulfilled by a suitable choice of the distance s between the center of gravity of the pendulum and its axis of rotation. We combine Eqns. (7.4) and (7.5) and find a relation for s:

$$\frac{F}{m(R-s)} = \frac{F \cdot s}{\Theta_S} \quad \text{or} \quad \Theta_S = ms(R-s)\,. \tag{7.3}$$

For the pendulum body (rod) chosen for the demonstration experiment, with mass m and length l, we find

$$\Theta_S = \frac{1}{12} m l^2. \tag{6.11}$$

Inserting this value into Eq. (7.3) yields $l^2 = 12s(R-s)$. – *Numerical example* for Fig. 7.18: $R = 50\,\text{cm}$, $s = 5\,\text{cm}$, $l = 52\,\text{cm}$.

"This strange demonstration plays a role in transportation systems".

This strange demonstration plays a role in transportation systems (cf. Sect. 7.5).

7.5 The Gravity Pendulum as a Plumb Bob in Accelerated Vehicles

Navigation of an aircraft without visual contact to the ground demands a secure knowledge of the true vertical direction at all times, or of the horizontal plane perpendicular to it. Without this knowledge, a pilot lacking visual information can't even distinguish a curved flight path from a straight one; the tension of his muscles and his kinesthetic sense leave him clueless. They tell him only the direction of the resultant of weight and centrifugal force, but never the true vertical direction, which is identical to the local radius vector of the earth.

At rest on the ground, one can use a gravity pendulum as a plumb bob to find the vertical direction. In accelerated vehicles, this would at first seem useless; we have all watched a pendulum in an accelerated vehicle. Think of a strap hanging from the baggage rack in a railroad car. It swings back and forth, at the mercy of inertial forces. Nevertheless, one can in principle use a pendulum as a plumb bob even in arbitrarily accelerated vehicles! We can see this by first carrying out a thought experiment.

To this end, we modify the experiment described in Fig. 7.17 so that the center of rotation does not move on a fixed circle, but instead on the surface of a sphere of radius R. Every motion of the center point O thus takes place *momentarily* on a great circle, and in our thought experiment, we replace the circle from Fig. 7.17 *momentarily* by a great circle (with the torsion axis perpendicular to it). When Eq. (7.3) is fulfilled in this arrangement, the pendulum points continually towards the center of the sphere, even when the direction of the great circle changes in the course of the motion.

Now, in order to carry out this thought experiment in fact, we move the point O into the vehicle and attach a physical gravity pendulum as in Fig. 6.15 at this point; it is suspended at only a single point, so that it can swing in the horizontal plane. Every motion of the vehicle now takes place *momentarily* on a great circle whose radius R is the earth's radius ($= 6.4 \cdot 10^6$ m). The gravity pendulum can move freely in the plane of this great circle. To keep it at rest, we have only to ensure that Eq. (7.3) is fulfilled. Then the pendulum points continually towards the center of the earth, even when the vehicle drives or flies through a sharp curve! We have thus achieved our goal.

For a gravity pendulum, in contrast to an elastic torsion pendulum (mass-and-spiral-spring system), the moment of inertia is directly related to the torsion coefficient D^*. The choice of D^* is no longer free; it is determined by the weight of the pendulum mass, mg. According to Sect. 6.5, we have

$$D^* = mgs \quad (g = 9.81\,\mathrm{m/s^2})\,.$$

As a result, the period of oscillation of this pendulum, from Eq. (6.13), is

$$T = 2\pi\sqrt{\frac{\Theta_0}{D^*}} = 2\pi\sqrt{\frac{msR}{mgs}} = 2\pi\sqrt{\frac{R}{g}} = 84.6\,\mathrm{min}\,. \qquad (7.6)$$

It is called the SCHULER *period*, corresponding to a mathematical pendulum (Sect. 6.5) whose length is equal to the earth's radius R ! (**Exercise 7.5**). For a physical pendulum of dimensions suitable for installing in an airplane, that is of order of 0.1 m, a period of this magnitude would require a distance of $s = 1$ nm (Exercise 7.5), which is practically impossible to construct. But even physical pendulums which are excited to oscillations with the SCHULER frequency are independent of accelerations and can be used as plumb bobs. This is an important fact for "inertial navigation"[C7.10].

C7.10. See for example R. P. G. Collinson, "Introduction to Avionics Systems", Kluwer Academic Press, Boston, 2nd ed. (2003), Chaps. 5 and 6.

7.6 Earth as an Accelerated Frame of Reference. Centrifugal Acceleration of Bodies at Rest

As our final accelerated frame of reference, we will consider the "Earth Carousel". We now take into account the daily rotation of the earth relative to the system of fixed stars. A complete rotation by $360° = 2\pi$ takes 86 164 s. The angular velocity of the globe is thus small; it has the value

$$\omega = \frac{2\pi}{86\,164\,\mathrm{s}} = 7.3 \cdot 10^{-5}\,\mathrm{s}^{-1} .$$

This angular velocity ω produces a centrifugal force on every body at rest on the earth's surface, $F = ma_\mathrm{c}$, and the corresponding centrifugal acceleration a_c.

Consider a body located at the geographic latitude φ (Fig. 7.19). $r = R\cos\varphi$ is the radius of the associated *circle of latitude*. Then the centrifugal acceleration is

$$a_\mathrm{c} = \omega^2 r = \omega^2 R\cos\varphi = 0.03\cos\varphi\ \mathrm{m/s^2} \tag{7.7}$$

(rounded off!).

This centrifugal acceleration is directed outwards, parallel to the radius of the circle of latitude. In the vertical, i.e. in the direction of the radius R of the earth, only one component of this centrifugal acceleration acts, namely

$$a_\mathrm{R} = a_\mathrm{c}\cos\varphi = 0.03\cos^2\varphi\ \mathrm{m/sec^2} . \tag{7.8}$$

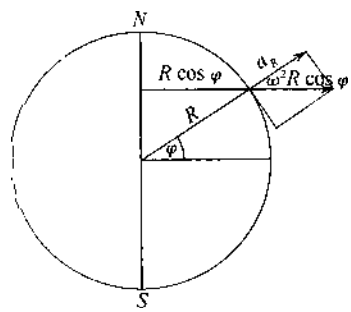


Figure 7.19 Gravitational attraction and centrifugal force as a function of the latitude φ on the earth's surface

It is directed outwards from the center of the earth and is thus opposite to the acceleration of gravity g which results from gravitational attraction. On the *rotating* earth, therefore, the acceleration of gravity at the geographic latitude φ is somewhat smaller than it would be if the earth were at rest. We find

$$g_\varphi = g - 0.03\cos^2\varphi\ \mathrm{m/s^2}\,. \tag{7.9}$$

Here, g, the value of the acceleration of gravity, applies to the earth *at rest*. Now we encounter a difficulty: The centrifugal force acts by no means only on bodies *on the surface* of the earth; indeed, every particle of the earth itself is subject to a centrifugal force, directed radially outwards from its circle of latitude. The sum of all these forces produces an elastic deformation of the earth; the globe is somewhat flattened at the poles, where its axis is about 1/300 shorter than its diameter at the equator. As a result of this polar flattening of the earth, the change in the acceleration of gravity $g_\cdot$ at the latitude φ is somewhat greater than one would calculate from Eq. (7.9). The experimentally-measured acceleration of gravity is described by the expression

$$g_\varphi = (9.832 - 0.052\cos^2\varphi)\ \mathrm{m/s^2}\,. \tag{7.10}$$

At sea level and a latitude of 45°, one finds $g = 9.806\,\mathrm{m/s^2}$. For $\varphi = 0°$, i.e. on the equator, the correction term is maximal; it is then 0.5 %. This small correction can be neglected in many experiments. But a precise pendulum clock at the equator in fact loses about 3.5 minutes per day as compared to a similar clock at the poles.

The polar flattening of about 1/300 mentioned above holds for the solid body of the earth. The deformation of its liquid sheath, the oceans, due to centrifugal force is much greater. But this latter deformation never occurs alone; it is combined with the diurnally and periodically varying attraction of the oceans' water by the sun and the moon. The hydrosphere is thus deformed much more strongly by these tidal forces than is the solid body of the earth. The superposition of centrifugal forces and gravitational tidal forces yields the complicated phenomenon of the *tides*. It is an example of "forced oscillations" (Sect. 11.10). We can give only a rough sketch here.

Similar conclusions hold for our atmosphere, the 'ocean of air'. The tides in the atmosphere cause only small *pressure variations* at its bottom, on the surface of the earth, just like the ocean tides at the bottom of the oceans. But around 100 km above the surface, the air tides cause motions of the air of the order of kilometers! Up there, the tidal waves of the air are much higher than those of the water on the surface of the oceans.

7.7 Earth as an Accelerated Frame of Reference. CORIOLIS Force on Moving Bodies

For an observer looking down on the north pole, the earth rotates in a counter-clockwise sense. It thus has the same sense of rotation as our swivel chair in Sect. 7.3. The angular velocity ω_0 of the earth was computed in Sect. 7.6; it is $\omega_0 = 7.3 \cdot 10^{-5}\,\mathrm{s}^{-1}$.

In Fig. 7.20, an observer is shown at a location of latitude φ on the earth. $H - H$ indicates his/her horizontal plane. At this location, the angular velocity of the earth can be decomposed into two components, one parallel to the radius R of the earth, the *vertical* component

$$\omega_\mathrm{v} = \omega_0 \sin\varphi \tag{7.11}$$

and one parallel to the horizontal plane, the *horizontal* component

$$\omega_\mathrm{h} = \omega_0 \cos\varphi\,. \tag{7.12}$$

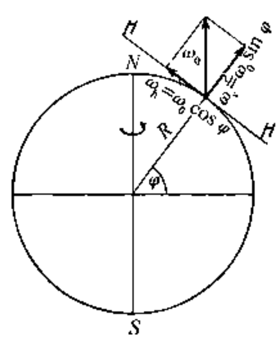
These two components of the angular velocity produce CORIOLIS accelerations of *moving* bodies. We start with the influence of the vertical component ω_v. In the northern hemisphere, it always leads to a deviation *to the right* of moving bodies. The most well-known example is provided by FOUCAULT'*s pendulum*. Its principle was already explained in Sect. 7.3 using a pendulum on a rotating chair: The pendulum followed a rosette orbit, which always curved to the right (Fig. 7.14).

A corresponding rosette path is followed by every long gravity pendulum, consisting of a cord and a weighted ball, on the surface of the earth. The end points of the rosette move forward, as seen from the rest point of the pendulum, by an angle $\alpha = \sin\varphi\,\frac{360^\circ}{24}$ per hour. In Göttingen ($\varphi = 51.5^\circ$), $\alpha \approx 12^\circ$.

The experimental demonstration presents no difficulties in any lecture room. Figure 7.21 shows a reliable setup. Its essential

Video 7.5:
"FOUCAULT's pendulum"
http://tiny.cc/luqujy

Figure 7.20 The two components of the angular velocity of the earth on its surface, for determining the CORIOLIS acceleration **(Video 7.5)**



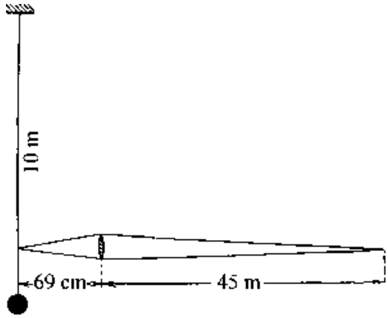


Figure 7.21 FOUCAULT's pendulum experiment **(Video 7.5)**

component is a good astronomical objective lens. The lens projects a strongly enlarged image of the cord of the pendulum at the turning points of the rosette loops. The figure contains the necessary numerical data. With the dimensions chosen, one can see the loops of the rosettes with their turning points spaced at around 2 cm in the enlarged image. *Thus, in a single cycle of the pendulum's swings, the earth's rotation around its axis can be verified!* (**Exercise 7.6**).

Still more transparent, but unfortunately more difficult to demonstrate, is an experiment to detect the earth's rotation carried out by J. G. HAGEN, S. J.[C7.11]. In Fig. 7.22, we explain this experiment using a swivel chair. An axle R held at a tilted angle carries a dumbbell-shaped object with a moment of inertia Θ_1. (We can initially ignore the spiral spring on the torsion axle.) The object is at rest on the rotating chair, and thus it has an angular velocity of $\omega_0 \sin\varphi$. When the string F is burnt through, two helical springs S pull the dumbbell weights close to the axle R and thereby decrease the moment of inertia to the value Θ_2. *During* this motion, the two weights experience a CORIOLIS acceleration and are deflected to the right. This causes the dumbbell to begin moving; it rotates relative to the swivel chair with an angular velocity ω_2. The magnitude of ω_2 can be calculated for an observer at rest relative to the ground by making use of the conservation of angular momentum. We require

$$\Theta_1\omega_0 \sin\varphi = \Theta_2(\omega_0 \sin\varphi + \omega_2)$$

or

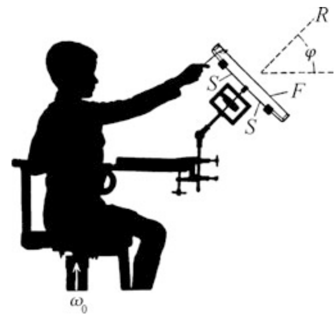
$$\omega_2 = \omega_0 \frac{\Theta_1 - \Theta_2}{\Theta_2} \sin\varphi\,. \tag{7.13}$$

ω_2 attains its maximum value for $\varphi = 90°$, thus at the "pole".
To stabilize the rest position, one attaches a spiral spring to the axle R as shown in Fig. 7.22 (torsion axle). Then the angular velocity ω_2 causes an *oscillation*, not a continuous rotation. In the original experiment of HAGEN, the axle and the spiral spring were replaced by a long torsion band, and the "dumbbell-shaped object" itself was 9 m long.

C7.11. J.G. Hagen: "*La rotation de la terre, ses preuves mécaniques anciennes et nouvelles*", *Tipografia Poliglotta Vaticana*, Roma 1912. J.G. HAGEN, S.J. (1847–1930) was director of the Vatican Observatory from 1907–1930. (The letters S. J. (*Societa Jesu*) indicate that he was a Jesuit.)

These two experiments can be neatly and quantitatively demonstrated. We mention also some more qualitative observations. In these, also, the *vertical* component of the angular velocity of the earth is relevant. In the northern hemisphere, it produces a deflection of moving bodies to the *right* due to the CORIOLIS force:

Figure 7.22 Model experiment to detect the earth's rotation, as carried out by J. G. HAGEN. The same torsion axis is used as in Figs. 6.5 and 6.6[C7.12]



C7.12. The shadow image of the small apparatus required a small person as the experimenter. In the picture: The author's son, R. O. POHL.

a) The air in the earth's atmosphere flows from the subtropical high-pressure regions to the equatorial low-pressure trough. In the northern hemisphere, this flow moves from the northeast to the southwest (deflection to the right); this is the origin of the northeast trade winds which are important for sailing ships and aircraft;

b) Bullets always deviate to the right;

c) For the wearing of railroad tracks and the erosion of river banks, the CORIOLIS forces due to the earth's rotation play no significant role. These examples, often cited in earlier times, can be discarded (**Exercise 7.3**).

CORIOLIS accelerations due to the *horizontal* component of the earth's angular velocity ω_0, i.e. $\omega_h = \omega_0 \cos\varphi$, can also be detected experimentally. But there is no experiment to demonstrate them which is as straightforward as FOUCAULT's pendulum experiment.

A further qualitative example is the easterly deflection of falling stones. But its demonstration requires dropping from considerable heights, preferably in a mine shaft[C7.13].

C7.13. … or in the drop tower at the University of Bremen (cf. comment C7.2).

Among the most important applications of the CORIOLIS force due to the earth's rotation in modern times is the *gyrocompass* (Fig. 7.15). For ships and aircraft, it is a serviceable compass only when it is constructed to be insensitive to yawing and rolling and to accelerations on takeoff and landing or braking, and to flying or sailing around curves. This can be accomplished by using three gyroscopes, whose horizontally-mounted axes are at angles of 120°, and by damping the precession oscillations whose periods are long (ideally, $T = 84$ min; cf. Sect. 7.5). But even perfectly-constructed gyrocompasses exhibit errors in their indication which depend on the direction and velocity of the moving vehicle. The reason for this is that all vehicles are moving momentarily on a great circle on the earth's surface; they thus have a momentary angular velocity ω_2, which adds vectorially to the rotational angular velocity of the earth, ω_1. The axis of the gyroscope therefore does not remain in the meridian plane, but rather moves to a plane containing the resultant vector of the angular velocity from the addition of ω_1 and ω_2. This plane is coincident with a meridian plane only when the vehicle is moving along the equator (see also Comment C7.10).

Exercises

7.1 A pistol is mounted on a rotating platform, whose rotational frequency is $\nu = 10\,\mathrm{s}^{-1}$. The barrel of the pistol points from the center of the platform radially outwards. The platform rotates, as seen from above, in a counter-clockwise sense (and thus as in Sect. 7.3, Point 5). The mass of the bullet is $m = 4\,\mathrm{g}$ and its velocity is $u = 100\,\mathrm{m/s}$; for simplicity, we assume it to be constant. a) Find the Coriolis force F_C which the bullet exerts against the barrel of the pistol. b) Determine the force F_H in the rest frame of the lecture room which the barrel of the pistol exerts against the bullet. (Sect. 7.3)

7.2 A pendulum hangs motionless from a cord attached to the ceiling of the lecture room and is observed by an experimenter who is sitting on a swivel chair that is rotating with the angular frequency ω. How does this observer describe the motion of the pendulum and the forces which act on the pendulum in the horizontal plane? (Sect. 7.3)

7.3 Find the horizontal component a_C of the Coriolis acceleration of a railroad train which is moving at a velocity of 60 km/h in the direction of the line of longitude through London (the latitude is $\varphi = 51.4°$). (Sect. 7.3)

7.4 A satellite circles the earth near its surface, i.e. on a circular orbit with a radius $R \approx R_E = 6371\,\mathrm{km}$ (earth's radius). Determine its period T. The acceleration of gravity is g; neglect air friction. (Sects. 2.5 and 5.9)

7.5 A physical gravity pendulum is to be mounted in such a way that it will oscillate with the Schuler period. How large must the spacing s be, i.e. the distance between the suspension point O of the pendulum and its center of gravity S (cf. Fig. 6.15)? In order not to have to take the details of the pendulum's shape into account, we describe its size by its *radius of inertia* ϱ, defined in the equation $\Theta = \int r^2 \mathrm{d}m = m\varrho^2$ (see e.g. K. Magnus, "*Kreisel*", Springer (1971), p. 12; English see e.g. W. Wrigley, W.M. Hollister, and W.G. Denhard (1969), "*Gyroscopic Theory, Design, and Instrumentation*" (MIT Press, Cambridge, MA).). Assume that $\varrho = 0.1\,\mathrm{m}$. (Sect. 7.5)

7.6 The pendulum shown in Video 7.5, "Foucault's pendulum", swings with an amplitude of $A = 1\,\mathrm{m}$ (measured at the point of the pendulum wire which is projected onto the scale on the wall of the lecture room). The pendulum wire has a diameter of $d = 0.4\,\mathrm{mm}$. Göttingen lies at a latitude of $\varphi = 51.5°$. With these three items of information and the results of the measurements which are shown in the video, derive the angular velocity of the earth's rotation. (Sect. 7.7)

For Sect. 7.3, see also Exercise 6.11; for Sect. 7.5, see also Exercise 2.7.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_7) contains supplementary material, which is available to authorized users.

Some Properties of Solids

8

Part I

8.1 Preliminary Remarks

From early on, children distinguish between solid and liquid bodies; the meaning of the word 'gaseous' becomes clear only much later. In physics, our understanding of gases is rather complete. The distinction between solid and liquid bodies, in contrast, is more difficult. This is not a matter of limiting cases, as in biology for the conceptual distinction between animals and plants: Major groups of materials that we encounter in everyday life, such as resinous and glassy substances, can be broken like solids, but even the layman notices their similarity to very thick, viscous liquids which flow extremely slowly. With increasing temperature, their liquid properties become more and more prominent, without passing through a well-defined melting point[C8.1].

"From early on, children distinguish between solid and liquid bodies; the meaning of the word 'gaseous' becomes clear only much later"

C8.1. Another example is *silly putty*, well known as a toy for children: When it is thrown onto the floor, it bounces like an elastic ball. Left at rest, it flows and spreads outwards on the floor slowly, under its own weight, like a liquid.

8.2 Elastic Deformation, Flow and Solidification

We repeat a few points from the preceding chapters: Every solid body can be deformed by forces acting on it. In the simplest cases, the deformation is not permanent. It disappears when the "load" is removed. The deformation is then called *elastic*. Now we want to discuss the deformation of solid bodies in somewhat more detail and treat it quantitatively.

We begin with the simple arrangement shown in Fig. 8.1. A copper wire, several meters long and 0.4 mm in diameter, is stretched by a constant force and its elongation is measured.

In order to describe this and similar measurements, we define the ratio

$$\frac{\text{Length change } \Delta l}{\text{Original length } l} = \varepsilon \tag{8.1}$$

$=$ *elongation* (for $\varepsilon > 0$) or *compression* (for $\varepsilon < 0$).

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_8

Video 8.1:
“Elastic deformation: HOOKE’s law”
http://tiny.cc/cvqujy
(**Exercise 8.1**)

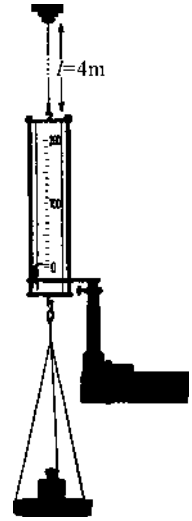


Figure 8.1 The stretching of a metal wire by a tensile force. The scale which is attached to the lower end of the long wire is magnified by about 15x and projected onto a screen. **(Video 8.1)**

C8.2. Like ε, σ can also be negative; i.e. the force and the area vector are directed oppositely. In this case, σ is called the *pressure*. In this chapter, we will consider mainly tensile stresses. Eq. (8.2) as a vector equation becomes $\boldsymbol{F} = \boldsymbol{\sigma} \boldsymbol{A}$, where σ can be a tensor in the most general case. (For more details, see e.g. Charles Kittel, *Introduction to Solid-State Physics*, John Wiley and Sons, Heidelberg, New York, 8th edition (2005), Chap. 3.)

In addition, the quotient[C8.2]

$$\sigma = \frac{\text{Force } F \text{ perpendicular to the area } A}{\text{Cross-sectional area } A \text{ of the wire or rod}} \tag{8.2}$$

is called the *tensile stress* (for short: *tension*, in general ‘normalized stress’; see Sect. 8.7).

We will treat exact and careful observations in Sects. 8.3–8.7. To start with, we describe some quick and less precise experiments. From these, we obtain the relatively simple results shown in Fig. 8.2. Initially, the elongation ε increases *proportionally* to the tensile stress σ; later, around β, it increases *faster* than proportionally. Up to this point β, i.e. up to an elongation of about 1/1000, the deformation remains elastic, or *reversible*; it disappears when the stress (load) is

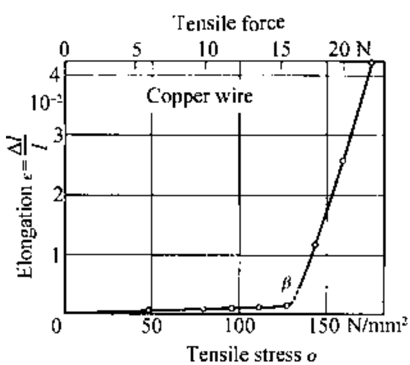


Figure 8.2 The relationship between elongation and tensile stress for a copper wire (diameter $d = 0.4$ mm, cross-sectional area $A = 0.126\,\text{mm}^2$). Modulus of elasticity $E = \sigma/\varepsilon \approx 10^5\,\text{N/mm}^2$[C8.3].

C8.3. Frequently, in “strain-stress plots” of this type, the stress is plotted against the strain (elongation).

removed. Beyond β, the elongation increases quickly with further increasing stress. This deformation is no longer reversible; at β, the *stretching* or *flow limit* is reached. Further stretching *hardens* the previously soft wire. Through heating (tempering), the hardened wire can again be softened (made ductile).

8.3 HOOKE's Law and POISSON's Relation

For small values of the stress, one finds a linear proportionality between the elongation ε and the tensile stress σ. This is HOOKE's law

$$\varepsilon = \frac{1}{E}\sigma\,. \tag{8.3}$$

The proportionality factor E is called the *modulus of elasticity* (or "Young's modulus"). (Examples are given in Table 8.1.)

Using thick wires, or rather rods, one can at the same time observe the elongation and the *transverse contraction*, defined by the ratio

$$\varepsilon_{\mathrm{q}} = -\frac{\text{Change in the diameter } \Delta d}{\text{Original diameter } d}\,. \tag{8.4}$$

For demonstration experiments, a rubber rod of several cm thickness is suitable. When the workpiece is sufficiently thick, one can demonstrate not only its elongation and transverse contraction under tensile stress, but also compression and the accompanying transverse thickening ($\varepsilon_{\mathrm{q}} < 0$) under pressure. Within certain limits, the transverse contraction ε_{q} and the elongation ε are mutually proportional; we find

$$\varepsilon_{\mathrm{q}} = \mu\varepsilon \tag{8.5}$$

(the relation of S. D. POISSON, 1781–1840).

The proportionality factor μ is called *POISSON's number* (Examples are found in Table 8.1).

Elongation and transverse contraction cause a change in the volume, just as compression and transverse thickening do. If the body is

Table 8.1 Elastic constants

Material	Al	Pb	Cu	Brass	Steel	Glass	Granite	Oak wood	
Young's modulus E	7.3	1.7	12	10	20	7	2.4	10	$10^4\,\frac{\mathrm{N}}{\mathrm{mm}^2}$
Poisson's number μ	0.34	0.45	0.35	0.35	0.27	0.2	–	–	–
Shear modulus G	2.6	0.8	4.5	4.2	8.1	2	–	–	$10^4\,\frac{\mathrm{N}}{\mathrm{mm}^2}$

a cube, its height is changed by the factor $(1+\varepsilon)$ and its cross-sectional area by the factor $(1-\mu\varepsilon)^2$. It follows for the *increase in volume* ("cubic dilation"):

$$\frac{\Delta V}{V} = (1+\varepsilon)(1-\mu\varepsilon)^2 - 1 \tag{8.6}$$

or, neglecting small quadratic terms,

$$\frac{\Delta V}{V} = (1-2\mu)\varepsilon\,. \tag{8.7}$$

2μ, twice Poisson's number, is according to Table 8.1 always less than 1. As a result, the volume is always increased by elongation ($\varepsilon > 0$), and always decreased by compression ($\varepsilon < 0$). When a compressive stress is applied from *all directions* (pressure), the volume change is three times as great as with a compressive stress in only *one* dimension (unilateral stress); then Eq. (8.7) combined with HOOKE's law (8.3) yields

$$\frac{\Delta V}{V} = 3(1-2\mu)\frac{1}{E}\sigma = \kappa\sigma\,. \tag{8.8}$$

The constant factor

$$\kappa = 3(1-2\mu)\frac{1}{E} \tag{8.9}$$

is called the "compressibility" and its reciprocal $K = 1/\kappa$ is the "bulk modulus" of the material.

The limiting case $\mu = 0.5$ means that there is no change in the volume due to stress. This limit applies generally to liquids; cf. Sect. 9.3.

8.4 Shear Stress

Thus far, we have assumed that the force F producing the stress acts *perpendicular* to the cross-sectional area A of the workpiece (wire or rod). In this case, the quotient F/A was called the tensile stress ($\sigma > 0$) or compressive stress ($\sigma < 0$). – In Fig. 8.3, in contrast, the force

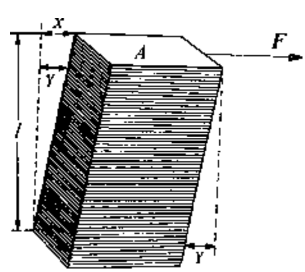


Figure 8.3 The definition of the shear stress

F acts *parallel* to the cross-sectional area A of a body. (One can think of this body as a pack of playing cards to model the situation!). Then the body will be *sheared* by the force F; the cards which originally formed an upright pile are tilted through the angle γ. In this case, we define the *shear strain* as the quotient

$$\frac{x}{l} = \tan\gamma \approx \gamma. \tag{8.10}$$

The quotient

$$\tau = \frac{\text{Force } F \text{ parallel to the area } A}{\text{Cross-sectional area } A \text{ of the body}} \tag{8.11}$$

is called the *shear stress* (see Sect. 8.5).

For *small* stresses, one finds experimentally that the shear strain γ is proportional to the shear stress τ, i.e.

$$\gamma = \frac{1}{G}\,\tau\,. \tag{8.12}$$

The proportionality factor G is called the *shear modulus*. This quantity is also a characteristic property of the material (examples are given in Table 8.1).

Thus, for isotropic bodies, we have all together three elastic constants, namely Young's modulus E, defined by Eq. (8.3); the shear modulus G, defined by Eq. (8.12); and Poisson's number μ, defined by Eq. (8.5). These three constants are however not mutually independent; instead, they are related by the equation

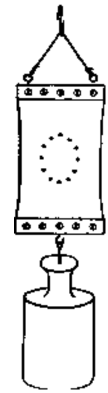
$$G = E\frac{1}{2(1+\mu)}\,. \tag{8.13}$$

Therefore, the elasticity of an isotropic body can be fully characterized by *two* elastic constants; the third is then determined by Eq. (8.13). Its derivation is given at the end of Sect. 8.5.

8.5 Normal, Shear and Principal Stress

Every load placed on a body, e.g. by a tensile stress, changes the state of the interior of the body. The state is characterized by the quality of *strain*. In the interior of a transparent body in the stress-free state, we imagine that a number of *small, spherical* regions are made visible by coloring. During loading, each of these small spheres is deformed into a small, three-dimensional *ellipsoid*. This can be made clear by a demonstration experiment (Fig. 8.4). It deals with the special case

Figure 8.4 The definition of the concepts ‘stress’ and ‘strain’



of a *planar* stress: In the plane of the page, we have a broad rubber band. A circle is drawn in the center of the unstressed band in the form of 12 dots. The ends of the band are clamped into frames, to which a tensile force is then applied in the plane of the page. During this loading, the circle is deformed into an ellipse. In making the transition from a circle to an ellipse, the 12 dots moved along the arrows in Fig. 8.5. A similar situation holds in the general case, that is for the transition from a sphere to a three-dimensional “deformation ellipsoid” (strain ellipsoid).

We can arrive at the concept of *strain* by carrying out the following thought experiment: We cut the ellipsoid out of its surrounding material, at the same time applying forces to its surface which maintain its ellipsoidal shape, thus replacing the effective influence of the previously surrounding material. Or, expressed differently: We transform the “internal forces” which are due to the surrounding material into “external forces” and make them (at least in principle) measurable. The directions of these forces are collinear with the arrows in Fig. 8.5 *only* along the three principal axes of the ellipsoid. Apart from these, their magnitudes are *not* proportional to the lengths of the transition arrows. – Then for each surface element dA of the deformation ellipsoid, we define the *stress* by the quotient dF/surface area dA, the force per surface area. The force is in general *not perpendicular* to the associated surface element dA. Therefore, the stress is decomposed into two components, one perpendicular and one par-

Figure 8.5 How the ellipse in Fig. 8.4 is formed

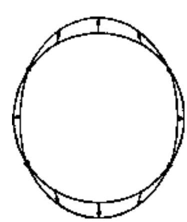
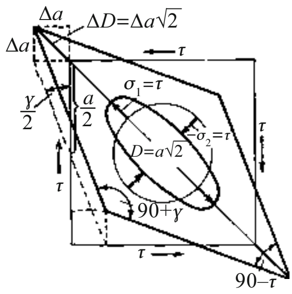


Figure 8.6 Deformation of a rubber plate by four equal forces, each of which produces a shear stress τ. – The edge length is a, the thickness of the plate is d, so that $\tau = F/A = F/ad$. The figure shows the relation between the shear and normal stresses and can be used to derive Eq. (8.13).



allel to the surface element. The perpendicular component is called the *normal stress*; the parallel component is the *shear stress*.

The three principal axes of the ellipsoid are special directions: In these directions, the force is exactly perpendicular to the surface of the ellipsoid. Along these directions, only normal stresses are present, and these are called the three *principal stresses*.

Shear stresses cannot be produced independently of normal stresses. This can be seen from a simple observation: In Fig. 8.6, we attempt to deform a square plate of thickness d by shear forces alone. We apply four equal forces F which act parallel to the sides a of the plate. Each of them produces the shear stress $\tau = F/ad$. The result is however the same as in Fig. 8.4, where a tensile load is applied; a circle is deformed into an ellipse. Thus, normal stresses are also produced. Their largest and smallest values, the principal stresses σ_1 and σ_2, occur in the directions of the diagonals. In the diagonal directions, two of the forces F combine to give a resultant $F\sqrt{2}$. These forces $F\sqrt{2}$ are each perpendicular to a diagonal cross-sectional area $ad\sqrt{2}$. As a result, the normal stresses σ_1 and σ_2 are likewise F/ad, thus exactly the same as the shear stresses τ. Thus, the deformation (strain) of the plate can be described in two ways: either by a *displacement* of the square's sides a by Δa, or by an *elongation* of the square's *diagonals* D by ΔD.

To calculate Δa, one can use the *shear* stress τ. It produces a shear strain of

$$\gamma = \frac{1}{G}\,\tau. \tag{8.12}$$

In simple terms, this means that the 90°-angles are changed to angles of $(90° \pm \gamma)$, and the sides of the square are tipped by angles $\gamma/2$ relative to the diagonals D. From Fig. 8.6, we can see the geometric relation $\tan\gamma/2 = 2\Delta a/a \approx \gamma/2$, or, with Eq. (8.12),

$$\frac{2\Delta a}{a} = \frac{1}{2}\frac{1}{G}\tau\,. \tag{8.14}$$

To calculate ΔD, we use the *normal* stresses, i.e. the tensile stresses $\sigma_1 = \tau$ and the compressive stresses $-\sigma_2 = \tau$. The tensile stresses lengthen the diagonals by an amount $2\Delta D_{\text{tens}} = \varepsilon D = \sigma_1 \frac{1}{E} D = \tau \frac{1}{E} D$. In addition, according to POISSON's relation (Eq. (8.5)), the compressive stresses also cause a lengthening of the diagonals by the amount $2\Delta D_{\text{comp}} = \mu\varepsilon D =$

$\mu\sigma_1 \frac{1}{E} D = \mu\tau \frac{1}{E} D$. The resulting overall lengthening of the diagonals from both sides is then

$$2\Delta D = 2\Delta D_{\text{tens}} + 2\Delta D_{\text{comp}} = \tau \frac{1}{E}(1+\mu)D\,, \qquad (8.15)$$

and, introducing the edge length a,

$$2\Delta a\sqrt{2} = \tau \frac{1}{E}(1+\mu)a\sqrt{2}$$

or

$$\frac{2\Delta a}{a} = \tau \frac{1}{E}(1+\mu)\,. \qquad (8.16)$$

Combining Eq. (8.14) with Eq. (8.16) yields

$$\frac{1}{G} = 2\frac{1}{E}(1+\mu)\,.$$

This is Eq. (8.13), as given previously without derivation.

To conclude, we read off another important fact for later use from Fig. 8.6: The directions of the principal stresses (the diagonals) and the directions of the largest shear stresses (along the edges) are at angles of 45° to each other.

8.6 Bending and Twisting (Torsion)

In applying the concepts of normal stress σ and shear stress τ, we have thus far limited ourselves to the simplest examples. Now we want to consider the bending of a rod by an external torque M.

Take a cuboid-shaped rubber eraser between the thumb and forefinger and bend it: Its sides are not only curved, but also arched outward. We will neglect this arching, and consider only the limiting case of a "planar" state of strain. In Fig. 8.7, a thin rod with a constant cross-sectional area A is bent by a constant torque $M = mg \cdot s$ (m = mass of a weight). We observe its shape to be that of a circular segment.

We want to compute the radius of curvature r of the bent rod. We make use of Fig. 8.8; it shows a cross-section along the length of the rod. The deformation produces tensile stresses along the upper side of the rod, and compressive stresses along the lower side. Both are normal stresses, i.e. they are perpendicular to the cross section A.

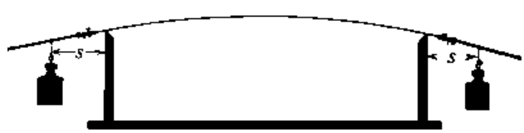
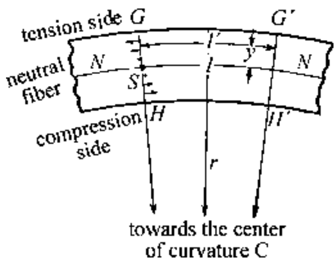


Figure 8.7 Bending load on a thin, flat rod produced by a torque M which is constant along its length **(Video 8.2)** (see also the numerical example in **Exercise 8.2**)

Video 8.2: "Bending a rod"
http://tiny.cc/4uqujy

Figure 8.8 The derivation of Eq. (8.22)



With the assumptions made above, which are well fulfilled for thin rods, the cross-sectional areas denoted as GH, $G'H'$ etc. should remain *planar* even during bending, that is they should simply *rotate* around their centers of gravity S. Then the transition from tensile to compressive stress occurs in a simply curved, and not arched, layer. It is stress free, stands perpendicular to the plane of the page, and cuts it along the line $N-N$. One calls this line the *neutral fiber* (cf. Vol. 2, Sect. 24.9, Strain birefringence).

Under these conditions, in Fig. 8.8 we have for the two radii of curvature r and $(r+y)$:

$$\frac{r+y}{r} = \frac{l'}{l} \,. \tag{8.17}$$

Furthermore, we have

$$\frac{l'-l}{l} = \text{elongation } \varepsilon \,. \tag{8.18}$$

According to HOOKE's law, this elongation is associated with a modulus of elasticity

$$E = \frac{\sigma}{\varepsilon} \,. \tag{8.3}$$

Combining Eqns. (8.3), (8.17), and (8.18) yields

$$\sigma = E\frac{y}{r} \,. \tag{8.19}$$

The integral $\int \sigma y \, \mathrm{d}A$ must be equal to the effective torque M, so that

$$M = \int E \frac{y^2}{r} \mathrm{d}A \,, \tag{8.20}$$

or, with the abbreviation

$$\int y^2 \, \mathrm{d}A = J \,, \tag{8.21}$$

we obtain

$$r = E\frac{J}{M} \,. \tag{8.22}$$

Part I

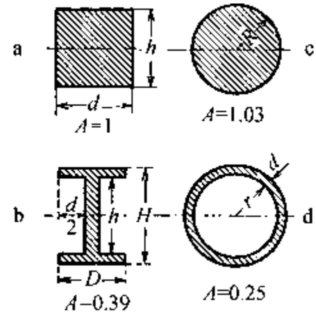


Figure 8.9 Cross sections with the same values of the geometrical moment of inertia J can have rather different areas A (A is normalized here to the rectangular cross section (a)). The neutral fiber which is perpendicular to the plane of the page and passes through S in Fig. 8.8 is dot-dashed here

The quantity J is composed formally just like the moment of inertia, i.e.

$$\Theta = \int y^2 \, \mathrm{d}m \,. \tag{6.4}$$

This value of Θ holds for a layer within the cross section of the rod and is referred to the center of gravity S of the layer. As a result, we can use the formulas for the moment of inertia which we established earlier (Sect. 6.4) to arrive at values for J: We need only replace the mass m in those formulas by the cross-sectional area A. For this reason, the unfortunate name *geometrical moment of inertia* has been adopted for J.

Examples

1. Rectangular cross section, $A = hd$ (Fig. 8.9a)

$$J = \frac{1}{12} dh^3 \,. \tag{8.23}$$

2. Double-T beam (Fig. 8.9b)

$$J = \frac{1}{12}(DH^3 - dh^3) \,. \tag{8.24}$$

3. Circular, ring-shaped cross section (Fig. 8.9c,d)

$$J = \frac{\pi}{4}(R^4 - r^4) \,. \tag{8.25}$$

4. Similar, but for *twisting around the long axis* of a tube whose wall thickness is $d = R - r$ (Fig. 8.11)

$$J = \frac{\pi}{2}(R^4 - r^4) \approx 2\pi R^3 d \,. \tag{8.26}$$

For the rod in Fig. 8.7, E, J and M were constant along its length. As a result, from Eq. (8.22), the radius of curvature r is also constant, i.e. the bent rod takes the shape of a circular segment. Its radius r as calculated from Eq. (8.22) agrees well with observed values (**Video 8.2**).

Video 8.2: "Bending a rod" http://tiny.cc/4uqujy.

The importance of the geometrical moment of inertia J is explained in Fig. 8.9. It shows profiles with the same values of the geometrical

Figure 8.10 Bending load on a rod of length l which is held at one end, produced by a force F acting on its free end

moment of inertia J, that is the same circular radius of curvature for the same loads. Below each profile is its cross-sectional area, given in arbitrary units. A smaller cross-sectional area A means that less material is required. In this regard, a tube is superior to a rod of similar diameter. This is the reason why the long bones in our appendages are formed as tubular bones.

The geometric moment of inertia also plays a decisive role in many other aspects of deformation. We offer two examples, the first without derivation. In Fig. 8.10, a *rod is clamped at one end*; a force F acts perpendicular to its long axis at its free end. Then for a moderate deflection y of the end of the rod, we have

$$y = F\frac{l^3}{3EJ}\,. \tag{8.27}$$

Furthermore, we briefly treat the *twisting (torsion) of a cylindrical rod*[C8.4]. It is also determined by a geometrical moment of inertia. We find the quotient

C8.4. See also the torsion experiment in Fig. 6.7 (**Video 6.1**, http://tiny.cc/csqujy) (**Exercise 8.3**)

$$\frac{\text{Torque } M}{\text{Angle of torsion } \alpha'} = D^* = \frac{GJ}{l} \tag{8.28}$$

(J = geometrical moment of inertia of the rod (Eq. (8.26) with $r = 0$), l = length of the rod, G is the shear modulus of the material; cf. Table 8.1).

D^* is the “torsion coefficient” introduced in Sect. 6.3. It can readily be measured, either directly or by using torsional oscillations. Equation (8.28) thus provides a handy method of determining the shear modulus G, an important quantity for materials science (Table 8.1).

To derive Eq. (8.28), we make use of a *special case* as shown in Fig. 8.11, a thin-walled tube. The torque is produced by two cords wound around the tube. We imagine the tube to be divided into flat ring-shaped sections; these experience a shearing (angle γ) relative to each other. The meaning of the angle γ can be seen from the figure. We find

$$\gamma \approx \tan\gamma = \frac{x}{l} = \frac{\alpha' R}{l}\,. \tag{8.29}$$

This shear deformation is caused by the shear stress τ, for which we have

$$\gamma = \frac{1}{G}\tau\,. \tag{8.12}$$

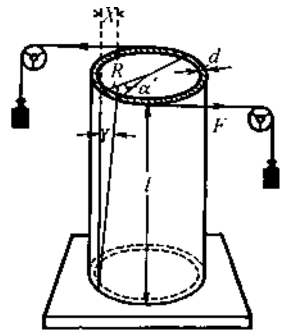


Figure 8.11 The derivation of Eq. (8.28) for the torsional deformation of a tube

The shear stress, in turn, is found from the torque M acting on the tube; it produces forces $F = M/2R$ acting tangentially on the ring-shaped sections, and thus the shear stress

$$\tau = \frac{2F}{\text{Cross-section}} = \frac{M}{2\pi R^2 d}\,. \tag{8.30}$$

The equations (8.29), (8.12) and (8.30), together with Eq. (8.26), yield

$$\frac{\alpha'}{l} = \frac{M}{G\,2\pi dR^3} = \frac{M}{GJ}\,. \tag{8.31}$$

To conclude this topic, we mention a technical application of Eq. (8.28). For transferring or conducting mechanical power ("kilowatts"), one often uses a *shaft*. It is simply a cylindrical rod which is loaded torsionally. For the transmitted power $\dot{W}$ we find for continuous motion

$$\dot{W} = Fu \tag{5.35}$$

(u = orbital velocity),

and thus for a rotation (row 10 in Table 6.1)

$$\dot{W} = M\omega = M 2\pi\nu \tag{8.32}$$

(ω = angular velocity, ν = rotational frequency = (number of revolutions/time).

Thus, instead of Eq. (8.31), we could write:

$$\text{Torsion angle } \alpha' = \frac{\dot{W}}{2\pi\nu}\frac{l}{GJ}\,. \tag{8.33}$$

In words: *At a given rotational frequency ν, the torsion angle α' is a measure of the mechanical power transmitted along the shaft.*

Numerical example

The hollow drive shaft of a ship: $l = 62$ m, outer diameter $2R = 0.625$ m, inner diameter $2r = 0.480$ m, geometrical moment of inertia J (from Eq. (8.26)) $= 9.77 \cdot 10^{-3}\,\text{m}^4$, material steel, with a shear modulus of $G = 8.1 \cdot 10^4\,\text{N/mm}^2$. The power transmitted to the ship's screw is

$\dot{W} = 2.4 \cdot 10^4$ kW, the rotational frequency is $\nu = 3.4$/s. – Inserting these values into Eq. (8.33) yields the torsion angle $\alpha' = 8.8 \cdot 10^{-2} = 5°$. This means that the front and back ends of the 62 m long shaft are twisted relatively to each other by 0.014 of their circumference.
Drill stems for deep vertical bore holes can have lengths of several kilometers. They must then be twisted through a number of rotations in order to transmit the power required for drilling to the drill bit at depth![C8.5]

C8.5. For deep drilling projects, such as for example the deep drilling program KTB in Southeastern Germany, with a depth of over 9 km, multiple twists indeed occur in the drill stem. However, hydraulic motors are also employed, directly above the drill bit, so than no twisting of the drill stem occurs (see e.g. https://en.wikipedia.org/wiki/German_Continental_Deep_Drilling_Program or https://en.wikipedia.org/wiki/Kola_Superdeep_Borehole)

8.7 Time Dependence of Deformation. Elastic Aftereffects and Hysteresis

For quantitative observations of elastic deformation, we have thus far considered metals, as in Figs. 8.1, 8.7 and 8.10. Glasses are also suitable. For demonstration experiments, a polymeric, plastic material is often more convenient, in particular rubber. We will demonstrate two important companion phenomena of elastic deformation using rubber, namely elastic aftereffects (also called the "memory effect"), and hysteresis. In our first *brief* demonstrations, we left these effects out of consideration.

We load a 0.3 m long and ca. 5 mm thick rubber tube alternately with around 1 and 6 N tensile forces (a 100-g weight and an additional 500-g weight), and record its elongation as a function of time. The result is shown in Fig. 8.12: its deformation occurs mainly at the same time as its loading or unloading; but a remainder, called the *elastic aftereffect*, takes a noticeable time to form after loading and to disappear after unloading. The new equilibrium values are approached to a good approximation exponentially with time. On loading, after the *relaxation time* τ has elapsed, about $1/e \approx 37\,\%$ of the full amount of elastic aftereffect is still missing. On removing the load, after a time τ, about $1/e \approx 37\,\%$ of the elastic aftereffect is still present.

Unfortunately, the separation of deformations into those with and without aftereffects would be an oversimplification, even for *small*

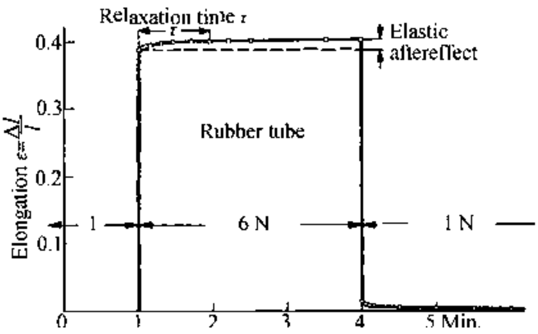


Figure 8.12 Elastic aftereffect on stretching a rubber tube (outer diameter ≈ 5 mm, inner diameter ≈ 3 mm)

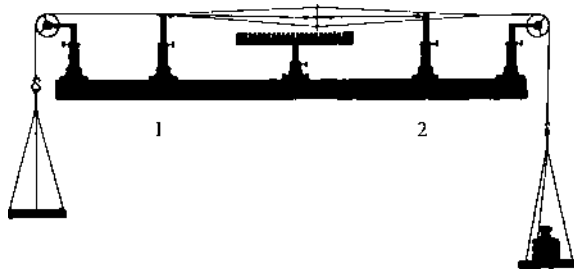


Figure 8.13 The demonstration of mechanical hysteresis. A stretched rubber tube is attached at its ends, *1* and *2*. In its center, its two halves are divided by a metal disk. A pair of cables can exert forces on this disk, which can be increased or decreased by adding or removing weights on one or the other of the two balance pans.

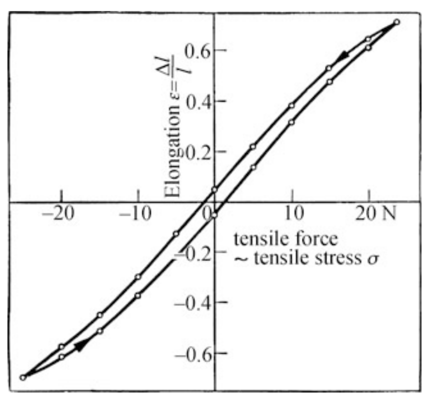
deformations. When the load is removed, some fraction of the preceding elongation remains as a *permanent* deformation. It can be eliminated only by loading in the *opposite* direction. This is called *hysteresis*. It can be demonstrated using the apparatus shown in Fig. 8.13.

A rubber hose which is clamped at both ends and already stretched to about twice its relaxed length can be further loaded with a tensile stress applied stepwise as increases or decreases from the left or the right (weights). Between two measurements, there is a pause of at least one minute. The results are indicated in Fig. 8.14. The relation between elongation and tensile stress along the increasing and the decreasing branch is shown by two curves, and we can see in Fig. 8.14 that they enclose a narrow area, the mechanical *hysteresis loop*. One observes similar behavior for nearly all solid bodies, and thus also for metals, glasses etc.[C8.6]

C8.6. Hysteresis occurs also in the electrical and magnetic properties of materials (see Vol. 2, "Matter in electric and magnetic fields", Sects. 13.4 and 14.4).

A small part of *every* deformation is thus not reversible; the material is not completely elastic. A small fraction of the tensile energy which drives the elongation is always "lost" in the form of heat. In an ε-σ

Figure 8.14 A hysteresis curve obtained with the apparatus shown in Fig. 8.13. Forces acting to the right are taken to be positive. The measurement series begins at the upper right corner.



plot, the area within the hysteresis loop is the work per volume which is transformed into heat during a cycle of loading and unloading:

$$\frac{\text{Loss per load cycle}}{\text{Volume of the deformed body}} = \frac{-\Delta W}{V}\,.$$

(The product $\varepsilon \cdot \sigma$, i.e. $(\Delta l/l) \cdot (F/A)$, has the units of work/volume).

The origin of elastic aftereffects and of hysteresis[C8.7] is related to the structure of the material. During elastic deformation, individual regions of the bulk can shift or rotate relative to one another and then act as *'latches'*. Releasing the latches can occur either through thermal motions alone (aftereffects), or it may require loading in the opposite direction (hysteresis).

C8.7. A quite different kind of hysteresis is exhibited by "shape-memory alloys", such as NiTi, when they pass through a structural phase transition during temperature changes. This can also be produced by deformation. They can, for example, be plastically deformed at room temperature, but on warming, they return to their original shape. They are used among other things in medical technology (see e.g. http://web.stanford.edu/~richlin1/sma/sma.html).

8.8 Rupture Strength and Specific Surface Energy of Solids

When the load becomes sufficiently large, every solid body will be torn or ruptured into smaller pieces **(Video 8.3)**. In idealized limiting cases, the tearing seams ("fissures") are found to lie either perpendicular to the direction of greatest normal stress, or parallel to the plane of greatest shear stress. For that reason, one distinguishes between *tensile strength* and *shear strength*; when they are exceeded, the result is *separation fractures* or *shear fractures*. The directions of the greatest tensile and shear stresses are tilted by ±45° relative to one another (see Fig. 8.6). Therefore, when brittle materials are pressed, we find rupture surfaces tilted at around 45° from the direction of the compressive stress.

Video 8.3: "Plastic deformation, rupture strength"
http://tiny.cc/kvqujy
A 40 cm-long Cu wire (of diameter = 0.4 mm) is placed under tensile stress by hanging weights on its end, until it breaks. The rupture strength of the copper used is found to be $Z_{max} \approx 240\,\mathrm{N/mm^2}$.

Between the regime of elastic deformation and rupture, many materials exhibit some other phenomena, namely a regime of *flow or slippage* of their individual volume elements and an accompanying *hardening*. Some metals can be rolled out to thin sheets or drawn through a die to make wire even at room temperature (cold working). The gradual change of shape associated with "plastic deformation" makes the process of rupture still more complex than with brittle materials, i.e. materials that rupture without previous plastic deformation (e.g. glass or cast iron). – Plasticity and brittleness are not well-defined properties of a substance; at high temperatures, every material becomes more or less plastic (malleable, ductile).

Table 8.2 contains some values of the *tensile strength* of materials for technical purposes. This is the term applied to the tensile stress Z_{max} which leads to rupture. These values were measured on standardized rods. – To properly evaluate these numbers, try the following simple experiment: Cut a strip about 20 cm long and 3 cm wide from a sheet of good-quality writing paper, hold it by its ends and try to tear it by pulling uniformly. This is seldom possible. Then cut a small notch into one side of the strip, barely 1 mm deep. Now, the paper strip can readily be made to tear: at the "tip of the notch", a kind of leverage

Part I

Video 8.3:
"Plastic deformation rupture strength"
http://tiny.cc/kvqujy

C8.8. Exceptionally large values of the tensile strength have been observed in various spider silk fibers: Z_{max} of 800–1600 N/mm^2. See B.O. Swanson *et al.*, *Applied Physics* **A82**, 213–218 (2006).

C8.9. This value originates from a publication by E. Orowan (*Z. Physik* **82**, 235 (1933)).

C8.10. Bending produces tensile stresses on the outside of the bent region which are inversely proportional to the radius of curvature of the bend (Sect. 8.6), and can therefore become especially large.

Table 8.2 Technical tensile strengths Z_{max} **(Video 8.3)**[C8.8] (To define the tensile stress, the original cross-sectional area of the rods was used, not the reduced area during elongation.)

Material	Al	Pb	Cu	Brass	Steel	Mica	Quartz glass	Wood fiber
Z_{max}	300	20	400	600	up to 2000	750	800	up to 120 N/mm^2

produces locally a very high tensile stress, and this tears the notch deeper and deeper. Even microscopic notches or other defects can play a decisive role in rupture.

In some cases, the perturbing effects of notches and surface defects can be avoided. With mica, for example, one can orient the cleavage planes parallel to the direction of tensile stress and also protect the edges, where notches could play a significant role, by using a suitable clamping frame. In this way, tensile strengths of up to $Z_{max} = 3180\,\text{N/mm}^2$ have been measured for mica.

With very thin (diameters of a few µm) fibers of glass or quartz, freshly prepared at high temperature, tensile strengths of up to $Z_{max} > 10\,000\,\text{N/mm}^2$ have been attained[C8.9].

For demonstrations, such fibers are loaded by bending them[C8.10]; one takes a piece a few centimeters long between the finger tips. Surprisingly sharp bends (small radii of curvature) can be attained without breaking the fiber. The smallest defects on the surface, however, cause premature breakage; it suffices to touch the bent fiber with another glass fiber.

In the interior (bulk) of a material, the molecules are surrounded on all sides by their neighbors; at the surface, however, the neighbors are missing on one side. As a result, work must be performed to move a molecule from the bulk to a surface position. The quotient

$$\zeta = \frac{\text{Work } \Delta W \text{ required to increase the surface area}}{\text{Area } \Delta A \text{ of the newly-formed surface}} \qquad (8.34)$$

is called the *specific surface energy*. It can be estimated from the tensile strength of a material measured without disturbance from notches or surface defects.

In the schematic view shown in Fig. 8.15, a wire of cross-sectional area A is broken with a separation fracture. This produces two new surfaces of area A, and requires that the work $W = 2A\zeta$ be performed. This work is performed by the force $F = Z_{max}A$ acting along a short distance x. We thus find

$$2\zeta A = Z_{max}Ax \quad \text{or} \quad \zeta = \tfrac{1}{2}Z_{max}x\,. \qquad (8.35)$$

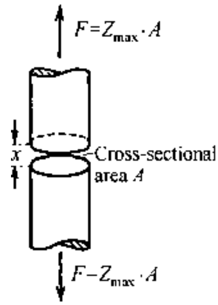


Figure 8.15 The derivation of Eq. (8.35)

The distance x must be of the same order of magnitude as the range of the atomic attractive forces or the distance between neighboring atoms in the material. This is of the order of 10^{-10} m. Then from Eq. (8.35), it follows that the specific surface energy of glass with $Z_{max} \approx 10\,000\,\mathrm{N/mm^2}$ is

$$\zeta \approx 5 \cdot 10^9 \,\frac{\mathrm{N}}{\mathrm{m}^2} \cdot 10^{-10}\,\mathrm{m} = 0.5\,\frac{\mathrm{W\,s}}{\mathrm{m}^2}\,.$$

The high values of the tensile strengths of materials implied by Eq. (8.35) are called the *theoretical* tensile strengths. They can be larger than the measured *technical* tensile strengths by more than an order of magnitude[C8.11]. The technical strength is determined essentially by perturbing secondary effects. "Notching" is indeed a rough simplification, but it is still an appropriate generic term for these effects.

C8.11. Even liquids can exhibit large tensile strengths (see Sect. 9.5 and **Video 9.4: "The tensile strength of water"**, http://tiny.cc/3vqujy).

8.9 Sticking and Sliding Friction

External friction (Sect. 3.1) occurs at the contact area between two solid bodies. It plays a fundamental role in everyday life and in technical applications. For physical experimentation, it is often a source of disturbance and of errors. Quantitative results depend strongly on the surface condition of the objects which are in contact. – We can distinguish between three forms of external friction: sticking friction, sliding friction, and rolling friction.

In Fig. 8.16, a smooth, box-shaped block is pressed against a smooth horizontal board by the force of its weight, $F_G = F_n$. A cord transmits a force F to the block, parallel to the board's surface. This tractive force must exceed a threshold value, F_{st}, before the block will start to slide along the board. From this we conclude that the two objects (block and board) *stick* to each other; a force of magnitude F_{st} acts in opposite directions on each of them. It is called *sticking friction*. This frictional force is independent of the size of the contact area between the two objects. It depends on the properties and condition of the contact surfaces and is proportional to the

Part I

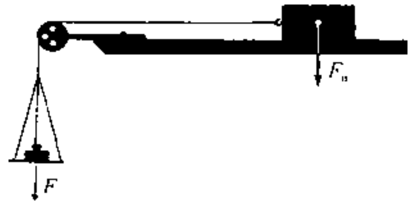


Figure 8.16 Sticking and sliding friction

magnitude of the normal force that presses them together (perpendicular to their contact surfaces), in this case F_n. The equation

$$F_{st} = \mu_{st} F_n$$

defines the coefficient μ_{st} of sticking friction (a number, usually between 0.2 and 0.7). Sticking friction plays an important role in technical applications. It determines the maximum possible value of the driving force for locomotive wheels and automobile tires, as well as for the shoe soles of a pedestrian. At the point of contact, a rolling wheel or the shoe sole are momentarily at rest relative to the ground. Therefore, it is *sticking* friction that applies here.

The force we call sticking friction arises through extremely small displacements of two objects relative to one another. In the simplest picture, one can compare the *smooth* surface of any solid body to a file or brush. The microscopic parts that stick out catch on each other, so that they have to be deformed to permit sliding. – A more refined model would have to take the adsorbed layers of other molecules on the surfaces into account. Without these layers, the sticking friction between two polished surfaces can become vanishingly small. An example: A glass block and a glass plate in high vacuum. The word 'sticking friction' is dubious; sticking is too strong an expression; it includes the concept of *adhesion*. Furthermore, one speaks properly of *friction* only during mutual *motions* and not before they begin.

Continuing our observations using the setup shown in Fig. 8.16, we make the traction force F pulling on the block larger than the sticking friction F_{st}. Then the block begins to slide in an accelerated motion. Its acceleration a however does not correspond to the tractive force F; it is smaller. Therefore, while the block is sliding, besides the force F there must be a smaller, oppositely-directed force F_{sl} acting on the block. This force F_{sl} is called the *sliding friction*.

The sliding friction F_{sl} is always smaller than the sticking friction F_{st}. It is, like the latter, proportional to the normal force F_n which presses the two bodies together, and is independent of the size of their contact area; thus:

$$F_{sl} = \mu_{sl} F_n$$

(μ_{sl} = coefficient of sliding friction, a number usually between 0.2 and 0.5).

The sliding friction is independent of the sliding velocity only to first order. With increasing velocity, it may decrease by up to ca. 20 % of the original frictional force which applies at very small velocities[C8.12].

C8.12. The velocity dependence of the frictional force which is mentioned here is quite different from that observed for objects moving in liquids or gases (see Chap. 10). In that case, the frictional force *increases* proportionally to the velocity, a dependence which is generally assumed in a first treatment of damped oscillations. In this connection, note the experiment which is described together with Fig. 5.22.

Measurements of the sliding friction are carried out at constant sliding velocities. For this purpose, one replaces the setup for the tractive force as shown in Fig. 8.16 by an electric motor with a winch, inserting a force meter in the cord which transmits the force.

In machines, *external friction* is usually reduced as far as possible; it is replaced by the *internal friction* of fluids. This is termed *lubrication*.

In some cases, one has to be satisfied with reducing the external friction to a minimum. This can be accomplished only for a particular direction of motion (Fig. 8.17). Perpendicular to this direction, one applies an ancillary force F_2 to maintain a constant velocity. During this ancillary motion, only the small component F'_{sl} opposes the force F_1. – Figure 8.18 illustrates a demonstration experiment.

When asked, 'Why, when we are *cutting* something with a knife, e.g. slicing bread, do we not only press down on the knife, but also move it back and forth?', even many physicists will give a wrong answer: 'The back-and-forth motion converts the large sticking friction into a smaller sliding friction'. In reality, the motion has two effects: First, the knife acts as a saw; and second, the sliding friction in the *direction of cutting*[1] is reduced by means of an ancillary force which acts perpendicular to it. One thus makes use of the scheme illustrated in Fig. 8.17. At the beginning of the cutting process, when the knife blade is resting on the crust of the bread, only the sawing effect is present.
Another example is pulling a wedge out of a crevice: One moves the wedge back and forth along the line of the crevice. Unintended, but – due to vibrations – often unavoidable relative motions between the sides of the wedge and the walls of the crevice produce the sometimes fatal *loosening of screws*. Screws are in fact simply "rolled-up wedges".

"When asked, 'Why, when we are cutting something with a knife, e.g. slicing bread, do we not only press down on the knife, but also move it back and forth?', even many physicists will give a wrong answer."

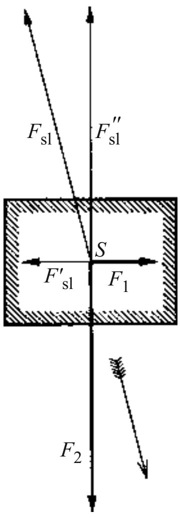


Figure 8.17 An object (center of gravity S) is to be moved to the right on its horizontal supporting plate. To produce a motion with a constant velocity, an ancillary force F_2 compensates the component F''_{sl} of the sliding friction F_{sl}. As a result, the tractive force F_1 need overcome only the small component F'_{sl} of the sliding friction.

[1] The direction of cutting is the direction in which the knife penetrates like a wedge into the object being cut.

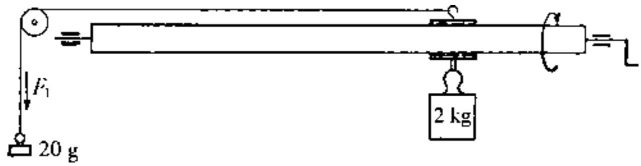


Figure 8.18 A demonstration experiment showing the reduction of sliding friction by means of an *ancillary force which performs work*: The tractive force F_1 is much smaller than the sliding friction $F_{sl} \approx 4\,\text{N}$, or the still larger sticking force F_{st}. Nevertheless, the block slides when the crank is rotated or swung, i.e. when the ancillary force F_2 is acting tangential to the cylinder (length of the cylinder 1 m, diameter 2 cm). For a 'free-hand' experiment, a pencil, held at an angle, suffices as rotatable cylinder, and a ring can be used instead of the block. **(Video 8.4)**

Video 8.4:
"Reducing sliding friction"
http://tiny.cc/mvqujy

8.10 Rolling Friction

A wheel (of radius r) is pressed onto a horizontal track with a force F_n which acts perpendicular to the track. In order to cause the wheel to *roll* along the track with a constant velocity, one must apply a torque

$$M = \mu_{ro} F_n$$

to the wheel; its magnitude is practically independent of the velocity. This equation defines the coefficient μ_{ro} of rolling friction. It is found to be a small *length* between ca. 10^{-2} mm and 1 mm.

Rolling friction, in contrast to sticking friction and sliding friction, has nothing at all to do with adhesion. It cannot be reduced by "lubrication". Rolling friction is caused by the elastic deformation of the track and the wheel at their point of contact. This point moves along the track and around the circumference of the wheel with the same velocity as the forward movement of the wheel. Since there is no ideally elastic deformation, aftereffects and hysteresis always lead to energy losses.

In order to generate the torque M to produce the rolling motion[2], one can for example apply a force to the axle of the wheel. As driving (tractive) force F_1, it must act parallel to the track (e.g. the rails of a railroad track). The resistance force F_2 acts as counter-force (i.e. $F_2 = -F_1$), and we will call it the rolling resistance in the following.

The rolling resistance is important for all vehicles with wheels. In all such vehicles, whether they are automobiles, locomotives, tractors or trailers, the motor has to perform work against the rolling resistance F_2 of all the wheels (including the drive wheels). Wagons are vastly superior to the sledges used before the invention of the wheel

[2] Around momentary axes of rotation (A_m in Fig. 6.4).

roughly 5000 years ago: To pull a wagon, one requires a considerably smaller force than to pull a sledge with the same weight F_G. A wagon requires the force $F_{Wa} = M/r = \mu_{ro} F_G / r$, while a sledge needs a force $F_{Sl} = \mu_{sl} F_G$. We thus find for the ratio of the forces

$$\frac{F_{Wa}}{F_{Sl}} = \frac{\mu_{ro}}{\mu_{sl} \cdot r} .$$

Example

$\mu_{ro} = 1\,\text{mm}$, $\mu_{sl} = 0.5$, $r = 50\,\text{cm}$, $F_{Wa}/F_{Sl} = 1/250$.

This ratio becomes very small when wheels with a large radius are used. Thus, replacing sledges by wagons with *large* wheels was an enormously important invention.

Exercises

8.1 In Video 8.1, "Elastic deformation: HOOKE's law", a copper wire of diameter 0.4 mm and length 4 m is loaded by a weight of 400 g and thereby reversibly stretched to an additional length of 1 mm. Find the modulus of elasticity E of the wire. (Sect. 8.2)

8.2 In Video 8.2, "Bending a rod", the brass rod (with a cross-sectional area of 12 mm width and 4 mm height) is lying on a table of length $L = 0.65\,\text{m}$. When each overhanging end is loaded by a weight of 1 kg, it is reversibly deformed into a circular segment, as shown in the video. The midpoint of the rod is raised by a height H. Measure H and the length s (s is required to determine the torque, as indicated in Fig. 8.7). From the quantities L and H, find the radius of curvature r of the rod. Use these values to calculate the modulus of elasticity of brass. For length measurements in the video, use your knowledge that the length of the table is $L = 0.65\,\text{m}$. (Sect. 8.6)

8.3 In Video 6.1, "Twisting a rod", the length l of the cylindrical steel rod (more precisely, the spacing of the clamps which hold the mirror) is $l = 9\,\text{cm}$. The diameter of the rod is $d = 1\,\text{cm}$. The length of the light pointer, which is just about equal to the diagonal of the experimental area of the lecture hall, is $L = 10\,\text{m}$. The torsion angle α can be found by comparing with the rod of diameter $d_S = 1.4\,\text{cm}$ which is mounted close to the projection screen on which the light pointer is seen. The torque M applied by the experimenter, which he measured by using a torque wrench (not shown in the video), is $M = 0.6\,\text{N m}$. From these data and the torsion angle α (to be measured), calculate the shear modulus G of the steel rod. (Sect. 8.6)

8.4 A wooden block of mass $m = 5\,\text{kg}$ is pulled along a horizontal, flat tabletop with a horizontal force of $F = 25\,\text{N}$. The coefficient of sliding friction between the block and the tabletop is $\mu_{sl} = 0.25$ and is independent of the sliding velocity. Calculate the distance s through which the block moves in 3 s, if its initial velocity was zero. (Sect. 8.9)

For Sect. 8.5, see also Exercise 12.5; for Sect. 8.9, see also Exercises 3.2 and 5.2.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_8) contains supplementary material, which is available to authorized users.

Liquids and Gases at Rest

9

Part I

9.1 The Free Displacements of Liquid Molecules

The distinction between solid and liquid bodies is based on their behavior when their shape is changed. The deformation of solid bodies *always* requires the action of forces. For liquids, in contrast, the force required to change the shape at constant volume becomes smaller and smaller, the more *slowly* the deformation is carried out. In the ideal limiting case of an arbitrarily slow deformation at constant volume, no force at all would be required. – We can conclude from this that in solid bodies, the smallest structural units, the molecules, are in the main bound to fixed rest positions. In liquids, in contrast, such fixed rest positions do not exist; all the molecules can be freely displaced relative to one another, they "slide around loosely".

In solid bodies, the invisible, "random" motions which are usually called *thermal motions* consist essentially of oscillations of the molecules around their *rest positions*. In liquids, however, translational and rotational motions are also possible. We can find a greatly oversimplified but still accurate picture of these thermal motions in liquids in the form of the *Brownian molecular motion*[C9.1] (**Video 9.1**).

The essentials can be summarized in an image of almost child-like simplicity: Imagine a bowl filled with live ants; we are looking at it from a certain distance (or nearsightedly). We cannot discern the individual squirming insects, but only a structureless, dark brown surface. A simple trick allows us to see more: We throw a few larger, readily visible, light objects into the bowl, for example downy feathers, snippets of paper, etc. These objects will not remain at rest; shoved and pulled by the indiscernible ants, they will be constantly moving, twisting and turning in a random way. We then see the restless movement of the individual insects in the form of a rather coarsened image.

The demonstration of BROWNian motion is carried out in an analogous manner. We use a fairly sophisticated microscope; between a microscope slide and a cover glass, we put a drop of some liquid, in the simplest case water. This liquid contains a fine, insoluble powder; conveniently, a small amount of India ink can be added to the water, containing a suspension of extremely fine carbon powder (grain diameter $\approx 0.5\,\mu$m).

C9.1. ROBERT BROWN, botanist (1773–1858), discovered this motion initially by observing pollen grains suspended in water; in the course of his investigations, he found that 100-year-old pollen grains also still move, and finally that even inorganic dust grains show similar motions! His description is contained in an article with the title, "A brief account of microscopical observations made in the months of June, July, and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies". This article is readily available in: R. Hardwicke, "*The Miscellaneous Botanical Works of Robert Brown*", London 1865, Vol. 1.

Video 9.1:
"BROWNian motion"
http://tiny.cc/4wqujy
(see also **Video 16.1:** **"Model experiments on diffusion and osmosis"**, http://tiny.cc/xhgvjy.)

"The essentials can be summarized in an image of almost child-like simplicity".

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_9

To show the demonstration to a large audience, one should use a powder with a high index of refraction, for example the mineral rutile (TiO_2). The large index of refraction guarantees a bright image of the powder grains.

Only a few physical phenomena can fascinate the observer like the BROWNian motion. It gives us a peek behind the backdrop of nature's workings. It shows us a new world, the restless, confusing hustle and bustle of a completely incomprehensible number of individual particles. Tiny grains shoot like arrows across the field of view, changing direction in a wild zig-zag. Larger grains jerk forward, constantly changing their course. The largest grains just stagger back and forth, staying nearly in the same place; their jagged edges and corners reveal their rotations around constantly changing axes. Nowhere is a trace of a systematic order to be seen. Random, blind chance rules this world; that is the overpowering impression of every objective observer. – The BROWNian motion belongs among the most important phenomena observed by science. No description in words can begin to replace the effect of first-hand observations.

An effective demonstration of BROWNian motion requires magnification of several 100× by a microscope. Such a high magnification leads us to overestimate the velocities involved. A different method of observation can save us from this error; it shows the suspended grains in the liquid only as a whole, as a *swarm* or a *cloud*, but no longer allows us to discern the individual grains. In Fig. 9.1, we see dust-containing water, e.g. highly diluted India ink, next to a layer of pure water. The boundary between the two liquids is initially sharp, but it becomes more and more washed-out in the course of time. Very slowly, in the course of weeks, the swarm of carbon particles "diffuses" into the previously clear water. *Diffusion* is defined in general as any movement of molecules caused by thermal motions. *Diffusion* and *Brownian motion* are *two different names for the same process*. The term *Brownian motion* presumes the *microscopic* observation of individual, larger particles. When dealing with *macro*scopic observations, we speak of *diffusion*, independently of the size of the particles observed. That is, the objects seen as a swarm or cloud can consist of *dust grains* or of extremely small *dissolved molecules*, invisible to any optical microscope.

For our purposes, the *velocity* of diffusion is the essential point. The boundary of the swarm moves surprisingly slowly. Depending on

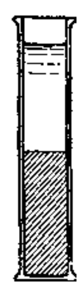


Figure 9.1 The advance of an interface layer by diffusion. The initially sharp boundary layer is prepared by putting a thin, flat cork disk onto the lower layer of water. Pure water can then be cautiously sprayed as a fine jet onto this disk.

the size of the suspended particles, a measurable displacement of the boundary requires days or even weeks.

The reason for the slowness of diffusion is the close packing of the swarming liquid molecules. The mean distance of the molecules in a liquid is of the same order of magnitude as in the corresponding solid material. This can be seen from two facts: The *density* of every substance is roughly the same in the liquid and the solid state; and furthermore, liquids have a very *small compressibility*. This property, with which we are familiar from everyday life, will be verified quantitatively in Sect. 9.3.

After these preliminary considerations, we can replace the real liquid by a *model* liquid and study the properties of liquids using the model. A good model would be a container full of live ants or round beetles with hard wing covers. But a container full of small, smooth steel balls will suffice[C9.2]. Then we will have to simulate the random motions of these model molecules – the "thermal motions" – in a somewhat clumsy way by shaking the whole container. In the following, we will not mention this shaking in every case, but take it for granted.

C9.2. Model concepts in physics are not only useful for didactic reasons; rather, they represent a basic aspect of the physical method, namely reducing complex situations by simplifying assumptions to known laws and descriptions (to "understand" them). Physical models range from point masses in mechanics up to purely mathematical formalisms in theoretical physics. Besides the model liquid made of steel balls described here, we will see a further example of a model concept in Sect. 9.8 (fundamental equation of the kinetic theory of gases).

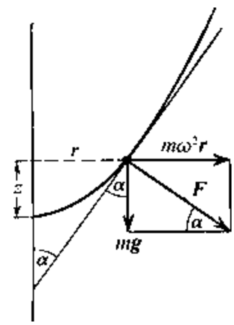
The free displacement of liquid molecules makes a number of properties of resting liquids at equilibrium understandable. They are treated extensively in school physics courses and repeated briefly here in the following Sects. 9.2 and 9.3. We begin by considering the positioning of liquid surfaces.

A liquid surface always adjusts itself to be perpendicular to the direction of the forces which act on its molecules. – In a flat, wide dish, only the *weight* of the individual molecules acts. The surface positions itself in a horizontal plane. In the wide basin of an ocean or a large lake, the forces of weight are no longer parallel at different locations; they point everywhere *radially* towards the center of the earth. As a result, the surface of the liquid (ocean or lake water) takes on the form of a sector of a spherical surface.

In a container which is rotating around a vertical axis, the liquid surface takes the form of a *paraboloid of rotation* (Fig. 9.2). We will explain this from the point of view of the accelerated frame of reference. Two forces act on every individual particle (molecule): its weight $m\boldsymbol{g}$, directed vertically downwards; and the centrifugal force $m\omega^2\boldsymbol{r}$, directed radially outwards. The two forces add vectorially to give the total force $\boldsymbol{F}$. The surface positions itself perpendicular to

Figure 9.2 The parabolic cross-section of a rotating steel-ball model liquid in a rectangular glass container (photographic snapshot)

Figure 9.3 The parabolic surface of a rotating liquid.
$\tan\alpha = \frac{mg}{m\omega^2 r} = \frac{\mathrm{d}r}{\mathrm{d}z}$,
$\frac{g}{\omega^2}\mathrm{d}z = r\mathrm{d}r$
and integrating, $z = \text{const}\cdot r^2$



this total force at every point. Quantitatively, from Fig. 9.3, we have

$$z = \text{const}\cdot r^2\,. \tag{9.1}$$

C9.3. POHL introduced the unit of pressure newton/meter^2 already in the 12th edition (1953); however, he did not yet use the unit name *pascal*. The pressure of the air (Sects. 9.9 and 9.10) is usually quoted in hectopascals (1 hPa = 100 Pa). In these units, atmospheric pressure is typically around 10^3 hPa = 10^5 Pa. In addition, the *bar* is used:
1 bar = 10^5 Pa,
1 mbar = 1 hPa.
The previously-used unit "atmosphere" is defined by:
1 atm = $1.013\cdot10^5$ Pa
$\approx$ 1 bar.
In medicine, e.g. for quoting blood pressure, the older unit mmHg (Torr) is still used (although the World Health Organization recommended as early as 1977 discontinuing the use of this unit by 1983!):
1 mmHg $\approx$ 1.33 hPa.
This is the *gravity pressure* (Sect. 9.4) of a column of mercury (Hg) with a height of 1 mm.

9.2 Pressure in Liquids. Manometers

External forces produce stresses not only in solid bodies (Sect. 8.5), but also in liquids. However, in liquids, we use a different terminology; the stress is called the *pressure p*. *For the pressure in liquids, the force is always perpendicular to the corresponding surface*[1]. This follows from the free deformability of liquids in every direction; or, put differently, the pressure in a liquid at rest is always a normal stress; there is no shear stress in liquids at rest. A spherical volume within a liquid at rest, made visible e.g. by coloring it with a dye, remains spherical under any external force. The sphere will not be distorted into an ellipsoid, but rather at most only changes its radius.

The unit of pressure $p = F/A$ is the

$$\text{newton/meter}^2 = 1\,\text{pascal (Pa)}^{\text{C9.3}}\,.$$

For pressure measurements, the molecules which produce the pressure are replaced by a wall which exerts the same pressure, i.e. one converts the "internal" forces within a solid body into "external" forces. This is accomplished in pressure meters or *manometers*. – In Fig. 9.4, on the left, we see a piston which can be displaced nearly without friction within a cylinder that is attached to the vessel holding a liquid. The piston is itself attached to a spring balance with a pointer and scale. – The piston and spring can be combined into

[1] For solid bodies, all the directions which point outwards from a given closed volume are considered to be positive. A tensile stress is therefore termed positive and a pressure (compressive stress) is negative. In liquids, usually the opposite sign is used by convention: A positive pressure p compresses the liquid volume, a negative p tends to expand it like a tensile stress.

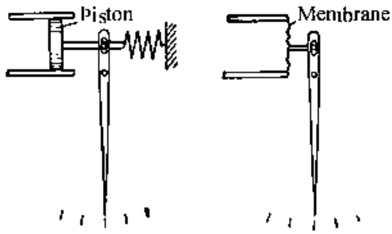


Figure 9.4 Schematics of a piston and a membrane manometer

a single component in the construction of the manometer. We thus arrive at a membrane which may be smooth or grooved (Fig. 9.4, right). The pressure acting on it produces a curvature which moves the pointer. The membrane, which is caused to bulge by pressure, may be replaced by a tube with an elliptical cross-section (Fig. 9.6, left). The tube stretches when it is filled with a liquid under pressure (think of the rolled-up paper "party horns" which unroll when blown into!). *Without calibration, such instruments can be used only to compare spatially or temporally separated pressures*. But we will describe in the next section how they can be calibrated to indicate absolute pressures.

As proud possessors of such an uncalibrated manometer, we will now consider the pressure distribution within liquids. For simplicity, we first distinguish two limiting cases:

1. The pressure in the liquid results only from its own weight. Keyword: *gravity pressure*.

2. The liquid is contained in a vessel which is closed on all sides. An attached cylinder with a smoothly-fitting piston (or "ram") produces an external pressure, much larger than the gravity pressure of the liquid itself. Keyword: *ram pressure*. We will start by considering this second limiting case.

9.3 The Isotropy of Pressure and Its Applications

Figure 9.5 shows an iron vessel of a complex shape, in cross-section; it is filled with water and has four identical manometers attached to it. On the right, we can apply ram pressure via a screw which presses a piston inwards. All of the manometers indicate the same readings and thus show that the pressure is the same in every direction (isotropic). – To explain this phenomenon, we imagine a model liquid (steel balls) which has been filled into a container and is being pressed by a piston through a suitable opening. The container will expand outwards on all sides. The free sliding of the steel balls in every direction will not permit a particular direction to predominate[C9.4].

We next deal with three important applications of the isotropy of the ram pressure.

C9.4. Alternatively, one can imagine that spheres are drawn within the liquid; they just touch each other and thus produce a connection between the piston which produces the pressure and any of the manometers. At the contact surface between two such imaginary spheres, the pressure from both sides must be the same (*actio* = *reactio*), and the same as the pressure in the interiors of the spheres. Therefore, the pressure measured by the manometer is just the same as that produced by the piston. – For a mathematically correct derivation of this fact, also known as PASCAL's law, see A. Sommerfeld, *Mechanik der deformierbaren Medien*, Akademische Verlagsgesellschaft Leipzig, 4th ed. (1957), Chap. 2, Section 6; English see e.g. www.brighthubengineering.com/naval-architecture/106499-hydrostatic-pressure-and-pascals-law-for-static-fluids/

1. *Calibration of a technical manometer* (Fig. 9.6). A tube leads from the manometer *R* to the cylinder *Z* with a tight-fitting piston *K*. All of the inner spaces are filled with a liquid, for example hydraulic oil. Pressure is force divided by area. The ram pressure of the piston is thus equal to its weight plus that of the mass sitting on it, divided by the cross-sectional area *A* of the piston. Now the essential point: The friction between the piston and the wall of the cylinder must be very small; otherwise, the force would be less than the weights just mentioned. Eliminating the friction is accomplished through a trick: The piston is surrounded by a thin film of liquid. This is guaranteed by a uniform rotation of the piston around its vertical axis[2] [C9.5]. To produce the rotation, the top of the piston is designed as a flywheel. Once it is set in rotation, it continues to rotate for a long time. The moving flywheel is given a strong kick from above; each time, the pointer of the manometer returns to the same reading. The position of the indicator is therefore in fact determined by the weights alone.

C9.5. At very low temperatures, such as in the expansion machine of a helium liquefier, this type of lubrication is not possible, since all lubricants solidify. Here, one makes use of helium *gas* as "lubricant"; it likewise forms a laminar flow around the piston. (cf. Sect. 10.3)

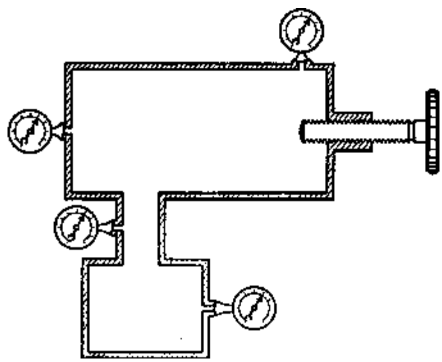


Figure 9.5 The pressure distribution within a liquid when the ram pressure is dominant

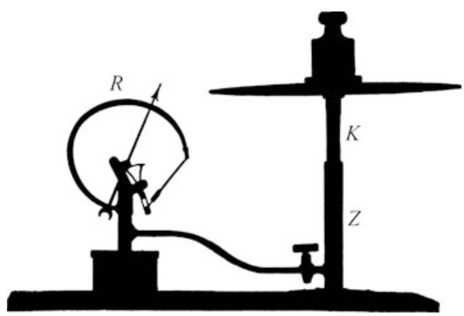


Figure 9.6 Calibration of a technical manometer *R* with a rotating piston *K*

[2] This experiment also illustrates the lubrication of bearings through the laminar flow of a fluid, as described in Sect. 10.3.

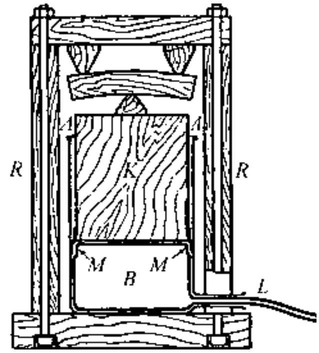


Figure 9.7 An improvised hydraulic press

2. *The hydraulic press.* This important tool is used to produce large forces by means of small pressures. This is a widely-used application in modern times. As an example, we mention the hydraulic jacks used to lift automobiles in repair shops.

We illustrate an hydraulic press (Fig. 9.7) in the form of an improvised setup. Its essential parts are a cylindrical cooking pot A, a thin-walled rubber bladder B, a wooden piston K, and a solid, rectangular frame R. The filling tube of the rubber bladder is attached to a water faucet. A leather sleeve M around the rim of the piston prevents the bladder from 'bulging' up between the piston and the side walls of the pot.

Numerical example

The water supply in the lecture room in Göttingen has a pressure of around $4 \cdot 10^5$ Pa. The cooking pot used has an inner diameter of 30 cm, giving the piston a surface area of about 710 cm^2. The press thus yields a force F of roughly $3 \cdot 10^4$ N. It can for example break oak blocks of $4 \times 5\,\mathrm{cm}^2$ cross-sectional area and 40 cm length.

3. *The compressibility of water.* The low compressibility of liquids can be measured readily using the isotropy of pressure in liquids. The principle is the following: One presses the liquid under high pressure into a measuring vessel, avoiding expansion of the vessel itself. To achieve this, the vessel is surrounded by a mantle of liquid at the same pressure as that inside. The resulting setup is sketched in Fig. 9.8. Initially, a volume decrease ΔV is found to be proportional to the increase of pressure, Δp, and to the volume V, i.e. $\Delta V = \kappa V \Delta p$; the proportionality factor κ is then measured. It is called the *compressibility*[C9.6],

$$\kappa \approx 5 \cdot 10^{-10}\,\mathrm{Pa}^{-1}\,.$$

Thus, the volume decrease $\Delta V/V$ of the pressurized water at 10^8 Pa (about 1000 times higher than atmospheric pressure) is only about 5 %. – This low compressibility of water can be shown in a variety of surprising demonstration experiments. They all illustrate the appearance of large forces and pressures when only a small degree of compression is produced.

C9.6. The compressibility $\kappa = (1/V)\,(\mathrm{d}V/\mathrm{d}p)$, which we introduce here for liquids, and likewise its reciprocal, the modulus of compression $K = 1/\kappa$, were already discussed as material constants of solids (Sect. 8.3) and will be applied also to gases in a later section (Sect. 14.10). Note that for liquids and gases, owing to the isotropy of pressure, a modulus of elasticity cannot be defined.

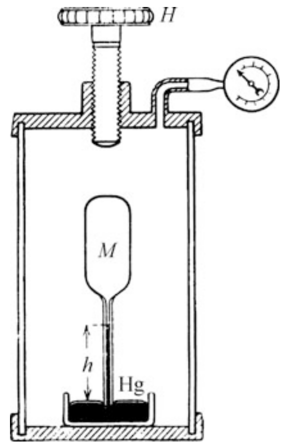


Figure 9.8 The compressibility of water. The thick-walled glass cylinder is filled with water, likewise the thin-walled measuring cuvette M. The handwheel H is employed to press down the piston (ram) by a screw drive. – Hg : Mercury as sealing liquid in the capillary tube (of cross-sectional area A). – The Hg column rises by a height Δh when the pressure is increased by Δp. This indicates a volume decrease of the water contained in M by $\Delta V = \Delta h \cdot A$.

Video 9.2:
"The Compressibility of water"
http://tiny.cc/6vqujy
The video shows a departmental seminar at the 1st Physics Institute of the University of Göttingen, where the properties of glass teardrops were demonstrated. A similar experiment was exhibited by LICHTENBERG in Göttingen many years earlier (see G.H. de Rogier, "*Verstreute Aufzeichnungen aus Georg Christoph Lichtenbergs Vorlesungen über die Experimentalphysik 1781*", Wallstein-Verlag Göttingen 2004, p. 48). The glass teardrops are known under the name of *Bologna* or *Batavia glass* drops. Their behavior also forms the basis for the shattering of safety glass.

Figure 9.9 Two glass teardrops (**Video 9.2**)

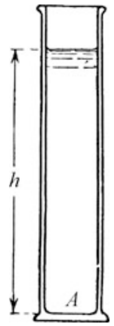
Example
We start with a rectangular, lidless wooden box which is suitably watertight and is filled with water. The liquid has a free upper surface. A bullet is shot through the side of this box, compressing the water by a volume equal to that of the bullet (the time of its passage is too short to allow the water to rise upwards). A considerable pressure results – the box is splintered into kindling (bladder shot!).
A variant of this experiment requires only a modest effort. It suffices to make a '*glass teardrop*' explode within a water-filled beaker. Glass teardrops are made by allowing molten glass to drop into water. They are solid, droplet-shaped pieces of glass with enormous internal stresses (Fig. 9.9). A glass teardrop is very insensitive to blows and impacts; one can hammer on it without effect. In contrast, it will not survive any sort of damage to its filamentary 'tail'. When the tail is broken off, the teardrop explodes into fine glass shards. If a teardrop is exploded in this way in a closed hand, one can clearly feel how the fragments fly apart, but there is no pain or damage (as when the safety glass in an automobile window shatters!). The harmlessness of this experiment in one's hand stands in surprising contrast to the complete destruction of a beaker filled with water ('water hammer').

9.4 The Pressure Distribution in a Gravitational Field. Buoyancy[3]

We use a cylindrical container which stands vertically and has a cross-sectional area A (Fig. 9.10). It is filled with a liquid of

[3] The isotropic atmospheric pressure of the air (Sect. 9.9) is left out of consideration in this section.

Figure 9.10 The gravity pressure of a liquid



density ϱ up to a height h. The weight of this column of liquid is

$$F_G = mg = Ah\varrho g\,. \qquad (9.2)$$

The weight divided by the area gives the pressure p which acts at the bottom of the container in all directions (isotropy):

$$p = \frac{F_G}{A} = h\varrho g\,. \qquad (9.3)$$

Numerical example for water

$h = 10^3$ m, $\varrho = 10^3$ kg/m^3, $g = 9.81$ m/s^2, $p = 10^3$ m $\cdot 10^3$ kg/m^3 $\cdot$ 9.81 m/s^2 = $9.81 \cdot 10^6$ N/m^2, i.e. about 100 times atmospheric pressure[C9.7]. – This pressure compresses the lowest layer of water by only 0.5 % of its volume (see above). Therefore, to a very good approximation, one can treat the density ϱ in Eq. (9.3) as independent of the depth h.

C9.7. The water pressure at the bottom of a 10 m high column of water is about the same as the normal atmospheric pressure (see also Fig. 9.31).

The shape and cross-sectional area of the container do not enter into Eq. (9.3). As a result, one can imagine that even the most strangely-shaped containers give the same pressure profiles as a simple vertical cylinder with a constant cross-sectional area. *The decisive factor for the gravity pressure at a given point within a liquid is only the perpendicular distance h* of the point from the surface of the liquid. Quantitatively, we find Eq. (9.3)[C9.8].

C9.8. BLAISE PASCAL (1623–1662). Published in "Traité de l'Équilibre des Liqueurs" (1663); excerpts from the original text are translated into English in W.F. Magie, *A Source Book in Physics*, Harvard U. Press (1963), p. 75.

Among the various applications of this principle which are often given in school books, we recall the familiar liquid manometers used to measure gas and vapor pressures. Their simplest form consists of a U-shaped glass tube containing water or mercury as sealing liquid (Fig. 9.11). These manometers can be calibrated using Eq. (9.3) (e.g. 1 mmHg $\approx$ 1.33 hPa, see Sect. 9.2).

The best-known result of the pressure distribution in a gravitational field is the *static buoyancy* of bodies in a liquid. We consider the buoyant force acting on a body which is immersed in a liquid. For simplicity, we assume that the body has the form of a flat cylinder (Fig. 9.12). The pressure of the liquid has no preferential direction (isotropy). This results from the free deformability of the liquid in every direction, due to the free displacement of its molecules. Therefore, an upwards-directed force $F_1 = p_1 A = h_1 \varrho g A$ presses on the

Figure 9.11 Liquid manometer

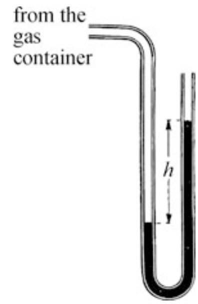
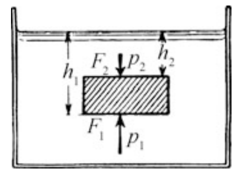


Figure 9.12 The origin of static buoyancy



lower surface A of the cylinder, while a smaller, downwards-directed force $F_2 = p_2A = h_2\varrho gA$ acts on the upper cylinder surface. All the forces acting on the side surfaces of the cylinder compensate each other exactly, pairwise. There remains only the difference of the two forces F_1 and F_2. It yields an upwardly-directed force F which acts on the body as a whole. This force is called the buoyant force or simply *buoyancy*:

$$F = \varrho gA(h_1 - h_2)\,. \tag{9.4}$$

The product on the right is just the weight of that amount of the liquid which has the same volume as the immersed body. We thus find in general: *The buoyant force on an immersed solid body is equal to the weight of the liquid which it displaces* (*Archimedes' principle*)[C9.9].

C9.9. This principle forms the basis of a very simple method of determining the mass density ϱ_O of an object. One need only weigh it once outside the liquid and again when fully immersed in a liquid of known density.

There are a variety of quantitative experiments on buoyancy. Instead of demonstrating them, we will elucidate the origin of buoyancy using our *model liquid*. Figure 9.13 shows the shadow image of a glass container half-filled with many small steel balls. We previously buried two larger balls within the small balls; one of them is made of wood, the other of aluminum. We simulate the thermal motion within our model liquid as usual by shaking the container vigorously. Immediately, buoyancy brings the two large balls to the surface. They *float*, the wooden ball rather high above the surface, the aluminum ball about at its midline.

Figure 9.13 Buoyancy in a steel-ball model liquid **(Video 9.3)**

Video 9.3: "Buoyancy in a model fluid"
http://tiny.cc/fwqujy

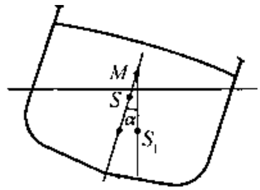


Figure 9.14 The metacenter

Of course, one cannot expect a quantitative verification of the formula for the buoyant force from such a model experiment. The simulation of thermal motions by shaking is too primitive for that.

The weight of an object and its buoyancy in a liquid oppose each other. When the *weight* is greater, the object sinks to the bottom of the liquid. When the *buoyancy* is greater, it rises to the surface and floats there. The transition between these two possibilities illustrates a special case: The object and the liquid which it displaces have exactly the same weight (equal densities). In this case, the object remains suspended at an arbitrary depth within the liquid. This special case can be demonstrated in a variety of ways. As a particular example, we mention a ball of amber in a solution of zinc sulfate of suitably-chosen concentration[C9.10].

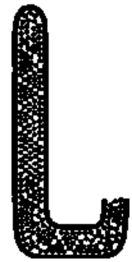
When the buoyant force is greater, a portion of the object rises above the surface of the liquid. The object comes to equilibrium when the displaced water has just the same weight as the object itself. Then we say that the object *floats*. For practical applications (ships), the *stability of the floating attitude* is very important. It is determined by the position of the *metacenter*. In Fig. 9.14, imagine that a ship is tilted (listing) by an angle α from its equilibrium position. Let S_1 be the center of gravity of the volume of water displaced by the ship *in this tilted position*, i.e. the point of action of the *buoyant force* in this position. We draw a vertical line through the point S_1. Its point of intersection with the central midline of the ship is called the *metacenter* M. This metacenter must remain above the center of gravity S of the ship itself during any tilting or listing; only then will the resulting torque due to the buoyant force bring the ship back to its equilibrium resting position. *Only when its metacenter is above its center of gravity can a ship float in a stable manner.*

C9.10. As a simple demonstration, we recall the "Cartesian diver" which is often used as a toy. A small glass figure (often a little devil) containing air is put into a test tube filled with water; it is closed at its top end by a rubber membrane (or a movable rubber stopper). The position of the figure can be readily projected as a shadow image onto the wall of the lecture room. The tail of the figure has a small opening, so that pushing on the membrane or stopper presses some water into the figure, thus increasing its overall density. The three cases rising, sinking and suspension can be convincingly demonstrated by regulating the pressure with the membrane.

9.5 Cohesion of Liquids: Tensile Strength, Surface Energy and Surface Tension

Our model liquid (steel balls) has yet to explain two well-known properties of real liquids: The molecules of a real liquid are subject to attractive forces, they exhibit *cohesion*. When the liquid is poured, they do not move apart in all directions, but rather form droplets of different sizes and shapes. Furthermore, real liquids *stick to* solid bodies (*adhesion*). This sticking can even lead to *wetting*; that is, the

Figure 9.15 The tensile strength of a model liquid



liquid cannot be pulled off the solid body, and attempts to do so only spread it around further. When wetting occurs, the attraction between the molecules of the *liquid* and those of the *solid* body is greater than that between the molecules of the *liquid* itself. – These shortcomings of the model liquid can be overcome; one need only use magnetic steel balls. Then they are attracted both to each other and to the walls of an iron container.

The model liquid which has been modified in this way leads us to another important fact that is not well known from everyday experience: Liquids have a considerable *tensile strength* (this is the limiting value of the tensile stress which leads to fracture or separation; cf. Sect. 8.8).

Figure 9.15 shows a sectional view along the length of an iron pipe which is closed at its top, and is filled with the steel balls of the model liquid. The magnetic "molecules" stick to its walls. They form a continuous *column*. This column carries its own weight, and thus it has a tensile strength in addition to its adhesion to the walls. Figure 9.16 illustrates the same experiment with a real liquid, here a column of water. The right arm B of the U-tube is evacuated. In this manner, we can hang columns of water up to many meters long. They have a tensile strength which is often surprising to observers. The long glass tube can be conveniently attached to a board; when the board is struck hard on the floor, causing strong downwards *inertial forces* to act on the water column, it may require several tries before the column breaks apart.

One point is central for this demonstration of the tensile strength of water, and for the model liquid: The molecules of the liquid must *stick firmly* to the walls of the tube (adhesion). Only then can a cross-sectional constriction of the column of liquid be avoided. This is the reason why no gas bubbles should be left on the walls of the tube; they would immediately act as nucleation points for the formation of constrictions (and resulting separation of the column). For water, a tensile strength of $Z_{\max} = 3.4\,\mathrm{N/mm^2}$ has been achieved; for ethyl ether, the maximum value is $Z_{\max} = 7\,\mathrm{N/mm^2}$ (comparison values for solids can be found in Table 8.2). From the considerations which will be discussed in Sect. 15.9, one would expect values at least 10 times larger. Most likely, the tensile strength is reduced by *nucleation centers*, tiny impurities in the liquid or on the walls of

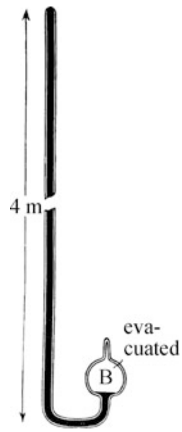


Figure 9.16 The tensile strength of a water column. The water has been purified of dissolved air by boiling in vacuum. The volume B thus contains gaseous water (water vapor) which adjusts to the corresponding vapor pressure (Sect. 15.7). At 20 °C, it is 23.2 hPa (see Fig. 15.11). From this we find, assuming that the water vapor behaves as an ideal gas, that its density is $\varrho = 1.7 \cdot 10^{-2}\,\mathrm{kg/m^3}$ (see Sect. 14.6). **(Video 9.4)**

the container, as discussed in in Sect. 15.9, for example, very small nitrogen-containing gas bubbles.

In our treatment of solids, we have already mentioned the essential relation between the tensile strength $Z_{\max}$ and the specific surface energy ζ, i.e. the quotient

$$\zeta = \frac{\text{Work necessary to increase the surface area } \Delta W}{\text{Size } \Delta A \text{ of the newly-formed surface}} . \qquad (8.34)$$

Here, both ΔA and ΔW can be either positive or negative. Positive means an increase in the surface area. Then, a force F must perform work, and it will be stored in the surface as potential energy. Negative means that the surface area decreases; then previously stored energy is released, performing work and giving rise to a force. In dealing with solid bodies, we treated only the first case; to demonstrate the second case in solids, one requires very high temperatures or long times. This is quite different for liquids: The free sliding of their molecules allows us to implement both cases.

A well-known example is shown in Fig. 9.17. A liquid film (e.g. a soap solution) is held from above and on both sides by a U-shaped wire frame, and from below by a freely sliding wire which is attached to the frame by loops and can slide along it. This "slider" or "strap" can be adjusted to any desired height by a suitably-chosen load (force F). A displacement by $\pm\Delta x$ produces a new surface area $\Delta A = \pm 2l\Delta x$ (front and back surfaces of the film!), so that the force F performs the work

$$\pm\Delta W = \pm F\Delta x = \pm 2\Delta x l\zeta .$$

The distance $\pm\Delta x$ cancels. The remaining result is

$$F = 2l\zeta . \qquad (9.5)$$

The strength of the force F is thus *independent* of Δx, i.e. of the magnitude of the extension which has already taken place. This is

Video 9.4:
"The tensile strength of water"
http://tiny.cc/3vqujy
To avoid gas bubbles on the wall of the tube, it must be swung back and forth several times in a horizontal position, which removes remaining gas bubbles. – A further interesting observation is the following: When the liquid hits the end of the tube while it is being swung, one may hear a loud bang, as if the tube had been hit by a hammer. This does not occur in the air, since a stream of water is slowed down by air resistance, for example before it hits the side of a wash basin. The water vapor at the end of the tube obviously does not show this effect; it immediately liquefies when the vapor is compressed. The speed of this phase transition from the gas phase to the liquid phase is also discussed in Comment C16.2.

Part I

Video 9.5:
"Surface work"
http://tiny.cc/tvqujy

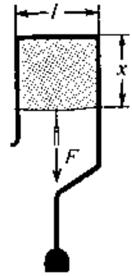


Figure 9.17 A soap-solution film in equilibrium; this is also an example of "reversible" surface work **(Video 9.5)**

an essential difference compared to stretching a rubber sheet. The popular comparison between a liquid surface and a rubber sheet must therefore be used only with caution.

Rewriting Eq. (9.5) yields

$$\zeta = \frac{\text{Force } F \text{ parallel to the surface and required to extend it}}{\text{Length } 2l \text{ of the movable edge of the surface}}\,. \tag{9.6}$$

For this reason, ζ is often referred to as the *surface work*. For liquids, both names are equally valid.

In measurements of ζ, friction between the slider and the frame is a perturbing effect. It is thus preferable to use a *cylindrical* liquid film instead of a planar film (Fig. 9.18). A ring with a *sharp* edge is dipped into the surface of the liquid. When the liquid level is slowly lowered, a cylindrical film is formed, similar to a thin-walled, short tube. The force F is measured with a balance. The circumference of the ring is $l = 2\pi r$. Table 9.1 lists some numerical values, which refer to liquid surfaces in air[C9.11]. When the liquid films are in contact with materials other than air, the values of ζ are generally smaller. Therefore, the term specific *boundary* energy or *boundary* tension would be more appropriate.

C9.11. except for liquid hydrogen

Without external influences, liquids often form spherical surfaces; think for example of a droplet of mercury or of a small bubble within a liquid. In both cases, for the filled sphere as well as for the empty

Table 9.1 The specific surface energy (surface tension) of some liquids

Liquid	Temperature in °C	Specific Surface Energy or Surface Tension in 10^{-3} W s/m^2 = 10^{-3} N/m
Mercury	18	500
Water	0	75.5
	20	72.5
	80	62.3
Benzene	18	29.2
Liquid air	−190	12
Liquid hydrogen	−254	2.5

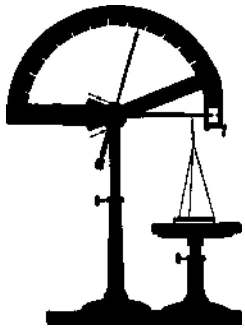


Figure 9.18 The measurement of the specific surface energy using a spiral spring balance. A *numerical example* for water: ring diameter 5 cm, circumference $2l = 0.31$ m, $F = 2.26 \cdot 10^{-2}$ N, $\zeta = 0.072$ W s/m^2.

sphere, the surface tension produces a pressure inside the sphere:

$$p = \frac{2\zeta}{r}\,. \tag{9.7}$$

Derivation

The radius r of the sphere is presumed to increase by the amount dr. Then the surface area of the sphere increases by $\mathrm{d}A = 8\pi r \mathrm{d}r$ and its volume by $\mathrm{d}V = 4\pi r^2 \mathrm{d}r$. During this volume expansion, the pressure performs the work

$$\mathrm{d}W_V = p\mathrm{d}V = p 4\pi r^2 \mathrm{d}r\,. \tag{9.8}$$

Producing the new surface area dA requires the work

$$\mathrm{d}W_A = \mathrm{d}A\zeta = 8\pi r \mathrm{d}r\zeta\,. \tag{9.9}$$

Setting equal these two values for the work yields Eq. (9.7).

This important equation (9.7) is often used in a strict context, but also for approximations. Examples:

1. A mercury droplet at the limit of microscopic visibility has a radius $r = 0.1\,\mu\mathrm{m} = 10^{-7}$ m. ζ for Hg $= 0.5$ W s/m^2; thus we find

$$p = \frac{2 \cdot 0.5\,\mathrm{W\,s/m^2}}{10^{-7}\,\mathrm{m}} = 10^7\,\mathrm{Pa} \quad \text{(i.e. 100 times atmospheric pressure!)}$$

2. Everyone knows the experiment sketched in Fig. 9.19. A wetting liquid is placed between two planar glass slides. It forms a *concave* surface, whose smallest radius of curvature r is $\approx d/2$. Surface tension produces a pressure p of order of magnitude given by Eq. (9.7). The direction of p is marked in the figure by arrows[4]. Glass plates which are "glued" together with water in this way cannot be separated by applied forces without damaging them. They can only be slid apart very slowly under water.

"This is a convenient but lax way of putting it. The pressure itself has no direction, but instead only the corresponding force."

3. A perfectly wetting liquid is pulled upwards into a capillary tube of radius r, up to a height h (Fig. 9.20). – Explanation: The liquid has a *hollow* (concave) upper surface (meniscus). Its smallest radius of curvature is $\approx r$. Then the pressure calculated from Eq. (9.7), $p = 2\zeta/r$, yields an upwardly-directed force of $F = A2\zeta/r$. The downwards-directed weight

[4] This is a convenient but lax way of putting it. The pressure itself has no direction, but instead only the corresponding force.

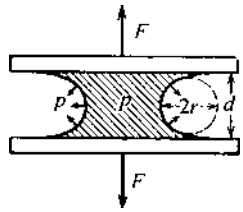


Figure 9.19 A water layer – intentionally drawn much too thick here – between two glass plates (referring to Eq. (9.7)). A *numerical example*: The wetted surface area is $A = 10\,\mathrm{cm}^2$, $d = 0.2\,\mu\mathrm{m}$, $r = 10^{-7}\,\mathrm{m}$, $\zeta \approx 8 \cdot 10^{-2}\,\mathrm{W\,s/m^2}$, $p = 16 \cdot 10^5\,\mathrm{Pa}$ (≈ 16 times atmospheric pressure), $F = 1.6 \cdot 10^3\,\mathrm{N}$.

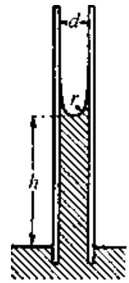


Figure 9.20 Applications of Eq. (9.10): "capillary elevation" h, diameter $d = 2r$

of the liquid column is equal and opposite to this force, with $F_G = Ah\varrho g$. The equilibrium between these two forces yields the *capillary elevation*:

$$h = \frac{2\zeta}{r\varrho g}\,. \tag{9.10}$$

In the case of a non-wetting liquid, e.g. Hg in glass, the meniscus is upwardly *convex*. As a result, the pressure calculated from Eq. (9.10) corresponds to a downwards-directed force. A tube dipped into mercury produces a *capillary depression* in its interior, of height h. – Equation (9.10) is often employed for the measurement of ζ; this effect also plays a role in the rise of sap in plants.

4. Strong inertial forces can produce bubbles, i.e. hollow spaces within liquids. This process, called *cavitation*, occurs for example behind rapidly rotating ship's propellers or in water turbines. – Water has a surface tension of $\zeta \approx 0.08\,\mathrm{W\,s/m^2}$. Therefore, each cm^2 of the surface area of a bubble has a potential energy of $8 \cdot 10^{-6}\,\mathrm{W\,s}$. The resulting pressure p causes the bubbles to collapse very quickly and compresses the energy of their surfaces into a region of only a few molecules. These concentrations of energy act like very high local temperature spikes. As a result, ship's propellers and turbine blades are "eaten up" by the water; they become pitted with deep holes. – Cavitation can also be produced by high-frequency sound waves. The local concentrations of energy can destroy small organisms which live in water, and can cause light flashes from water containing dissolved gases[C9.12].

C9.12. This phenomenon of so-called *sonoluminescence* has been studied in detail in recent years, but is still not completely understood (see for example D. Lohse, *Physikalische Blätter* **51**, 1087 (1995). English: e.g. https://en.wikipedia.org/wiki/Sonoluminescence)

Among the numerous other examples of surface tension, we give here only a brief selection. In the first, the surface of a liquid appears to resemble a slightly stretched wrapping or skin.

1. Water cannot wet greased or oily objects. Such objects can rest on a water surface as on a loosely-filled cushion, for example an air cushion. The surface is visibly slightly 'dented' below the object. For example, one can place a slightly oily sewing needle carefully onto

a water surface and it will rest there, similarly to the legs of a water strider.

> Liquid fuels wet every solid object. Therefore, one never sees dust on their surfaces. Furthermore, they drip very slowly out of a container hung in the open; the completely wetted walls of the container act like a *syphon*. This process occurs very quickly when the liquid is the 'ideal fluid' ^{4}He in its *superfluid phase* (found at temperatures below 2.17 Kelvin), which completely lacks internal friction[C9.13].

C9.13. For more information on the fascinating phenomenon of the *superfluidity* of liquid helium, see for example K. Lüders, "*Superflüssigkeiten*", in Bergmann/Schaefer, *Lehrbuch der Experimentalphysik*, Vol. 5, Chap. 5, Verlag de Gruyter 2nd ed. (2006). English: see e.g. https://en.wikipedia.org/wiki/Superfluid_helium-4

In our remaining examples, the surface tension produces the maximum possible *reduction* of the surface area of a free liquid surface (minimal surfaces).

2. Mercury is injected as a fine stream into a flat watch glass, filled with a liquid. It initially forms numerous droplets at the bottom of the glass, of roughly 1 mm diameter (Fig. 9.21). The resulting overall surface area of the mercury is thus quite large. But then the droplets begin to combine in a jerky manner; here and there, a small droplet is taken up by a larger one, which limits the lifetime of the small droplets[5]. After about 1 minute, only a single large puddle of mercury is present. The surface area of the mercury, driven by its surface tension, has been reduced to the minimum possible under the circumstances. This is an especially informative demonstration: The small droplets are "physical individuals". The fate of a particular individual cannot be predicted using physical methods; one can never say which of the droplets will be the next to disappear. Nevertheless, for the set of all the individuals, we can find a clear-cut physical rule: Their number decreases according to an exponential law[C9.14] with a given *mean lifetime* τ (in Fig. 9.21, $\tau = 10\,\mathrm{s}$). *One can thus make quite precise statements about a large set of individuals, even when such statements are completely meaningless for particular, isolated individuals.* This fact plays an important role in atomic physics (e.g. in radioactive decay processes).

"This is an especially informative demonstration."

C9.14. This is an example of the exponential function, which occurs very frequently in physics; here with a negative exponent: $N(t) = N_0\,\mathrm{e}^{-t/\tau}$. N is the momentary number of individuals at the time t. Other examples of exponential behavior are the "barometric pressure formula" (Sect. 9.10) and the damping of harmonic oscillations (Sect. 11.10).

3. A water surface is covered with a non-wetting powder. Then, for example with a needle, one adds a tiny amount of a fatty acid at the center of the water surface. Immediately, the surface breaks up, and a clear-cut, circular spot appears, which is free of the powder.

Explanation: The surface tension of the water is greater than that of the fatty acid. Therefore, a droplet of a fatty acid on a water surface forms a monolayer whose thickness corresponds to that of a single molecule. If N molecules of a fatty acid, each with an area a, are placed onto a water surface, they will spread out to cover a circular area of $A = Na$. This allows us to determine the molecular cross-sectional area a if we know the number N of molecules added. This apparently modest demonstration is in fact quite important. – For measurements, one uses a rectangular water surface and replaces the powder by a floating movable, rectangular frame (Agnes Pockels, 1891)[C9.15].

C9.15. Agnes Pockels (1862–1935): cf. *Nature* **43**, 437 (1891). She was self-taught and wrote fundamental articles about films on liquid surfaces, for which she was granted an honorary doctoral degree (Dr.-Ing. E.h.) from the Technical University of Braunschweig in 1932.

[5] Reflections of light simulate the appearance of bridges between neighboring droplets in some places in the image.

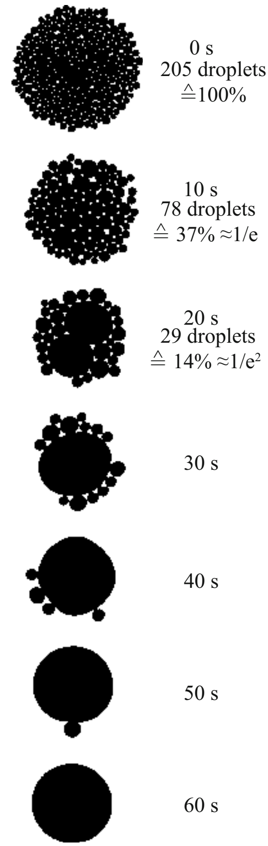


Figure 9.21 The coalescence of mercury droplets in alcohol containing a small amount of glycerin. This is a good example of a process which proceeds according to a statistical law: with a sufficiently large number of droplets (as in the upper three images), one can determine the mean lifetime to be $\tau = 10\,\text{s}$, i.e. after each 10 s, the number of droplets decreases to $1/e \approx 37\,\%$ of the preceding number (photographic images, each with an exposure time of $4 \cdot 10^{-3}$ s). The larger droplets are sometimes distorted because they have been set into oscillation by coalescing with smaller droplets. **(Video 9.6)** (**Exercise 9.8**)

Video 9.6:
"Coalescence of Hg droplets"
http://tiny.cc/mwqujy
The coalescence of small mercury droplets under water is demonstrated.

4. The addition of impurity molecules changes the value of the surface tension. This can be demonstrated with a grain of camphor and some water. The different faces of the grain dissolve at different rates; thus, the surface tension varies along different directions. The grain 'dances around' on the water surface. Similar processes play a role in the movements of microscopic life-forms (single-celled animals).

5. "Oiling the seas". An oil slick converts the "breakers" with their foam-covered crests into rolling swells. To produce the necessary change in surface tension, a ship need only release small amounts of oil in the form of droplets onto the ocean's surface.

When impurity molecules are involved, the phenomenon of surface tension becomes rather complex. The surface tension is then termed *abnormal*. That is, its magnitude becomes dependent on how much the surface area has already been increased, as with a rubber membrane. Furthermore, the increase in surface area is accompanied by warming. Kinetic energy is converted into heat. These potentially very interesting topics properly belong to *thermodynamics*.

9.6 Gases as Low-Density Liquids Without Surfaces. BOYLE's Law

The mass densities ϱ of gases are considerably less than those of liquids. As an example, in Fig. 9.22 (left side), we show a measurement of the density of room air, $\varrho = 1.29\,\mathrm{kg/m^3}$. It is thus only about 1/800th of the density of liquid water.

The molecules are the same in a gas and in the corresponding liquid. Therefore, the smaller density of the gaseous form can only result from greater average distances between the individual molecules. The following facts also support the hypothesis of large intermolecular distances in gases:

1. Gases, in contrast to liquids, have extremely high compressibilities (bicycle tire pump!). As a result, the density of a gas increases with increasing pressure. At $p = 160 \cdot 10^5\,\mathrm{Pa}$, we find for example for air a density of $\varrho \approx 200\,\mathrm{kg/m^3}$, i.e. about 1/5th of the density of water (Fig. 9.22, right side).

2. The BROWNian molecular motion can be observed in gases at a much lower magnification than in liquids. To provide visible fine particles, tobacco smoke can be conveniently used.

3. The molecules of a gas fly apart in complete disorder in all directions. They distribute themselves randomly in any available volume. Imagine for example a leaking gas-line in a room, or the gaseous, fragrant components of a perfume. In contrast to liquids, no attractive forces or cohesion of the molecules can be observed in gases without resorting to very sophisticated methods. In any case, gases do not form free surfaces. The cohesive forces between the molecules are

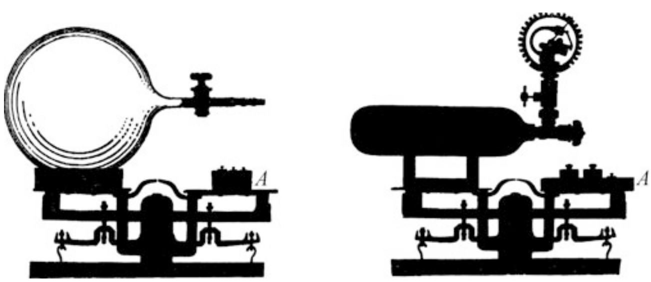


Figure 9.22 The dependence of the density ϱ of air on the pressure p. *Left-hand image*: $p = 10^5\,\mathrm{Pa}$ (normal atmospheric pressure), the glass balloon with $V = 7$ liter is evacuated and the balance is tared (with the weights A). Then we allow room air to flow into the balloon; in order to restore the equilibrium of the balance, we have to add 9 grams to the balance pan. Thus, $\varrho = m/V = 9\,\mathrm{g}/7$ liter $= 1.29\,\mathrm{kg/m^3}$. *Right-hand image*: $p = 160 \cdot 10^5\,\mathrm{Pa}$ in a steel pressure cylinder with $V = 1$ liter. After the compressed air is released, we have to remove 205 grams from the right pan, i.e. $\varrho = 205$ g/liter $= 205\,\mathrm{kg/m^3}$.

clearly not very effective at the large intermolecular distances typical of gases.

This completes our first overview. – The *relationship between pressure and density*, that is of pressure, mass and volume for gases, has been thoroughly investigated (Fig. 9.22), carefully keeping the temperature constant. The results of such measurements follow a simple relation over a large range of values; it is called BOYLE*'s law* (sometimes referred to as the BOYLE–MARIOTTE law, or MARIOTTE's law):

$$p = \frac{M}{V} \cdot \text{const}\,. \qquad (9.11)$$

In words: *The pressure p is directly proportional to the total mass M of the confined gas and is proportional to the reciprocal of the volume V of its container*. Two other formulations are somewhat more compact:

$$p = \varrho \cdot \text{const} \qquad (9.12)$$

and

$$pV_\text{s} = \text{const} \qquad (9.13)$$

($\varrho = M/V =$ mass density and $V_\text{s} = V/M =$ specific volume of the gas).

BOYLE'*s law is obeyed to a good – and often excellent – approximation by all gases at sufficiently high temperatures and low pressures*. This is demonstrated by the examples collected in Fig. 9.23: The product pV/M is represented over a large range of pressures and temperatures by horizontal lines parallel to the abscissa; it is thus independent of the pressure over the range shown. In this range of pressures and temperatures, substances in the gas phase are called *ideal gases*[C9.16]. If noticeable deviations from BOYLE's law occur, then one is dealing with a *real* gas. At normal pressures and temperatures, for example air, hydrogen, the noble gases, etc. all behave as *ideal* gases, while CO_2, NO_2, and chlorine behave as *real* gases. This distinction loses its validity at sufficiently low pressures and/or sufficiently high temperatures: Then all substances behave as *ideal* gases. BOYLE's law is thus a typical *limiting-case law*. Because of its importance, we will formulate its content once more in words: In an ideal gas, pressure and *mass density* are proportional to each other, or, equivalently, the product of pressure and *specific* volume is constant.

C9.16. Taking the temperature dependence into account leads directly from BOYLE's law (in the form of Eq. (9.13)) to the equation of state for ideal gases (Eq. (14.19) in thermodynamics).

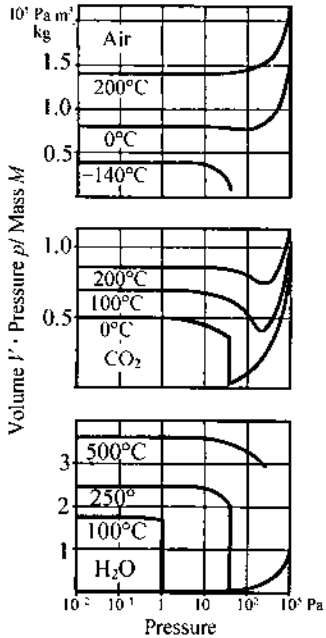


Figure 9.23 The horizontal straight lines are examples of the range of validity of BOYLE's law for ideal gases. The deviations which can be seen outside this range will be treated in Sect. 15.1. The vertical sections of the curves occur when part of the gas condenses (liquefies).

9.7 A Model Gas. Pressure Due to Random Molecular Motions (Thermal Motions)

The facts presented above can be well illustrated by a model gas. This will be shown in the present section and in the following sections. – As molecules, we again use steel balls, as in our tried and tested model for liquids. But now, we give these "molecules" much more free space within a voluminous "gas container". The latter is a flat box with glass side windows (Fig. 9.24). Furthermore, we now produce the random motions of the model molecules (the "thermal motions") by means of a vibrating steel piston *A*. It serves as one end wall of the gas container. The other end wall *B* is also a piston; it can be displaced with little friction, and together with a connecting rod and a helical spring *S*, it serves as a pressure meter (manometer).

Now, when the apparatus is in operation, all of the steel-ball molecules fly back and forth with a lively motion. They continually collide with each other and with the walls of the container. These collisions are *elastic*. Every "molecule" is constantly changing its velocity, in magnitude and direction. It presents a picture of truly *random* thermal motion.

This random molecular motion produces the pressure of the model gas against the walls of the container. We start by measuring this pressure using our manometer. *The pressure of a gas against its con-*

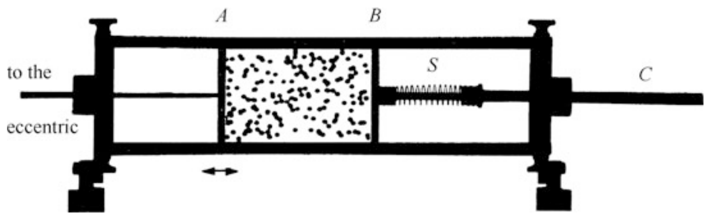


Figure 9.24 A model gas consisting of steel balls. The piston A ('ram') vibrates to produce "thermal motions", while the piston B on the right can be displaced along the tube C. Piston B and the helical spring S together form a manometer. The rod visible within the spring S can move freely within the tube C and serves to guide the piston B. **(Video 9.7)**

Video 9.7:
"A Model Gas: Its barometric density distribution"
http://tiny.cc/rwqujy
(see also **Video 16.1**, http://tiny.cc/xhgvjy)

tainer walls results from a quite different process than the pressure of a liquid. In the case of a liquid, the pressure is produced by "loading", e.g. by the weight of the liquid itself (gravity pressure), or by pressing a piston into a closed system containing the liquid (ram pressure). We made no mention of a liquid pressure resulting from the random collisions of its molecules against the walls of its container. *In this sense, gases exhibit a completely new phenomenon, related to their lack of cohesion and of free surfaces*. The gas molecules continually patter against the container walls. Every reflection of a molecule by a wall entails an impulse ($\int F\mathrm{d}t$) against the wall. The sum of all these collisions acts like a steady force of magnitude pA (A = area of the wall). The wall can remain at rest only if it is acted upon by an equal but opposite force, directed inwards towards the gas, produced for example by the spring S in Fig. 9.24.

9.8 The Fundamental Equation of the Kinetic Theory of Gases. Velocity of the Molecules

The origin of the pressure of a gas explained above can be described quantitatively. We need only fulfill one *postulate*: All N molecules are presumed to have the same kinetic energy, averaged over time; it is independent of the volume of the gas container: $E_{\text{kin}} = \frac{1}{2}mu^2$. Then with a brief computation (given in the fine-print paragraph that follows immediately), we can derive the fundamental equation of the kinetic theory of gases:

$$p = \frac{1}{3}\varrho\overline{u^2} \quad \text{or} \quad p = \frac{1}{3}\frac{\overline{u^2}}{V_\text{s}} \tag{9.14}$$

(p = pressure, ϱ = density, V_s = specific volume of the gas, $\overline{u^2}$ = mean value of the squared velocities of the molecules).

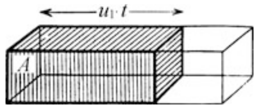


Figure 9.25 Derivation of the pressure of a model gas

Derivation

In Fig. 9.25, the gas container of volume V is supposed to hold a total of N molecules, each of mass m; the total mass of the gas is $M = N \cdot m$. Then the density of the model gas in the container is

$$\varrho = \frac{Nm}{V} = \frac{M}{V}. \tag{9.15}$$

We wish to compute the pressure acting on the left-hand side wall of the container (of area A). A molecule with the velocity u_1 travels a distance $s = u_1 t$ in the time t. As a result, only those molecules which are within the shaded volume, $As = Au_1 t$, can reach the left side wall within the time t. In the whole volume, there are N_1 molecules with the velocity u_1; then in the smaller shaded volume, only a number $Au_1 tN_1/V$ is present. The molecules fly around in a disordered manner. None of the six spatial directions ($\pm x$, $\pm y$, $\pm z$) is preferred. Therefore, only 1/6 of the molecules is moving on average in the direction ($-x$) towards the left side wall of area A. Thus, within the time t, only 1/6 of the molecules within the shaded volume will collide with the wall A, that is $\frac{1}{6}\frac{N_1}{V}Au_1 t$ molecules. To simplify the computation, we assume that these molecules hit the wall in a direction perpendicular to it. Then *each individual* molecule transfers an impulse $\int F_1 \mathrm{d}t = 2mu_1$ to the wall (Sect. 5.5), since the collisions are all *elastic*. The sum of all these impulses within the time t is

$$2mu_1 \frac{1}{6}\frac{N_1}{V}Au_1 t = \frac{1}{3}\frac{N_1 m}{V}Au_1^2 t. \tag{9.16}$$

We can replace this sum by an impulse $F_1' t$, which acts during the time t with the *constant* force F_1'. Then for the pressure produced by the N_1 molecules with the velocity u_1, we find

$$p_1 = \frac{F_1'}{A} = \frac{1}{3}\frac{N_1 m}{V}u_1^2.$$

We would obtain corresponding values p_2 for the N_2 molecules with the velocity u_2, and so forth. Finally, we add up all the partial pressures $p_1, p_2, p_3 \ldots$ from the N_1, N_2, $N_3 \ldots$ molecules with velocities u_1, u_2, $u_3 \ldots$ We set $p = p_1 + p_2 + p_3 \ldots$ and $N = N_1 + N_2 + N_3 \ldots$ and define $\overline{u^2}$ as the arithmetic mean of the squared velocities, so that

$$\overline{u^2} = \frac{(N_1 u_1^2 + N_2 u_2^2 + N_3 u_3^2 + \cdots)}{N}.$$

We then obtain

$$p = \frac{1}{3}\frac{Nm}{V}\overline{u^2}. \tag{9.17}$$

According to our postulate, the kinetic energy of a molecule is on average constant, and thus so is $\overline{u^2}$, the average value of the squared velocity. Furthermore, $Nm = M$, i.e. it is equal to the total mass of the gas in the container, and $Nm/V = M/V = \varrho$, the density of the gas. This yields with Eq. (9.17)

$$p = \varrho \cdot \text{const} \qquad (9.12)$$

This means that our simple model leads quantitatively to BOYLE's law! The constant likewise follows from Eq. (9.17); we obtain Eq. (9.14) as given above (A.K. KRÖNIG, 1856; a high-school teacher in Berlin).

Equation (9.14) makes it possible to calculate the root-mean-squared velocity of the gas molecules, defined as $u_{\text{rms}} = \sqrt{\overline{u^2}}$, from the corresponding values of the pressure p and the density ϱ of the gas[C9.17]. For air under normal conditions, for example, we find

$$p \approx 10^5\,\text{Pa}\,, \quad \varrho \approx 1.3\,\text{kg/m}^3\,.$$

C9.17. A precise statistical definition of u_{rms} will be given in Sect. 16.3 (Eq. (16.3) and Fig. 16.9). It will be abbreviated as $\overline{u}$ in later sections (e.g. Sect. 17.10).

Inserting these values into Eq. (9.14) yields for the velocity of the air molecules at room temperature $u_{\text{rms}} = 480\,\text{m/s}$. Likewise, for hydrogen at room temperature, we find a molecular velocity of $u_{\text{rms}} \approx 2\,\text{km/s}$. In terms of the order of magnitude, this calculation is certainly reliable. As we have emphasized, it yields *mean values*. The true (momentary) velocities of the molecules are distributed widely around the mean value (details in Sect. 16.3).

9.9 The Earth's Atmosphere. Atmospheric Pressure in Demonstration Experiments

The air, just like our model gas, expands to fill the available volume. Lacking a free surface, it has no fixed volume. How does the earth retain its atmosphere? Why do the air molecules not fly out into space? – Answer: Like all bodies, the air molecules are attracted towards the center of the earth by their *weight*. Each air molecule obeys the same conditions as a projectile (Sect. 4.9): To leave the earth, it would require a velocity of at least 11.2 km/s (escape velocity). The *average* velocity of the air molecules is much less than this value. As a result, the vast majority of the air molecules is bound to the earth by its weight.

Without their *thermal motion*, all the air molecules would fall to the earth like stones, and – incidentally – they would form a layer on the earth's surface of around 10 m thickness. Without their *weight*, they would immediately fly off the earth, never to return. The competition between thermal motion and weight however keeps the air molecules suspended and leads to the formation of the free air mantle of the earth, its atmosphere. The solid surface of the earth prevents them

Figure 9.26 Two Magdeburg hemispheres are being pulled apart by 8 (not 16!) horses **(Video 9.8)**

Video 9.8: "Magdeburg hemispheres" (Otto von Guericke's experiment)
http://tiny.cc/vwqujy
The entire lecture, although in German, can be seen at G. Beuermann, http://lichtenberg.physik.uni-goettingen.de.

from falling closer to its center of gravity. Therefore, the earth's surface carries the full *weight* of the air contained in its atmosphere. The quotient of this weight divided by the surface area is the normal *gravity pressure of the air*, for short 'air pressure' or 'surface atmospheric pressure'. It is around 1000 hPa $= 10^5$ Pa (this corresponds to the gravity pressure at the earth's surface of a mercury column 76 cm high).

"We humans lead a deep-sea life at the bottom of the enormous ocean of air". Every school child knows this today. The – at the time sensational – experiments carried out a few centuries ago to prove the existence of "air pressure" are today among the most elementary topics of school physics[C9.18]. Nevertheless, for reasons of historical deference, we will describe a classic demonstration. The mayor of Magdeburg, OTTO VON GUERICKE[6] (1602–1686), pressed together two copper hemispheres of 42 cm diameter with a greased leather seal and pumped out the air inside them. The hemispheres were then seen to be pressed firmly together by the air pressure of the surrounding atmosphere. We can compute the force as the product of the cross-sectional area of the hemispheres ($A \approx 1400\,\mathrm{cm}^2$) and the air pressure ($p \approx 10^5\,\mathrm{N/m}^2$): $F = 1.4 \cdot 10^4\,\mathrm{N}$. Thus, GUERICKE needed 8 horses to pull the hemispheres apart. The woodcut shown to a very small scale in Fig. 9.26 illustrates a demonstration of this famous experiment. The picture indeed shows not 8, but 16 horses; that was of course a bluff intended to impress the spectators, who were for the most part laymen in matters of physics. Eight of the horses could just as well have been replaced by a solid wall; since even back then, *force* equaled *counter-force*.

"Every school child knows this today."

C9.18. POHL often mentions school physics. That reveals what sort of previous knowledge was expected of beginning students during his lifetime; and from *all* students. His lectures were required not only of physics majors, but also of many non-majors including pre-med students. This highlights a problem that today's physics lecturers must deal with, namely the often very different levels of knowledge of elementary physics among beginning students, due to the different standards of their schools, even among natural-science majors.

"Since even back then, force equaled counter-force."

Today, the Magdeburg hemispheres still carry on a modest but useful existence in a puny form – that is as the well-known canning jars, consisting of a glass jar with a lid and rubber gasket. They are evacuated not with

[6] An excellent excerpt from his principal work, "*Nova experimenta (ut vocantur) Magdeburgica*", was published in 1912 by R. Voigtländer, Leipzig, as a German translation. No beginning physicist should miss reading this book. The experimental skills of GUERICKE and his descriptions, which strive for a clear simplicity, are exemplary.

Figure 9.27 Chain 'syphon'

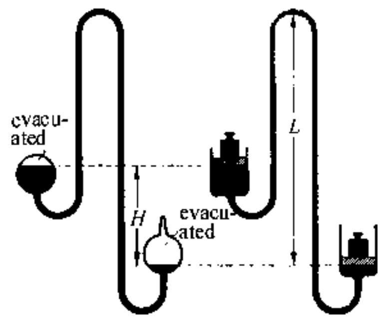


Figure 9.28 *Left:* A liquid syphon functions in a vacuum. *Right:* Loading of a liquid surface holds together a column of liquid, even when it contains gas bubbles.

a pump, but rather by using hot steam to displace the air (this also sterilizes their contents against anaerobic bacteria!). After cooling and condensation of the steam, a "vacuum" results.

In elementary physics classes, one often also encounters the well-known *syphon* as an effect of air pressure. However, *the syphon principle has nothing to do with air pressure.* It is explained in Fig. 9.27. A chain is hanging over a frictionless pulley; each of its ends is lying coiled up in a glass jar. When one of the jars is raised or lowered, the chain runs into the lower of the two jars; it is pulled down by the weight of its overhanging end H. Exactly the same principle applies to liquids, since they, like solid bodies, have a finite tensile strength (Sect. 9.5). As a result, a water syphon will function quite well in a vacuum, so long as no visible gas bubbles are present on the walls of the tube. A vacuum syphon of this type is shown in Fig. 9.28 (left side). The overhanging end of the stream of water is marked with its length H. *In principle, a liquid syphon functions entirely without any action of the air pressure.*

The liquids we encounter in our daily lives, especially water, are never completely free of small air bubbles. Therefore, a column of water containing air will separate easily. This difficulty can be overcome by loading the liquid surfaces equally on both ends, for example in principle by using frictionless pistons as in Fig. 9.28, right side. In practice, this loading can be most simply accomplished by making use of the pressure of the atmosphere. It can hold together

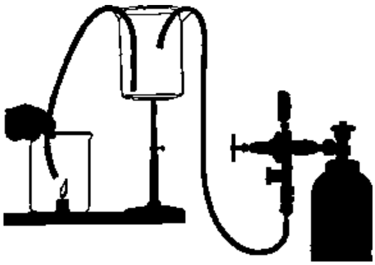


Figure 9.29 Gas syphon. On the right, a carbon dioxide cylinder with a pressure-reducing valve and hose fill the upper beaker with the heavy gas

a column of length $L \approx 10\,\mathrm{m}$ even when bubbles interrupt the column across the whole cross-section of the pipe containing it. 10 m is however only a small fraction of the column length corresponding to the tensile strength of gas-free water ($Z_{\max} = 3.4\,\mathrm{N/mm^2}$). This can be rather clearly seen from the fact that air pressure plays only a modest supporting role in the operation of a syphon, although this role is important in technical applications.

In the case of the *gas syphon*, the situation is quite different. Gases have no tensile strength. In contrast to liquids, gases can never form a *column* on their own. For this reason, gas syphons cannot operate in a vacuum. Figure 9.29 shows a gas syphon in operation. It permits the invisible gas *carbon dioxide* to flow through a syphon hose from the upper into the lower beaker. The arrival of the gas (and the resulting displacement of the air) in the lower beaker is indicated by a candle flame; it is extinguished by the CO_2 gas, which does not burn or support combustion.

The gas syphon clarifies the useful supporting role which the atmospheric pressure plays in many demonstration experiments: *Gases have no free surfaces, but the presence of the atmospheric pressure provides a certain substitute!* In place of the missing free surface, the *diffusion boundary* of the gas *vs.* the surrounding air acts as a 'surface'. Therefore, we can handle for example ether vapor just as though it were a liquid. We tip a bottle containing diethyl ether, keeping the tipping angle small, so that no liquid can pour out of the bottle. But we can see the ether *vapor* flowing out like a stream of liquid; this stream is especially noticeable in a shadow image ('schlieren').

We can also collect this flowing ether vapor in a beaker resting on a tared balance (Fig. 9.30). As the beaker is filled, the balance tips in the direction of "increasing weight". This is because ether vapor has a greater density than the air which is displaced from the beaker. At the end of the experiment, we can empty out the beaker by tipping it; again, we see the ether vapor flow out like a broad stream of liquid and drop to the floor.

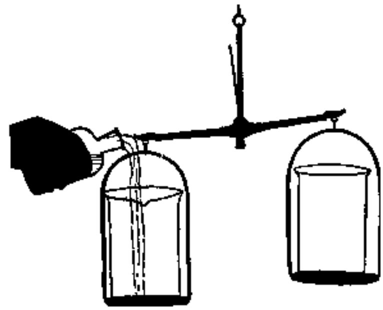


Figure 9.30 A stream of ether vapor in a shadow image

9.10 The Pressure Distribution of Gases Under Gravity. The Barometric Pressure Formula

Thus far, we have considered the atmospheric pressure of the air only on the surface of the earth. At sea level, it is practically constant and equal to 10^3 hPa, apart from small variations due to weather conditions. It is the same as the water pressure at the bottom of a freshwater pond which is 10.33 m deep.

In every liquid, the pressure decreases on going from the bottom of the container towards higher levels. This decrease is *linear* in the case of liquids. In water, the pressure decreases for example by about 100 hPa for every meter of increased height (Fig. 9.31). The reason for this is that lower layers of the liquid are not noticeably compressed by the weight of the layers above them; thus every layer of water of thickness dh makes the *same* contribution d$p = -\varrho g$dh to the overall pressure.

The story is quite different for gases. Gases are strongly compressible; the lower layers in a gas column are pressed together by the weight of the upper layers. The density ϱ within each individual

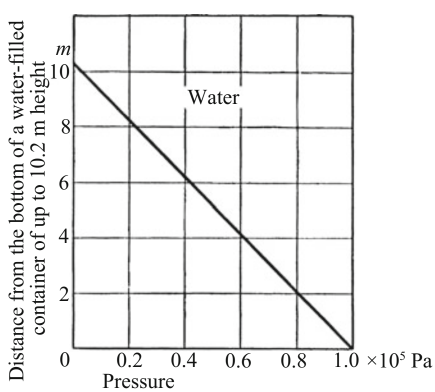
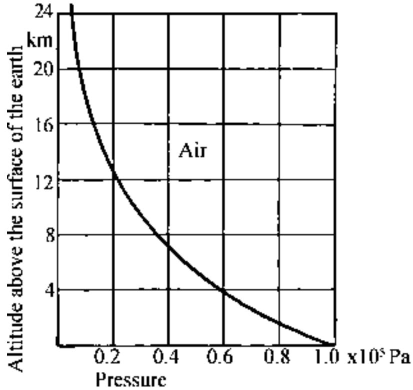


Figure 9.31 Distribution of the gravity pressure in a water column

Figure 9.32 Distribution of the gravity pressure in the air at a uniform temperature of 0 °C



layer is proportional to the pressure p acting on that layer. We thus find

$$\frac{\varrho}{\varrho_0} = \frac{p}{p_0} \quad \text{or} \quad \varrho = \varrho_0 \frac{p}{p_0}\,. \tag{9.18}$$

Here, ϱ_0 is the density of the gas at normal atmospheric pressure, p_0. Then the contribution to the pressure of each layer of gas of height dh is given by

$$\mathrm{d}p = -\varrho_0 \frac{p}{p_0} g\,\mathrm{d}h \tag{9.19}$$

$$(g = 9.81\,\mathrm{m/s^2}).$$

Integrating over height up to a maximum h yields:

$$p_\mathrm{h} = p_0 \mathrm{e}^{-\frac{\varrho_0 g h}{p_0}} = p_0 \mathrm{e}^{-\mathrm{const}\cdot h}\,. \tag{9.20}$$

Inserting the values which hold at a temperature of 0 °C, we obtain for the air pressure at a height h:

$$p_\mathrm{h} = p_0 \mathrm{e}^{-\frac{0.127h}{\mathrm{km}}}\,.$$

This *barometric pressure formula* is represented graphically in Fig. 9.32. This is the counterpart of the distribution of gravity pressure in a column of water as shown in Fig. 9.31.

The logic of this "barometric pressure formula" can be made intuitively clear by our model gas with steel balls. To show this, we set up the apparatus that we have seen already in Fig. 9.24 in a vertical position and look at it under stroboscopic illumination. On the projection screen, we then see a series of momentary images of the type shown in Fig. 9.33. In the lowest layers, we can see a large number of molecules, which rapidly decreases as we go upwards. We can recognize the competition between weight and thermal motions. Already at a height of 2 m (on the screen) above the vibrating piston, the

Video 9.7:
"A Model Gas: Its barometric density distribution"
http://tiny.cc/rwqujy

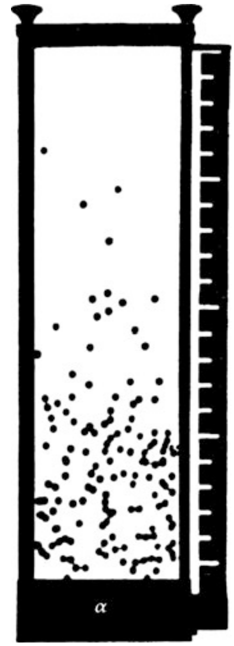


Figure 9.33 Snapshot image of a steel-ball model gas, demonstrating the barometric pressure formula (exposure time $\approx 10^{-5}$ s). **(Video 9.7)**

molecules are rather rare. Only the occasional molecule wanders up to a height of 3 m. *Our "artificial atmosphere" gradually thins out at increasing altitude, but without any discernible upper boundary.*

"But even at several hundred km above the surface of the earth, some molecules from the atmosphere can be found wandering about."

The situation in the earth's atmosphere is completely analogous; only the height scale is considerably greater[7]. A well-defined upper limit to the earth's atmosphere cannot be specified, any more than for our artificial atmosphere. 5.4 km above the earth's surface, the density of the air has decreased by about half ($e^{-0.69} = 0.5$); at about 11 km to 1/4, etc. (cf. Fig. 9.32). But even at several hundred km above the surface of the earth, some molecules from the atmosphere can be found wandering about[C9.19]. For even at this high altitude, we can observe the flashes from meteorites; they begin to glow when they enter the atmosphere. Auroras are also produced at a similar altitude; they arise when high-energy particles from the sun penetrate the earth's atmosphere and interact with the gas molecules in it.

C9.19. In fact, the exponential law is valid up to an altitude of ca. 100 km, where the pressure has dropped to about 0.03 Pa and the number density of the molecules is around $10^{19}/\mathrm{m}^3$. At still greater altitudes, the pressure decreases more slowly than predicted by the exponential law. At 300 km, it is around 10^{-5} Pa, with a number density of the order of $10^{15}/\mathrm{m}^3$ (*CRC Handbook*, 83rd edition (2002), pp. 14–19).

To conclude, we add some larger bodies to our artificial atmosphere, e.g. some wooden chips. They simulate dust particles in the air. We see the dust particles dancing around with lively "BROWNian molecular motions". But yet they always remain close to the "earth's surface"; for the weight of one of the wood chips is much greater than that of a steel-ball molecule.

[7] Furthermore, the composition of the atmosphere and its temperature also depend on the altitude. The true distribution of these quantities can be determined only through measurements. At high altitudes, Eq. (9.20) may fail even as an approximate description.

9.11 Static Buoyancy in Gases

From the results of the preceding section, we see that the gravity pressure in gases, as in liquids, decreases as we go upwards. Thus, there must also be a "buoyant force" in gases, as well as in liquids. As an example, we will describe the principle of the free balloon. It is drawn schematically in Fig. 9.34.

Formally, we could again apply the principle derived in Sect. 9.4: The buoyant force of the balloon is equal to the weight of the air which it displaces. But it is expedient to consider the distribution of pressure inside the balloon. This makes the situation intuitively clearer.

A free balloon is open at its bottom. There is no pressure difference between the air and the filling gas at their boundary. Naturally, this boundary is not completely sharp; it is simply a diffusion boundary between two gases. The effective pressure difference can be observed in the upper half of the balloon. There, the pressure of the filling gas on the inner side of the skin of the balloon is greater than the pressure of the air on its outer surface at the same height. Here is also where the bleed valve of the balloon is attached (*a* in Fig. 9.34).

The upwards-directed force acting on the skin of the balloon is proportional to the difference in density between the air and the filling gas. Both densities decrease with increasing altitude. For the filling gas, this decrease occurs within a loose balloon skin by gradually inflating the lower parts of the skin. When the filling pressure is exceeded, the excess filling gas escapes through the opening at the bottom of the balloon. As the densities decrease, the magnitude of their *difference* also decreases. At a certain limiting value of the densities, the upwards-directed force is equal to the weight of the balloon, and then the balloon remains suspended at a constant altitude. A further increase in altitude requires a reduction of its weight, e.g. by dropping off ballast.

The same pressure distribution as in a free balloon can be found in gas lines in buildings. Just like balloons, they are surrounded by air. Normally, the gas in the pipes is pressurized by a given level of ram pressure; sometimes, however, this pressure is too low. Then the gas doesn't "want" to flow out of an open valve in the basement. On the fifth floor of the building, however, this disturbance is not noticeable; when a valve is opened there, the gas streams out forcefully.

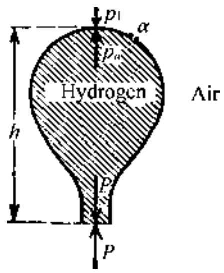


Figure 9.34 The buoyancy of a free balloon. The density ϱ_0 (Eq. 9.20) of the filling gas must be smaller than that of the surrounding air (compare Fig. 9.12), (**Exercise 9.10**). Note that the difference in pressure $(p_2 - p_1)$ is greatly exaggerated by the arrows.

Part I

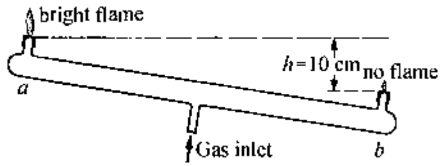


Figure 9.35 With increasing height, the gravity pressure of natural gas decreases more slowly than that of air (BEHN's pipe)[C9.20] **(Video 9.9)** (**Exercise 9.11**)

C9.20. U. Behn, *Phil. Mag*. **13**, 607 (1907).

Video 9.9:
"BEHN's tube"
http://tiny.cc/0vqujy

This situation can be elucidated by a demonstration experiment: Figure 9.35 represents the gas-pipe system as a glass tube. This tube has a small opening for a gas flame at each of its ends; the right-hand flame opening is 10 cm lower than the left-hand opening. We now pass natural gas (mainly CH_4) into the tube through a filling lug somewhere along its length; its flow is restricted by a throttle valve. Then we can readily light a flame at the upper opening a, but not at the same-sized lower opening at b. At this opening b, there is no pressure difference between the gas in the tube and the surrounding air. However, at a, 10 cm higher, there is a noticeable pressure difference, which allows us to ignite a bright flame. When the tube is held horizontally, so that both openings are at the same height, they can both be ignited. When the tipping is reversed, so that opening b is higher, a flame can be ignited only there. The setup is thus surprisingly sensitive – it does not show the decrease in pressure with increasing height, but rather only the *difference* in the decrease of the gravity pressure in an air column and in a column of natural gas.

Finally, we mention the further example of chimneys in houses and factories. They contain warm air and combustion gases with a lower density than the surrounding atmosphere. The higher the chimney, the greater is the pressure difference at its lower opening, and the better is its "draft" (**Exercise 9.9**).

9.12 Gases and Liquids in Accelerated Frames of Reference

Referring to our detailed treatment in Chap. 7, we can be brief here. We first give some examples for a radially-accelerated frame of reference. Thus, in this whole section, we adopt the point of view of an observer on a carousel or a swivel chair. Again, the swivel chair as seen from above is rotating in a counter-clockwise sense.

1. *Static buoyancy due to centrifugal forces. The principle of technical centrifuges.* On the carousel, we set a horizontal sealed box which is filled with water oriented in the radial direction (Fig. 9.36). Under its lid, a ball is floating; its density is thus less than that of water. When the carousel is rotating, the ball moves towards the center

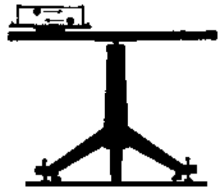


Figure 9.36 The centrifuge principle

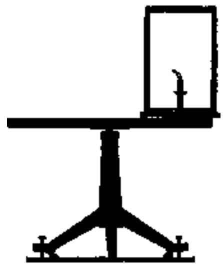


Figure 9.37 A flame under the influence of inertial forces

of the carousel (axis of rotation). Conversely, a ball which is lying on the bottom of the box (density greater than that of water) will move outwards, towards the rim of the carousel[C9.21].

C9.21. A further example of buoyancy in an accelerated frame of reference is a toy balloon in an airplane during takeoff or landing.

Explanation: The weight of the balls and their buoyancy due to the weight of the water they displace are compensated by the bottom and the lid of the box; the CORIOLIS forces are likewise compensated by its *side* walls. The only remaining force is the centrifugal force. It acts within the *horizontal* box just like the weight within a *vertical* box. For the centrifugal force, the axis of rotation at the center is "up", and the outer rim of the carousel is "down". An object in a liquid experiences a buoyant force in an "upwards" direction, i.e. here towards the axis of rotation. This buoyant force may be stronger or weaker than the centrifugal force which is acting on the object. When the latter predominates, the object will move towards the outer rim, that is, figuratively, it "sinks to the bottom". When the buoyant force predominates, it will on the other hand "rise to the top".

This static buoyancy in radially accelerated liquids forms the basis of technical centrifuges, for example for separating butter-fat or cream from milk. The fat, owing to its lower density, moves towards the axis of rotation (cf. Sect. 16.6, last part.)

2. *The deflection and curvature of a candle flame by centrifugal and* CORIOLIS *forces.* On the carousel, we set up a candle in a large glass box, carefully protected against air currents. The flame tips towards the axis of rotation (Fig. 9.37). In addition, it becomes curved towards the right as seen from above.

Explanation: The vector addition of the weight and the centrifugal force leads to a net force which slants *downwards and outwards*. The gases in the flame have a lower density than the air, therefore the force of *buoyancy* pushes them *inwards and upwards*. This buoyancy also gives the flame gases a *velocity*, and therefore, in addition to

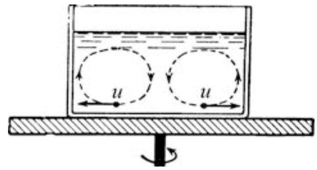


Figure 9.38 Radial circulation in a dish filled with liquid

centrifugal force, there will be a CORIOLIS force acting on the flame. It will cause the flame to curve towards the right.

3. *Radial circulation in liquids when individual liquid layers have different angular velocities.* In the center of our rotatable table, we place a flat dish filled with water (Fig. 9.38). Then we give the table a spin (constant angular velocity) and observe how the water gradually arrives at a stationary state. The water, driven by friction with the walls of the dish, gradually takes on the same angular velocity, initially in the neighborhood of the bottom and the side walls of the dish. As a result, at first only the water molecules near the bottom of the dish u are rotating and, driven by the centrifugal force (heavy arrows), they begin to flow outwards. This flow starts up the circulation shown by dashed lines. It can be readily verified using paper snippets placed at the bottom of the dish.

After a certain time, the molecules of the upper water layers also attain the full angular velocity of the dish, and then they also flow outwards. The dashed circulation is thus slowed, the water level in the center of the dish is lowered, while it rises at the outer circumference, until finally the stationary parabolic surface shape is established (see Figs. 9.2 and 9.3).

A reversal of this behavior can also be observed. In a teacup, stirring with a spoon initially gives the whole liquid the same angular velocity. However, friction with the bottom and sides of the cup reduces the angular velocity of the lowest layers as soon as one stops stirring. A radial circulation begins, but now in the opposite sense from that shown in Fig. 9.38. It causes the tea leaves which had been lying on the bottom of the cup to swirl towards the center.

This discussion of the circulation of water has already led us beyond the topic of liquids and gases at rest. This takes us through a smooth transition to our next chapter: Liquids and gases *in motion*.

Exercises

9.1 The two cylinders of a hydraulic press have diameters of $d_1 = 4\,\mathrm{cm}$ and $d_2 = 75\,\mathrm{cm}$. A force $F_1 = 10^{-1}\,\mathrm{N}$ acts on the smaller cylinder. Find the force F_2 which would have to press on the larger cylinder in order to maintain equilibrium. (Sect. 9.3)

9.2 The smaller piston of a hydraulic press has a cross-sectional area A_1, and the larger piston has an area A_2. The short end of a lever is connected to the smaller piston. A force $F_1 = 200\,\mathrm{N}$ is acting at the other end of the lever, whose long arm is 10x longer than its short arm. Find the area A_2 required if the press is designed to lift a block of mass $10^3\,\mathrm{kg}$. (Sect. 9.3)

9.3 Determine the buoyant force F acting on a lead ball of diameter d in water. (Sect. 9.4)

9.4 A hollow cylindrical buoy of height $l = 2\,\mathrm{m}$ and diameter $2r = 1\,\mathrm{m}$ is made of sheet iron with a thickness of $d = 0.5\,\mathrm{cm}$, and is designed to be held under water. Determine the required force F. Iron has a density of $\varrho = 7.8\,\mathrm{g/cm^3}$. (Sect. 9.4)

9.5 A rigid balloon of volume $1000\,\mathrm{m^3}$ is filled with hydrogen. The mass of the empty balloon skin and its basket is $m = 200\,\mathrm{kg}$. Find the force F which would be required to hold the balloon on the ground. (Densities: air, $1.3\,\mathrm{kg/m^3}$; hydrogen, $0.09\,\mathrm{kg/m^3}$.) (Sect. 9.4)

9.6 A hydrogen-filled, sealed, rigid balloon with a volume of 10 liters is hanging from a balance, which indicates a mass of 7 g (i.e. a weight of 0.0687 N). The air pressure is $1.013 \cdot 10^5\,\mathrm{Pa}$ (= 1 bar). Which mass m would the balance indicate if the air pressure were to decrease by 2.6 % to $0.986 \cdot 10^5\,\mathrm{Pa}$ while the temperature remained constant? (The density of air at the higher pressure is $\varrho = 1.3\,\mathrm{kg}/m^3$.) (Sect. 9.4)

9.7 Compare the energy which is necessary to bring a water molecule from the bulk to the surface of the liquid (surface work), with the energy which is required to expel a molecule from the liquid into the gas phase (energy of evaporation). The latent heat of vaporization of water (see Sect. 13.4 and Fig. 14.3) is $l_v = 2.5 \cdot 10^6\,\mathrm{W\,s/kg}$ at $T = 0\,°\mathrm{C}$ (Fig. 13.7), and the number density in the liquid phase of water is $N_V = 3.3 \cdot 10^{22}\,\mathrm{cm^{-3}}$. (Sect. 9.5)

9.8 In Video 9.6, "Coalescence of Hg droplets", find the number of mercury droplets as a function of time, $N(t)$. Show that $N(t)$ is an exponential function and find the average lifetime τ (note that in the video, the liquid is not alcohol, as mentioned in the text, but rather water). (Sect. 9.5)

9.9 A chimney of length $l = 30\,\mathrm{m}$ contains air at a temperature of $t_c = 200\,°\mathrm{C}$. The air outside the chimney has a temperature $t_o = 0\,°\mathrm{C}$ and a pressure of $p_o = 1.013 \cdot 10^5\,\mathrm{Pa}$. Calculate the pressure difference $p_o - p_c$, where p_c is the pressure inside the chimney at ground level. (This pressure difference drives the gas inside the chimney upwards. For simplicity, we assume in this calculation that the gas inside the chimney is air and that effects due to its flow are negligible.) Note: Take the value of the constant in BOYLE's law (Eq. (9.13)) from Fig. 9.23 at the two temperatures and compute the pressure differences $\Delta p = \varrho g \Delta h$ with $\Delta h = 30\,\mathrm{m}$, both inside and outside the chimney. (Sects. 9.6 and 9.11)

9.10 In Fig. 9.34, the upwardly-directed force which pushes up the balloon is explained in terms of the difference between the decrease in the barometric pressure of the surrounding air and that of the hydrogen inside the balloon from its bottom to its top surface. Calculate this pressure difference ($p_2 - p_1$ in Fig. 9.34), and compare the force that it produces with the static buoyant force (Fig. 9.12). To simplify the calculation, choose the shape of the balloon to be a cylinder with its symmetry axis upright, with its bottom surface open. Let its cross-sectional area be A and its height h. Neglect the weight of the empty balloon. (Sect. 9.11)

9.11 When BEHN's pipe (Fig. 9.35) is tipped so far that the lower flame nearly goes out, the pressure difference between the air outside and the gas inside the pipe at that position is practically zero, and the pressure is thus equal to $p_0 = 10^3\,\mathrm{hPa}$ ($= 1\,\mathrm{bar}$). Compute the pressure difference Δp between the gas and the surrounding air at the *upper* opening. Take the height difference to be $h = 0.1\,\mathrm{m}$, and assume that the gas is pure methane (CH_4). The densities of air and methane are $\varrho_{\mathrm{air}} = 1.3\,\mathrm{kg/m^3}$ and $\varrho_{\mathrm{methane}} = 0.72\,\mathrm{kg/m^3}$ (these values hold at 0 °C, which we assume for simplicity here). (Sect. 9.11)

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_9) contains supplementary material, which is available to authorized users.

Motions in Liquids and Gases

10

10.1 Three Preliminary Remarks

1. Liquids and gases differ in the formation of their surfaces and in their volume elasticity. Nevertheless, many aspects of the phenomena in both liquids and gases at rest can be treated in a similar manner. – One can go even further in unifying the treatment of motions in liquids and gases. Up to velocities of around 70 m/s, air can for example be considered as a non-compressible liquid; this velocity is still much smaller than the velocity of sound in air (340 m/s, cf. Sect. 14.10). *In this chapter, for brevity we will use the word "liquid" as a collective term*[C10.1]. *It includes liquids both with and without free surfaces.*

C10.1. In the literature, the word "fluid" is often used as a collective term for liquids and gases, and will be used in this book when *both* are expressly meant.

2. At high velocities, gases are compressed and their *temperatures* change. Processes of this type cannot be understood without the concepts of thermodynamics. They will therefore be discussed later, in Sect. 18.7.

3. In the mechanics of solid bodies, motions in basic experiments are perturbed quantitatively – more or less – by friction, but they are not changed qualitatively. We therefore treated friction initially as a side effect and only in Sects. 8.9–8.10 did we make some quantitative statements about it. – In the motions of liquids and gases, however, even the *qualitative* aspects of the phenomena are influenced decisively by friction. Therefore, we will treat them differently from the motions of solid bodies: We place a quantitative discussion of friction right at the beginning and treat motions in light of its crucial influence.

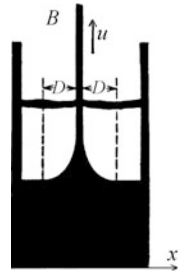
10.2 Internal Friction and Boundary Layers

Friction between solid bodies, i.e. *external* friction, is hard to grasp physically. In contrast, the friction which occurs within liquids, *internal* friction, can be rather clearly summarized. We demonstrate the essentials with two experiments:

In Fig. 10.1, we show a flat metal plate *B* which is being pulled slowly upwards from a glass basin filled with glycerine. Before the experiment was started, the lower half of the glycerine was colored

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_10

Figure 10.1 The two dashed lines show the two boundary layers of thickness D which have formed on either side of a moving plate B ('snapshot' image)



(e.g. with $KMnO_4$), so that at least *one* horizontal plane within the liquid was made visible (at the interface between the colored and uncolored liquid). One can imagine that a number of other horizontal planes could be similarly marked. During the motion, all of these planes on both sides of the moving plate are distorted within a broad region. This region is called the *boundary layer*. Its thickness D increases in the course of the motion. The innermost part of the boundary layer sticks to the moving solid object and moves with the same velocity u. The next layers, further out, are also pulled along, but their velocities become smaller with increasing distance from the plate. Thus, within the boundary layer, there is a velocity gradient, $\partial u/\partial x$.

When friction plays a role in a motion, one requires a force not only to accelerate the moving object up to its final velocity u, but also to maintain that velocity at a constant value (Sect. 5.11). In simple cases, this force $\boldsymbol{F}$ is proportional to the velocity $\boldsymbol{u}$, that is

$$\boldsymbol{F} = k\boldsymbol{u} \,. \tag{10.1}$$

The equally strong counter-force $F_2 = -F$ which is directed oppositely to u is the *frictional force* or resistance (in aerodynamics, it is also called 'drag'). The sum of these two forces $F + F_2$ is zero, so that the velocity u remains constant. The proportionality factor k can be expressed as a formula when the geometric situation is not too complicated. We will give some examples here and in the following section.

In Fig. 10.2, the boundary layer from the upwards-moving plate reaches out to the walls of the basin, so that their distance x from the plate is less than the thickness D which the boundary layer would have in a larger basin. In this case, the velocity gradient within the layer is practically linear; this is indicated in the figure by arrows whose lengths decrease uniformly with increasing distance from the moving plate. We find that k is proportional to the surface area A of the moving plate and inversely proportional to x, i.e. $k = \eta A/x$. The proportionality factor η depends on the liquid; it is a material constant and is called the *coefficient of viscosity* (or simply the "viscosity"; see Table 10.1). We then have

$$\boldsymbol{F} = k\boldsymbol{u} = \eta\frac{A}{x}\boldsymbol{u} \,. \tag{10.2}$$

Figure 10.2 The definition of the viscosity η for a planar flow. This means that the flow follows the same pattern in all the planes parallel to the page.

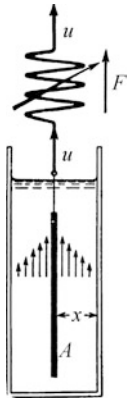


Table 10.1 The viscosities of some fluids

Substance	Temperature in °C	Viscosity η in N s/m^2
Air	20	$1.7 \cdot 10^{-5}$
CO_2, liquid	20	$7 \cdot 10^{-5}$
Benzene	20	$6.4 \cdot 10^{-4}$
Water	0	$1.8 \cdot 10^{-3}$
	20	$1.0 \cdot 10^{-3}$
	98	$0.3 \cdot 10^{-3}$
Mercury	−21.4	$1.9 \cdot 10^{-3}$
	0	$1.6 \cdot 10^{-3}$
	100	$1.2 \cdot 10^{-3}$
	300	$1.0 \cdot 10^{-3}$
Glycerine	0	4.6
	20	$8.5 \cdot 10^{-1}$
Pitch	20	10^7

The internal friction within liquids may be compared to the *shear forces* in solids. The quotient F/A could be termed the 'shear stress τ'. But there is a fundamental difference: The shear stress in solids increases with increasing strain or distortion. The internal friction in liquids, however, is proportional to the *velocity* of the distortion. Liquids at rest exhibit nothing at all comparable to a shear stress. In them, only normal stresses can occur (cf. Sect. 9.2).

The thickness D of the boundary layer can be estimated. One finds

$$D \approx \sqrt{\frac{\eta l}{\varrho u}} \tag{10.3}$$

(l = length of the moving body, ϱ = density of the liquid).

Derivation

The liquid within the boundary layer (Fig. 10.1) is being accelerated. It receives a kinetic energy $E \approx \frac{1}{2}mu^2$ (simplifying the velocity distribution). The accelerating force, the force F which is equal and opposite to the force of friction in Eq. (10.2) performs the work $W \approx \eta \frac{A}{D} ul$ along the length l. Setting E and W equal yields

$$\frac{1}{2}mu = \eta \frac{A}{D} l .$$

We replace the mass m of the accelerated liquid by the product of its volume AD and its density ϱ to obtain

$$\frac{1}{2}\varrho ADu = \frac{\eta Al}{D} ,$$

or, for the thickness of the boundary layer (with the prefactor $\sqrt{2}$), equation (10.3). – Numerical example for water: $\eta \approx 10^{-3}$ N s/m^2, $\varrho = 10^3$ kg/m^3, $l = 0.1$ m, $u = 10^{-2}$ m/s, $D \approx 3$ mm.

10.3 Laminar Flow: Fluid Motions Which Occur when Friction Plays a Decisive Role

The fluid motion observed in Fig. 10.2 is an example of a "laminar" (i.e. 'layered') flow. It is generally characterized by a velocity distribution which is constant over time, and it occurs when the velocity u is sufficiently small. We offer three additional examples:

First, we consider a liquid flowing through a *narrow tube* with a circular cross-section and length l. Due to friction, the velocity distribution within the liquid is not uniform. At the walls of the tube, the liquid is not moving; its velocity increases towards the center of the tube and is maximal there. Figure 10.3 shows an example. Just like the planar layers of liquid in Fig. 10.2, the coaxial liquid layers within the tube move past each other. The force necessary to overcome the frictional resistance is found in this case to be

$$F = k\,u_\mathrm{m} = 8\pi\eta l u_\mathrm{m} . \tag{10.4}$$

u_m is the mean value of the flow velocity, defined by the equations

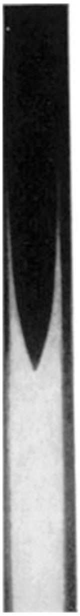
$$\text{Mean flow velocity } u_\mathrm{m} = \frac{\text{Volume current } \dot{V}}{\text{Cross-sectional area of the tube, } \pi R^2}$$

and

$$\text{Volume current } \dot{V} = \frac{\text{Volume } V \text{ that flows through the area } \pi R^2}{\text{Time } t \text{ required for this flow}} .$$

The force F is often produced by two unequal pressures p_1 and p_2 at the ends of the tube or pipe. We then have $F = \pi R^2(p_1 - p_2)$; and

Figure 10.3 The velocity distribution in laminar flow through a tube with an open cross-section of 6 mm × 6 mm[C10.2]



C10.2. In a tube of circular cross-section and also in a channel formed by planar parallel plates, in the stationary state of flow, one expects a parabolic velocity profile (see e.g. Noakes, Cath and Sleigh, "*Real Fluids. An Introduction to Fluid Mechanics*" University of Leeds (2009)). This is shown in Fig. 10.3, but only as an approximation. A possible explanation for this is the experimental difficulty of establishing precise initial conditions for the flow.

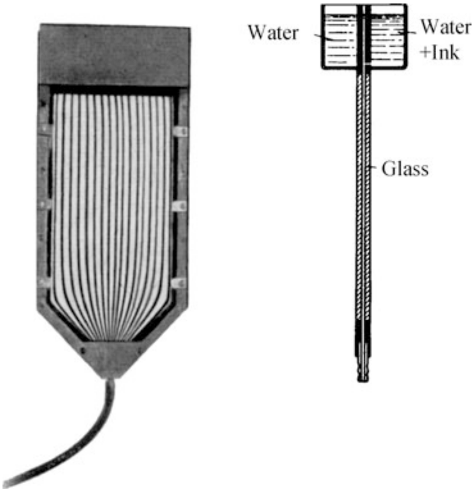


Figure 10.4 A streamline apparatus for demonstrating flow fields[C10.3], shown at the *left* as a plan view from above and on the *right* as a profile view. The upper chambers are connected through holes with the flat channel (the spacing of the glass plates is 1 mm). These holes are offset by one-half their spacing in the two chambers. – First, both chambers are filled with water, then some ink is added to the water in the right chamber. – *Left:* An example of parallel streamlines. For observing the shadow image on a screen, one can easily show the flow in a horizontal direction; two right-angle prisms can be used to rotate the optical ray path by 90°. **(Video 10.1)**

C10.3. The fluid flow shown in the "streamline apparatus" (Fig. 10.4) is a so-called HELE-SHAW flow, which occurs in flat channels as a laminar and practically two-dimensional flow field. The "streamline images" thus obtained (Fig. 10.5 and the later figures 10.10, 10.14, 10.16, 10.17 and 10.35) are the same as those of potential flows (i.e. ideal, frictionless flows; see e.g. Etienne Guyon, Jean-Pierre Hulin, Luc Petit, and Catalin D. Mitescu, *Physical Hydrodynamics*, Oxford University Press, 2nd edition (2015), Chap. 6.)

Video 10.1:
"Model experiments for streamlines"
http://tiny.cc/8wqujy

Part I

for the *volume current*, we obtain

$$\dot{V} = \frac{\pi}{8}\frac{R^4}{\eta}\frac{p_1 - p_2}{l}\,. \tag{10.5}$$

This equation is called the HAGEN-POISEUILLE *law*[C10.4], and it plays an important role in the physiology of our blood circulation.

The system of blood capillaries in the human body has a length of $\approx 3 \cdot 10^4$ km (this is of the order of magnitude of the circumference of the earth!). Increased muscular activity demands an increase in blood flow $\dot{V}$. This is very effectively accomplished by an expansion of the capillaries ($\propto R^4$!). The expanded network of capillaries must then be filled with blood; the required amount of blood is taken out of the *blood reservoirs* (in the splanchnicus region) (cf. H. REIN, *Physiology*)[C10.5].

As a second example, we replace the tube by a very *flat channel*, formed by two glass plates (of length l, width B, spacing of the plates $= d$). In a channel of this type, we can readily visualize the paths of individual liquid 'volume elements'. We dye the liquid and obtain an impressive image of the *streamlines* (Fig. 10.4). Quantitatively, we have[C10.6]

$$F = k\,u_\mathrm{m} = 3\eta\frac{Bl}{d}u_\mathrm{m}\,. \tag{10.6}$$

As a third example, we add a circular obstacle to this laminar fluid flow. The streamlines then show the picture reproduced in Fig. 10.5. If we imagine it to be three-dimensional, it would show the *laminar flow around a sphere* in a fluid (compare Fig. 5.21). If the sphere or ball (of radius R) is moving slowly enough at a velocity u relative to the fluid, then the required force is given by STOKES' law,

$$F = k\,u = 6\pi\eta R u\,. \tag{10.7}$$

When the fluid is at rest, F is usually the weight of the ball, reduced by its static buoyancy. Equation (10.7) finds many applications.

Examples
1. Measurement of the viscosity η.
2. Measurement of the radii of small spheres which are suspended in the air (droplets). This method is often more convenient than making microscopic measurements.
3. Without the frictional resistance of their tiny water droplets, the clouds would fall on our heads. As it is, they just sink quite slowly, evaporating from their lower sides and reforming at their upper sides.

C10.4. It was discovered independently by G. HAGEN (1797–1884), hydraulic engineer in Königsberg (published in *Poggendorff's Annalen der Physik und Chemie* **46**, 423 (1839)), and by J.-L.-M. POISEUILLE (1799–1869), a physician in Paris (*Comptes Rendus des Séances de l'Académie des Sciences* **11**, 961 and 1041 (1840)).

C10.5. POHL refers here to the physiology text by H. REIN. Modern edition: R.F. Schmidt, G. Thews, and F. Lang (eds.), *Physiologie des Menschen*, Springer-Verlag, 28th ed. (2000), Chap. 24. English: see e.g. *Textbook of Medical Physiology*, Guyton and Hall, 12th ed. by John E. Hall, Saunders-Elsevier, Philadelphia (2011).

C10.6. The derivation of Eq. (10.6), and of the other equations in this section, can be found for example in A. Sommerfeld, *Mechanik der deformierbaren Medien*, Akademische Verlagsgesellschaft, Leipzig 1957. English: *Mechanics of Deformable Bodies* (Lectures on Theoretical Physics), Arnold Sommerfeld, Academic Press (1950).

"Without the frictional resistance of their tiny water droplets, the clouds would fall on our heads."

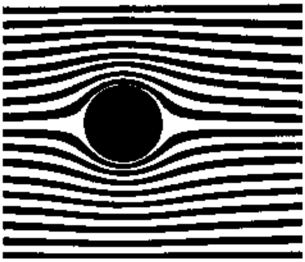


Figure 10.5 Laminar flow around a sphere or a cylinder (photographic positive image with bright-field illumination) **(Video 10.1)**

Video 10.1:
"Model experiments for streamlines"
http://tiny.cc/8wqujy

10.4 The REYNOLDS Number

The flow motions described in the previous section are observed only at sufficiently low velocities and/or dimensions. When velocities and dimensions are larger, the flow becomes *turbulent*. – *Turbulence* refers to a strong *mixing of the liquid*[C10.7] with swirling and eddies (vortices). It can be most simply observed using a colored stream of water in a transparent tube. Figure 10.6 shows a liquid flow before and after the onset of turbulence. The turbulent motions produce an additional viscosity, called the *apparent viscosity*, which can be orders of magnitude larger than the normal viscosity (Eq. (10.2)). Equation (10.4) is no longer applicable, and the required force F increases approximately as the square of the velocity.

C10.7. See for example S.B. Pope, *Turbulent Flows*, Cambridge University Press (2000).

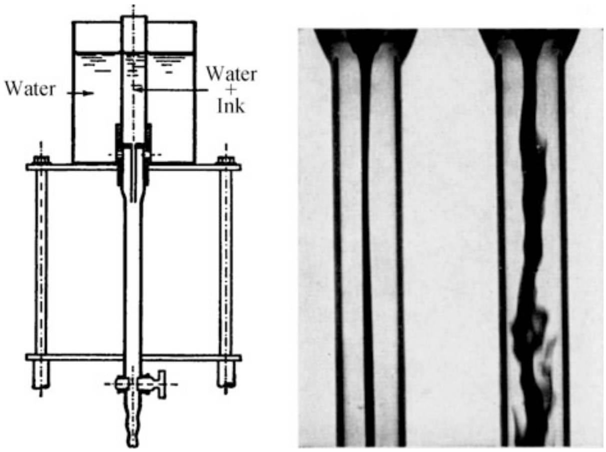


Figure 10.6 The formation of turbulence in a stream of water within a tube, and measurement of the REYNOLDS number as a demonstration. On the *left* is an apparatus suitable for shadow projection (tube: 15 mm × 15 mm clear opening). In its center is a laminar flow. On the *right*, a turbulent flow is shown. The flow velocity is computed from the amount of water that flows out, the time and the cross-sectional area of the tube **(Video 10.2)**.

Video 10.2:
"Turbulence of flowing water"
http://tiny.cc/cxqujy

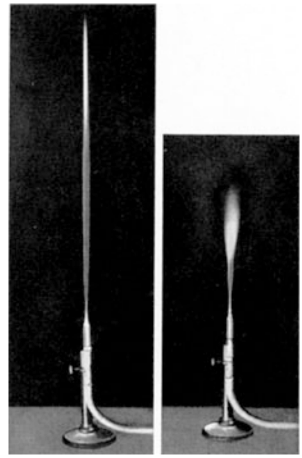


Figure 10.7 The "sensitive flame"; at the left laminar, at the right turbulent (and noisy). The onset of turbulence can be produced by making a soft hissing sound or shaking a key ring (at a distance of several meters).

The turbulence of a stream of gas which is burning as a "sensitive flame" (Fig. 10.7) is quite impressive.

The turbulent motions in the boundary layer between the earth and its atmosphere are familiar; they are known as *wind*. When the turbulence is strong, we speak of *gusts*. The altitude of the boundary layer can be several kilometers. In a *blizzard*, the turbulence is especially apparent[1].

In a turbulent flow, regions of the liquid which are randomly varying in terms of size and composition are combined into "individuals of higher order". During their lifetimes, which depend strongly on their sizes, they are subject to common motions and rotations. When they decay or shed off pieces, these can again combine into new, short-lived individuals.

The transition from laminar to turbulent flow is determined by a "critical" value of the ratio

$$Re = \frac{\text{Work of acceleration}}{\text{Frictional work}} = \frac{lu\varrho}{\eta}\,. \tag{10.8}$$

l is a length which is characteristic of the size of the body, e.g. the radius of a tube, boundary-layer thickness, etc.; u is the velocity of the liquid relative to a solid body; in a tube, for example, it is the mean value of the flow velocity as defined above; ϱ = density of the liquid; η = viscosity of the liquid. The ratio η/ϱ is often called the *kinematic* viscosity, and, to distinguish it, η is then called the *dynamic* viscosity.

[1] In the evening, the wind "goes to sleep" (but only near the ground!). – The reason: Turbulent motions lift cooler and therefore more dense air upwards, displacing the warmer air there downwards, with its lower density. Both require work, which is performed at the cost of the kinetic energy of the air, slowing the winds near the ground.

This was discovered in 1883 by O. REYNOLDS; therefore, *Re* is called the REYNOLDS *number*.

To derive Eq. (10.8), we make use of "dimensional analysis"[C10.8]. This means that we take all of the lengths which occur in the problem to be proportional to a length l which characterizes the size of the body. Furthermore, numerical factors are left off. For the work of acceleration, from Sect. 5.2, we have

$$W_\mathrm{a} = \tfrac{1}{2} m u^2 = l^3 \varrho u^2 \,. \tag{10.9}$$

The frictional work is found using Eq. (10.2):

$$W_\mathrm{r} = F\,l = \eta A \frac{u}{l} l = \eta l^2 u \,. \tag{10.10}$$

Taking the ratio of these two expressions then leads to Eq. (10.8).

C10.8. Details of the derivation of the REYNOLDS number can be found in textbooks on hydrodynamics, e.g. Etienne GUYON, Jean-Pierre HULIN, Luc PETIT, and Catalin D. MITESCU, *Physical Hydrodynamics*, Oxford University Press, 2nd edition (2015), Chaps. 1 and 2.

Small REYNOLDS numbers mean that the frictional work is predominant, while large numbers characterize the predominance of the work of acceleration. The ideal, frictionless fluid corresponds to a REYNOLDS number of ∞. – The "critical" values of the REYNOLDS number which lead to turbulence can be determined only from experiment[C10.9]. In smooth-walled tubes, *Re* must be greater than 1160 to cause turbulence.

C10.9. Also the "sufficiently low velocities" mentioned at the beginning of this section can be determined only from experiments.

For small spheres in air, $Re < 1$ is the condition for avoiding turbulence and guaranteeing the validity of STOKES' law (Eq. (10.7)). – In the streamline apparatus (Fig. 10.4), we are dealing with REYNOLDS numbers of around 10.
The air we breathe flows through our nasal channels without turbulence. In abnormally expanded noses, however, the critical values of the REYNOLDS number and the velocity can be exceeded, and then strong turbulence can occur, and it increases the frictional flow resistance. Noses which are internally abnormally large thus seem to be continually blocked.

The REYNOLDS number plays an important role in every quantitative treatment of fluid flow. Experiments on particular geometric shapes can be first carried out with models of convenient dimensions, and the results can then be scaled up to full-size dimensions. To this end, one must simply choose the flow velocity, density and dimensions to guarantee the same REYNOLDS number[C10.10]. For aircraft, the REYNOLDS numbers are on the order of several 10^7. This has an annoying consequence for measurements; it makes the study of technically important questions on small models more difficult. By lowering the temperature to that of liquid nitrogen (77 K) and increasing the working pressure to around 10 bar, the viscosity of the air can be decreased while simultaneously increasing its density. This allows realistic measurements on small models at around the same velocities as for the full-size aircraft.

C10.10. This holds as long as one can consider the fluid to be incompressible, i.e. the velocities must remain small compared to the velocity of sound in the fluid.

10.5 Frictionless Fluid Motion and BERNOULLI's Equation

From now on, we will follow the same path that we have already taken in discussing the mechanics of solid bodies: We first attempt to observe motion that is, as far as possible, free of the influence of friction; that is, we will try to reduce the effects of the boundary layer. For this purpose, we use a liquid container whose dimensions are large compared to the thickness of the boundary layer resulting from the motion that we wish to study.

A suitable "flow apparatus" is illustrated in Fig. 10.8. It consists of a basin which is 1 cm wide and filled with water. Aluminum flakes are added to the water as suspended particles. Objects of various shapes (profiles) can be moved through the basin, loosely touching the glass walls. In Fig. 10.8, we see an object with a circular profile, and in Fig. 10.9, there are two objects, *a* and *b*, which are held by invisible rods and together form a bottleneck. For photographic images, the basin is moved along a track at a constant velocity[2]. The Al flakes show the magnitude and direction of the flow velocity at every moment throughout the whole basin in the projection. In an exposure of around 0.1 s, the path of each flake is seen as a short streak. Each of these streaks is practically a straight line and represents, briefly stated, the velocity vector of a single water 'volume element'. With longer exposure times, the streaks combine to show *streamlines*. They indicate the overall pattern of velocity directions, that is the *flow field*. – the image of the flow can be stationary, i.e. in-

Video 10.3:
"Fluid Flow around obstacles"
http://tiny.cc/fcgvjy

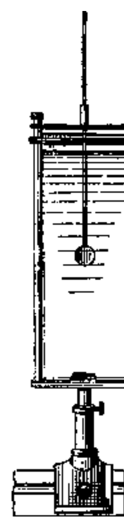


Figure 10.8 Flow apparatus. Here, again, for projection onto a wall or screen, it is advisable to rotate the image by 90°, as e.g. in Figs. 10.9, 10.26, 10.32 and 10.34. **(Video 10.3)**

[2] For observations projected on the wall, the fixed setup of the basin sketched in Fig. 10.8 is sufficient. The eye of the observer *follows* the object, and thus sees the liquid flowing past it.

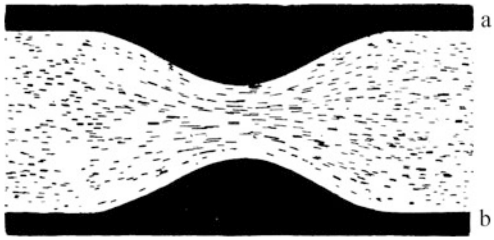


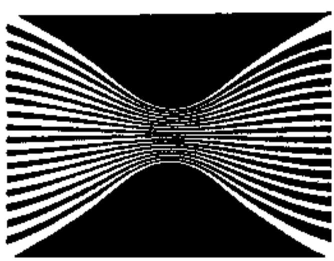
Figure 10.9 Streamlines passing through a bottleneck; the observer (camera) and the tube are at rest, while the liquid flows through the latter

dependent of time, or at a fixed position. Then the streamlines show in addition the whole path followed by the volume elements of the liquid *one after another*.

The time-exposure image shows the flow field in its clearest form, as seen in Fig. 10.9. The image projected onto a wall or screen is more lively. Often, however, one is striving for an image without too many details, with only a few clear streaks. In this case, a strange circumstance comes to our aid: We can imitate the flow field of a stationary flow which is practically free of the effects of friction in a *model experiment*. The streamline apparatus which we have already seen in Fig. 10.4, with its laminar flow, can serve this purpose. In spite of the completely different conditions for producing the flow, the course of the streamlines is that of an ideal, frictionless liquid flow. Figure 10.10 shows an image made in this way. It corresponds to Fig. 10.9. But in contrast to Fig. 10.9, it shows only a *model* experiment; we should keep that in mind. Formally, however, the image is accurate, and its simplicity makes it clear and easy to remember.

These liquid flow patterns, practically free of the effects of friction, which we can illustrate with such model experiments, can be maintained for only very short times. They are roughly analogous to the example from the mechanics of solid bodies of a ball moving free of forces at constant velocity; this is an idealized limiting case. But an important law holds for this case, which forms the basis for all that follows. It concerns the "static" pressure, i.e. the pressure of the liquid against a surface placed parallel to its streamlines. In a first, qualitative form, this law can be formulated as follows:

Figure 10.10 Streamlines in a model experiment (positive photographic image in bright-field illumination, like Figs. 10.14, 10.16, 10.17 and 10.35) **(Video 10.1)**



Video 10.1: "Model experiments for streamlines"

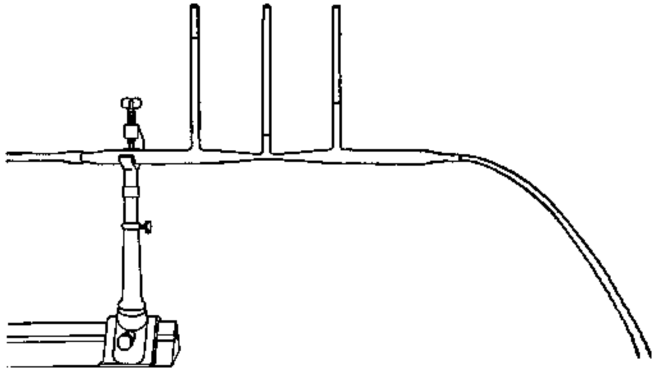


Figure 10.11 Distribution of the static pressure around a flow through a bottleneck or waist. The three vertically-mounted glass side tubes serve as water manometers.

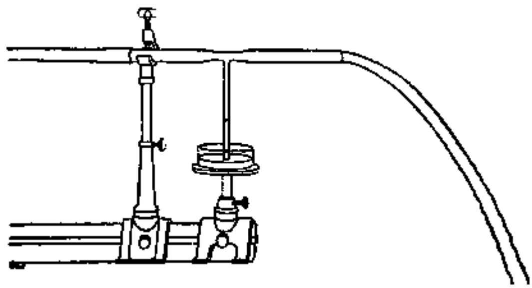


Figure 10.12 The static pressure in a waist. It is lower than atmospheric pressure in this case. A mercury column serves here as manometer.

In regions where the streamlines are pressed together, or the flow velocity is increased, the static pressure p of the fluid is lower than in the surrounding regions.

In order to make this law intuitively clear, the two experiments shown in Figs. 10.11 and 10.12 are useful[C10.11]. Figure 10.11 shows the static pressure of the flowing liquid *in front of, within* and *behind* the bottleneck. The figure is schematic. The diameter of the tube is not yet large compared to the thickness of the boundary layer (Sect. 10.2); the influence of friction is therefore only partially reduced. As a result, the static pressure behind the bottleneck does not attain precisely the same value as before it. – In Fig. 10.12, a considerably higher flow velocity is shown. At this velocity, the static pressure of the water in the bottleneck is less than the surrounding air pressure. The flowing water can "suck in" mercury within a manometer and pulls up a mercury column of several centimeters in height ("jet pump" principle).

C10.11. Here, and in the following experiments in this section, we are dealing with laminar, stationary flows.

The quantitative relation between the pressure and the velocity of the flow is found by applying the law of conservation of energy. We consider a certain quantity of liquid of mass m, volume V and density ϱ. Its static pressure and its velocity before the bottleneck are p_0 and u_0; within the bottleneck, they are p and u. The liquid must be *accelerated* from u_0 to u on entering the bottleneck, in order to pass through it. This requires the work

$$V(p_0 - p) = \tfrac{1}{2}m(u^2 - u_0^2)\,, \tag{10.11}$$

or, after dividing by the volume V,

$$p + \tfrac{1}{2}\varrho u^2 = p_0 + \tfrac{1}{2}\varrho u_0^2 = \text{const.} \tag{10.12}$$

$\frac{1}{2}\varrho u^2$ is added here to the pressure p; it must therefore itself represent a pressure. It is called the *dynamic* or *stagnation pressure*. The sum on the right of the equation is constant. It must also represent a pressure, and it is called the 'dynamic head' or *total pressure* p_1. We have thus obtained the important BERNOULLI equation,

$$\underset{\text{static pressure}}{p} \quad + \quad \underset{\text{stagnation pressure}}{\tfrac{1}{2}\varrho u^2} \quad = \quad \underset{\text{total pressure}}{p_1}\,. \tag{10.13}$$

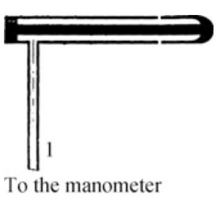
Measurement of the static pressure p in a flowing liquid is accomplished by the setup shown in Fig. 10.12. The opening which leads to the manometer lies *parallel* to the streamlines of the fluid flow. For measurements *within* wide flow paths, one places the opening, usually in the form of a sieve or slit, in the side of a *pressure probe*. It is connected to a manometer via a hose or tube. This is illustrated in Fig. 10.13.

The total pressure p_1 is measured in a *stagnation region*. One is illustrated in the model experiment in Fig. 10.14: *At the center of the stagnation region* (stagnation point), *a streamline strikes an obstacle in a perpendicular direction*. The connection to the manometer (PITOT tube) is made at this point. Here, the liquid is at rest, that is $u = 0$. The static pressure is, from Eq. (10.13), equal to the total pressure p_1 at this point. The manometer indicates the total pressure p_1.

The stagnation pressure is determined as the difference of the total pressure p_1 and the static pressure p. The stagnation pressure, from Eq. (10.13), is given by

$$\tfrac{1}{2}\varrho u^2 = (p_1 - p)\,.$$

Figure 10.13 Sectional view of a pressure probe with a ring-shaped slit for measuring the static pressure within a flowing liquid



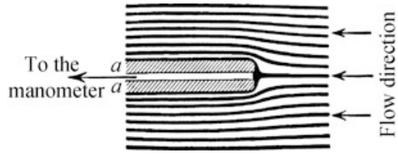


Figure 10.14 A PITOT tube for measuring the total pressure in a stagnation region; *in natura*, the tube is bent at a right angle, usually with an outer diameter of only 2 to 3 mm (model experiment with the streamline apparatus (Fig. 10.4); the contours of the tube were shaded after exposing the photo)

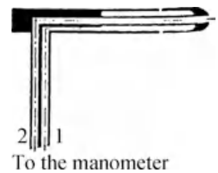


Figure 10.15 Cross-setional view of a PRANDTL tube, a combination of a PITOT tube and a pressure probe. The liquid-column manometer which is attached on both sides to the tubes *1* and *2* indicates the stagnation pressure directly as the difference of the total pressure p_1 and the static pressure p.

The total pressure p_1 is measured with a PITOT tube, and the static pressure p with a pressure probe. For technical measurements, it is expedient to combine the two devices ("PRANDTL tube", Fig. 10.15). Stagnation-tube measurements are the preferred method for determining the velocity in flowing fluids.

Figures 10.11 and 10.12 elucidate the decrease of the static pressure p with increasing flow velocity u. Numerous other demonstration experiments can also show this. We give two examples: In each one, the flow field is imitated by a model using the streamline apparatus (Fig. 10.4, HELE-SHAW flow).

1. Figure 10.16 shows a circular disk in model experiments with three different orientations within a flow field. Even the slightest tilting causes an asymmetry in the distribution of the static pressure, which produces a torque[C10.12]. When the tilting is stronger (Fig. 10.16b), the torque can be readily seen: The regions of expanded streamlines press against the disk on one side, while the regions of compressed streamlines pull on the opposite sides. In the real experiment, the disk in Fig. 10.16b would be rotated in a clockwise sense. The first position (Fig. 10.16a) proves to be labile; the disk swings back and forth, until it reaches a stable orientation perpendicular to the flow (Fig. 10.16c). We can observe this by letting let a stiff sheet of paper fall to the floor.

C10.12. An example of an application is the RAYLEIGH disk (Sect. 12.24).

2. Two spheres are moving within a fluid. The line connecting their midpoints is perpendicular to the direction of the unperturbed streamlines. The model experiment shown in Fig. 10.17 indicates an increased flow velocity u between the two spheres. Therefore, the static pressure p between the spheres is reduced, and they "attract" each other via "hydrodynamic forces". Figure 10.18 shows a similar experiment. In a water basin, the wooden ball at left stands upright

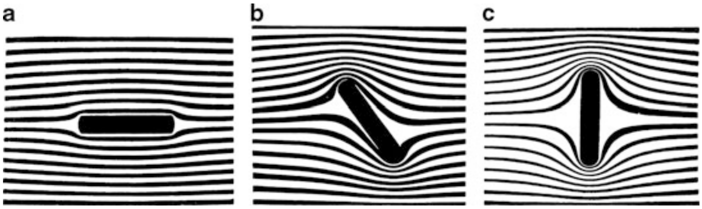


Figure 10.16 Three model experiments on the flow around a disk. The disk and the observer are at rest, the liquid is flowing. In the center image, one can see the two stagnation points. This demonstrates the occurrence of a *torque* around the center of gravity of the disk. **(Video 10.1)**

Video 10.1: "Model experiments for streamlines" http://tiny.cc/8wqujy (see also Comment C10.6 on Fig. 10.4)

Figure 10.17 The streamlines between spheres or cylinders; model experiment

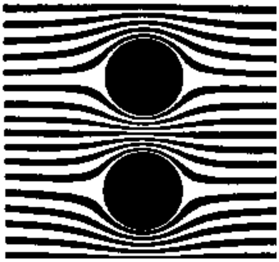
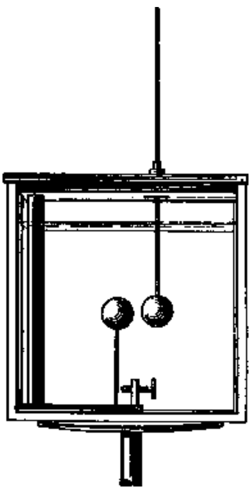


Figure 10.18 The attraction of a ball at rest and a moving ball in water



as an inverted gravity pendulum. A second ball is moved past the first one at a certain distance by a pushrod. The attraction of the two wooden balls is clearly visible. A buffer (at bottom) keeps them from colliding. – We can imagine two ships passing each other as a large-scale version of the two balls. In narrow navigation channels, for example in a canal, there is always a danger that the ships will be pulled together and collide[C10.13]. It can be reduced only by proceeding very slowly; for, as seen in Eq. (10.13), the stagnation pressure $\frac{1}{2}\varrho u^2$ increases as the square of the velocity, and with it the attractive force between the balls (or ships).

C10.13. The same attraction affects a ship which passes too close to the bank of a canal.

Part I

10.6 Flow Around Obstacles. Sources and Sinks. Irrotational or Potential Flows

In the flow fields that we have considered thus far, it is clear that two different flows overlap. First of all, there is a *parallel flow* of the liquid without the object placed in it (as shown in model form in Fig. 10.4); secondly, after the object is placed in that flow, there is an additional flow around it. This additional *avoidance flow* (or 'flow around obstacles') can be observed *by itself*. One need only modify the method of observation: Up to now, we have considered the object and the observer (camera) to be at rest, and the liquid was allowed to flow past them. We now adopt the opposite possibility: The liquid (basin) and the observer are at rest, but the object is moving. (For projection of the observations, one makes use of limited back-and-forth motions.) – With this second method of observation, we find for example the left-hand part of Fig. 10.19 instead of Fig. 10.16c, and the right-hand part instead of Fig. 10.16b. Corresponding images for the avoidance flow around a sphere and a cylinder can be seen in Figs. 10.20 and 10.21. In the direction of motion, the edges of the object are washed out and appear as half-tones in the pictures. This was repaired after the fact by shading in the figures. The streamlines begin at *one* of the shaded regions; there are *sources* there. They end at the other shaded region, where there are *sinks*. The flow fields move along with the objects; they are thus no longer stationary.

The flow field of a single point source (+) or sink (−) has spherical symmetry; a cross-section is sketched in Fig. 10.22. The two shaded areas denote the same small volume in two positions which follow each other in time. The liquid thus flows in a radial direction and *without rotating*. – The source is assumed to release a liquid volume V within the time t. The quotient $V/t = q$ is called its *productivity*[C10.14]. For a source, the productivity has a positive sign; for a sink,

C10.14. The 'productivity' q mentioned here again refers to a volume current, as already defined in Sect. 10.3. Its connection to the flow velocity u leads directly to Eq. (10.14).

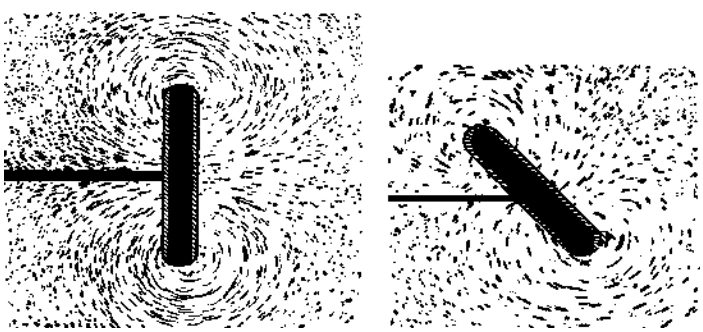


Figure 10.19 The parallel flow around a perpendicular and a tilted plate; observer and liquid are at rest, and the plate is moving **(Video 10.3)**

Video 10.3:
"Fluid flow around obstacles"
http://tiny.cc/fcgvjy

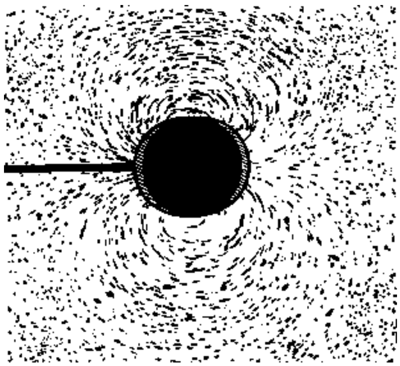


Figure 10.20 Avoidance flow around a sphere; observer and liquid (basin) are at rest, and the sphere is moving

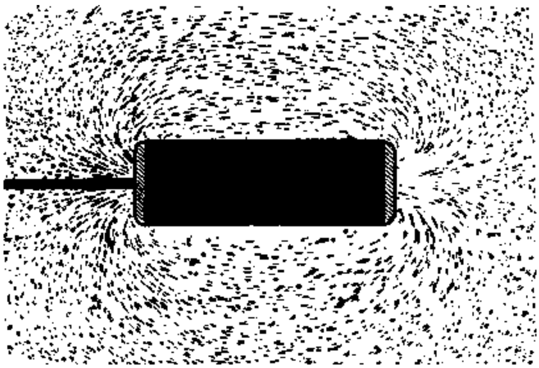


Figure 10.21 The avoidance flow around a cylinder which is parallel to a parallel flow; the observer and the liquid (basin) are at rest, and the cylinder is moving

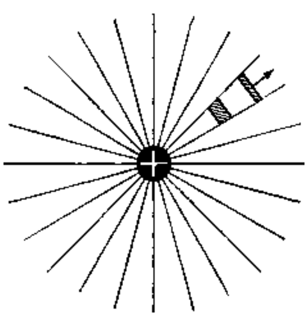


Figure 10.22 The flow field of a source (or a sink, with reversed directions)

a negative sign. Then for the velocity u of a volume element at a distance r from the source or sink, we find

$$u = \frac{\pm q}{4\pi r^2} \,. \tag{10.14}$$

Irrotational flow fields can be combined by simple superpositions. This provides a considerable simplification in their mathematical treatment. Thus, in Fig. 10.23, the two radially-symmetric

Part I

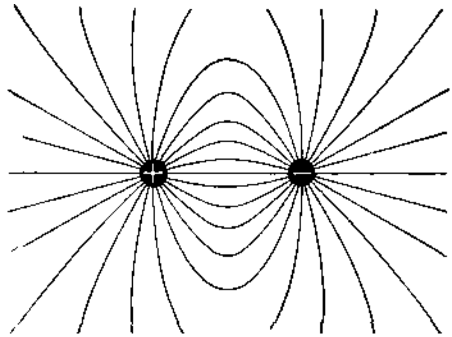


Figure 10.23 The flow field of a source and a closely neighboring sink (a "dipole")

fields of a source (+) and of a neighboring sink (−) have been superposed[C10.15]. The flow field which results is called a *dipole field*. It has many applications. At large distances, the flow fields of Figs. 10.19 through 10.21 are the same as a dipolar flow field. They can all be replaced by dipole fields.

C10.15. Since the velocity is a vector, flow (velocity) fields are vector fields, and their superposition is subject to the rules for vector addition. Thus, for example, the image of the field lines of a dipole in Fig. 10.23 is formed by superposition of the fields of a source and a sink.

We will encounter Eq. (10.14) again later when we are dealing with electricity and magnetism (Vol. 2). Then, it will not contain a dependence on a velocity u, but rather that of an electric or magnetic field on a distance r. Instead of the productivity $\pm q$ (m^3/s), the electric charge $\pm q$ (in ampere seconds) or the magnetic flux $\pm \Phi$ (in volt seconds) will enter the equations. Therefore, the streamline patterns of the avoidance flow are formally just the same as the field-line patterns in electrodynamics. Figure 10.21 is thus similar to the magnetic field lines from a long coil through which an electric current is flowing, and Fig. 10.19 (left) is like the electric stray field of a parallel-plate condenser (see Vol. 2, Chap. 2, Fig. 2.5). Similarly, Fig. 10.20 resembles the field of an electrically- or magnetically-polarized sphere.

All these vector fields are characterized by a potential field from which they can be derived by taking its gradient. The flow field of a frictionless liquid is also such a vector field; we thus refer to it as a *potential flow field*[C10.16].

C10.16. A good introduction to the theory of potential flows, which describes the flow of frictionless fluids, can be found for example in R.P. Feynman *et al.*, *Lectures on Physics*, Addison-Wesley, Reading, Massachusetts, U.S.A. (1964), Vol. II, Chap. 40. They can be read online at http://www.feynmanlectures.caltech.edu/.

10.7 Rotations of Fluids and Their Measurement. The Irrotational Vortex Field

Within a solid body, all its parts (volume elements) are bound rigidly together. This has three consequences: First: The shape of an arbitrarily chosen volume element within the body remains *unchanged* during motions. Second: Every point within a volume element has the *same* angular velocity ω. Third: The rotation of *all* the elements is uniquely defined by their *common* angular velocity ω.

In a liquid, on the other hand, all the volume elements can flow freely past each other. This leads to quite different consequences from the case of solid bodies: First, marked regions (e.g. colored with a dye) within a liquid *change* their shapes during motions[3]; think for example of Fig. 10.1. Second, different points within a volume element can have *differing* angular velocities. Therefore, we can *not* define the rotation of such an element by quoting a common angular velocity, as in solids. Instead, we must introduce a *new* measure for the rotation of a volume element within a fluid. It must summarize the motion by defining a reasonable mean value for the different angular velocities within the volume element. The measure of rotation used for fluids is called the *curl of the path velocity* $\boldsymbol{u}$, or more simply, "curl $\boldsymbol{u}$".

The *experimental* definition of the curl is simple: One puts a little *float* marked with an arrow into or onto the liquid, choosing its diameter to be *small* compared to the radius of curvature of its orbit. During the motion, the arrow on the float will change its direction with an angular velocity $\boldsymbol{\omega}_{\mathrm{fl}}$. Then we define[C10.17]

C10.17. The factor 2 in Eq. (10.15) has no physical meaning. It simply results from the mathematical definition of the curl.

$$2\,\boldsymbol{\omega}_{\mathrm{fl}} = \operatorname{curl} \boldsymbol{u}\,. \tag{10.15}$$

To obtain the *mathematical* definition of the curl, one starts with the *circulation* Γ. This is defined as the path integral of the orbital velocity $\boldsymbol{u}$ along an arbitrary closed curve, i.e.

$$\Gamma = \oint \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{s} \tag{10.16}$$

(the circle on the integral sign indicates a closed path or loop).

One then supposes that the path encloses a surface element dA, and computes the quotient $\mathrm{d}\Gamma/\mathrm{d}A$ for this limiting case. This defines a new *vector* which is perpendicular to the surface element; it is called the *curl* of the orbital velocity $\boldsymbol{u}$: curl $\boldsymbol{u}$. It describes the rotation of the fluid within this surface element. If the surface element lies e.g. in the xy plane, then we find for the z component of the curl

$$(\operatorname{curl} \boldsymbol{u})_{\mathrm{z}} = \left(\frac{\partial u_{\mathrm{y}}}{\partial x} - \frac{\partial u_{\mathrm{x}}}{\partial y}\right). \tag{10.17}$$

Derivation
Referring to Fig. 10.24, we calculate the circulation around the z axis along the four sides of a rectangular surface element $\mathrm{d}A = \mathrm{d}x\ \mathrm{d}y$. The order of the summation is in the clockwise sense as seen by an observer looking along the positive z-axis direction. The circulation then consists of four individual terms, namely

$$\begin{aligned}\mathrm{d}\Gamma &= u_{\mathrm{x}}\mathrm{d}x + \left(u_{\mathrm{y}} + \frac{\partial u_{\mathrm{y}}}{\partial x}\mathrm{d}x\right)\mathrm{d}y - \left(u_{\mathrm{x}} + \frac{\partial u_{\mathrm{x}}}{\partial y}\mathrm{d}y\right)\mathrm{d}x - u_{\mathrm{y}}\mathrm{d}y\\ &= \mathrm{d}x\mathrm{d}y\left(\frac{\partial u_{\mathrm{y}}}{\partial x} - \frac{\partial u_{\mathrm{x}}}{\partial y}\right) = (\operatorname{curl}\boldsymbol{u})_z\,\mathrm{d}A\,.\end{aligned}$$

[3] With the exception of the special case treated in Eq. (10.20).

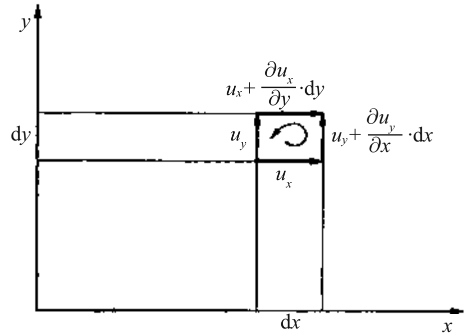


Figure 10.24 The derivation of Eq. (10.17); the z direction points from the plane of the page towards the eye of the reader

The curl of the orbital velocity in its general (3-dimensional) form is a somewhat difficult concept. Therefore, we offer several examples of its application:

In Fig. 10.2, we see the *boundary layer* of a planar flow field; the volume elements of the liquid move along *straight-line* paths. u_y is their upwardly-directed velocity (called u in the figure); its horizontal component is $u_x = 0$. Then equation (10.17) yields

$$(\mathrm{curl}\,\boldsymbol{u})_z = \frac{\partial u_y}{\partial x}\,. \tag{10.18}$$

In this case, the curl is thus simply the *gradient* or slope of the velocity $\boldsymbol{u}$ in a direction perpendicular to $\boldsymbol{u}$. (The vector [curl $\boldsymbol{u}$] in Fig. 10.2 at the left of the plate points *towards* the observer, and at the right of the plate, *away from* the observer.)

In general, the volume elements of the liquid move along *curved* paths. Figure 10.25 shows a planar, circular flow in the xy plane. Then we have

$$(\mathrm{curl}\,\boldsymbol{u})_z = \frac{u}{r} + \frac{\partial u}{\partial r}\,. \tag{10.19}$$

Derivation
We compute the circulation along the path shown by heavy lines. It again consists of four terms; they are

$$\mathrm{d}\Gamma = -ur\mathrm{d}\alpha + 0\,\mathrm{d}r + \left(u + \frac{\partial u}{\partial r}\mathrm{d}r\right)(r+\mathrm{d}r)\mathrm{d}\alpha - 0\,\mathrm{d}r = \left(u + r\frac{\partial u}{\partial r}\right)\mathrm{d}r\mathrm{d}\alpha\,.$$

Furthermore, $\mathrm{d}A = r\mathrm{d}r\mathrm{d}\alpha$. Then we find for the quotient $\frac{\mathrm{d}\Gamma}{\mathrm{d}A} = (\mathrm{curl}\,\boldsymbol{u})_z = \frac{u}{r} + \frac{\partial u}{\partial r}$.

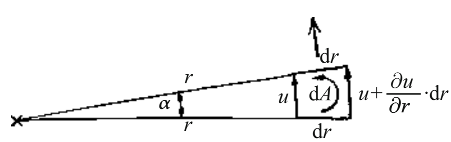


Figure 10.25 The derivation of Eq. (10.19). The arrow above dA indicates its motion

We will now apply Eq. (10.19) to two limiting cases. In the first case, the liquid is presumed to stick to a rotating solid disk and to have the *same* angular velocity ω as the disk in all its regions. Then we have

$$u = \omega r \quad \text{and} \quad \frac{\partial u}{\partial r} = \omega \,. \tag{10.20}$$

With this, from Eq. (10.19) we find for the whole liquid a *constant* value of the curl, namely

$$(\operatorname{curl} \boldsymbol{u})_z = 2\,\omega \,. \tag{10.21}$$

This thus represents the limiting case of a "pseudo-solid rotation".

In liquids, a different limiting case is more important; it is characterized by the condition:

$$u\,r = \text{const} \,. \tag{10.22}$$

Then $\dfrac{\partial u}{\partial r} = -\dfrac{\text{const}}{r^2} = -\dfrac{u}{r}$, and Eq. (10.19) yields

$$(\operatorname{curl} \boldsymbol{u})_z = 0 \,. \tag{10.23}$$

Therefore, when the condition (10.22) is fulfilled, a liquid moves along a *curved* path in a *rotation-free* manner; the arrow on a small float would continually maintain its direction. This curious sort of flow field is called an *irrotational vortex field*[C10.18, C10.19]. It is a further example of a potential field (see Sect. 10.6).

Such a flow field can however exist only when the liquid is circling around a *core*. A well-known example is the *hollow vortex* around a bathtub drain. The *core* in this case is a liquid *surface* which is rotating about its axis like a tube. It surrounds a column of air which is not moving with the revolving liquid and which narrows like a funnel in the downwards direction. – The core of an irrotational vector field can thus be the boundary layer at the surface of a rotating cylinder[4].

Imagine that the diameter of the core is continually decreasing. Then the flow velocity in its immediate neighborhood must be increasing, and will approach ∞ as a limit. This of course cannot occur. Instead, the central parts of the liquid begin to *rotate*. They thus form a *liquid* core, a *vortex tube* or, when the cross-section is very small, a *vortex fiber*. Examples of this kind are given in Sect. 10.8.

The strength of the vortex, known as its *vorticity*, is defined as the circulation Γ along an arbitrary path which *encloses* the core just

C10.18. The magnetic field *outside* a long straight wire carrying an electric current is also an irrotational vortex field (see Vol. 2, Chap. 6, Eq. (6.14)). *Within* the wire, the curl of the magnetic field equals the current density in the wire (Ampère's law).

C10.19. Such irrotational (or rotation-free) vortex fields, which mathematically describe potential flows in the ideal case of frictionless fluids (ideal liquids), can in fact occur in real liquids with finite viscosities, as long as the latter are not too great and can be neglected. In any case, however, friction is necessary to *produce* the vortices. The following paragraph will give some examples; and the observation of vortex fields will be the subject of the rest of this chapter.

[4] To produce and maintain a vortex in water, it suffices to set the cylinder in rotation around its symmetry axis. The thickness of the boundary layer increases without limit as a function of time (Eq. (10.3)). The velocity distribution approaches more and more that of an irrotational vortex field with increasing distance from the surface of the cylinder (i.e. curl $\boldsymbol{u} = 0$). – In air, with its small dynamic viscosity, vortices can be generated by the process described in the caption of Fig. 10.36 (see Fig. 10.37).

once. – Example: An irrotational vortex field surrounds a rotating cylinder and the *boundary layer* which is sticking to it. The cross-sectional area of the cylinder is taken to be A, and its angular velocity is ω. Then the vortex has a vorticity of $\Gamma = \oint \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{s} = 2\pi\, ru = 2\pi\, r\omega r = 2\,\omega A$. One finds the same value along every closed path which circles the core exactly *once*. When the path does *not* circle the core, we find $\Gamma = 0$, the vector field is irrotational, and curl $\boldsymbol{u} = 0$.

An irrotational vortex field, and a core around which the liquid performs its rotation, together form a "vortex". Every vortex thus consists of two components[C10.20].

C10.20. The liquid cores of vortices, as described in Sect. 10.8, are thus regions in which curl $\mathbf{u} \neq 0$ can occur even without friction.

10.8 Vortices and Separation Surfaces in Nearly Frictionless Fluids

Thus far, we have restricted our considerations of motions in fluids to two limiting cases. In the first, we considered motions within a boundary layer. In that case, the internal friction of the fluid played a decisive role (Sects. 10.2 to 10.4). In the second limiting case, we tried to implement motions in fluids which are not influenced by friction or boundary layers (potential flows, Sects. 10.5 and 10.6). We accomplished this by using a flow apparatus of sufficient size, i.e. large compared to the boundary layer. Of greatest importance, however, was that we restricted our observations to short times at the beginning of the motion.

At longer observation times, we find in all fluids, even in those with very low viscosities (tiny values of the internal friction, e.g. gases!), that new phenomena appear: *Vortices and separation layers* are formed[C10.21]. These two phenomena are shown experimentally in the following. – We again use the broad flow apparatus which we have seen in Fig. 10.8, and again observe flow through a bottleneck, as in Fig. 10.9. Initially, the flow field is symmetric in front of and behind the bottleneck, both for the avoidance flow and for the overall flow. These symmetrical flow fields can be observed only immediately after beginning the motion; after a short time, their symmetry is lost. Behind the bottleneck, two large vortices which rotate outwards are formed (Fig. 10.26). These *initial vortices* are rapidly cast off in the direction of the overall flow, and a *stream* remains (Fig. 10.27). It is itself delimited from the surrounding liquid, which is at rest, by a *separation layer*. In this separation layer, several small vortices can be clearly seen. Such a separation layer can be idealized in the limit of vanishing thickness as a *separation surface*. All of the liquid 'volume elements' which it contains must rotate. This is sketched schematically in Fig. 10.28.

C10.21. Examples were already given at the end of the previous section. Note that vortices are formed in a *boundary layer* which borders on a bounding surface of the fluid. They can then propagate in a *separation layer* between regions of the fluid with different velocities.

Vortices can *end* only at the *walls* of a container, at its *bottom*, or at the *surface* of the liquid. *In the interior of a fluid, there are only*

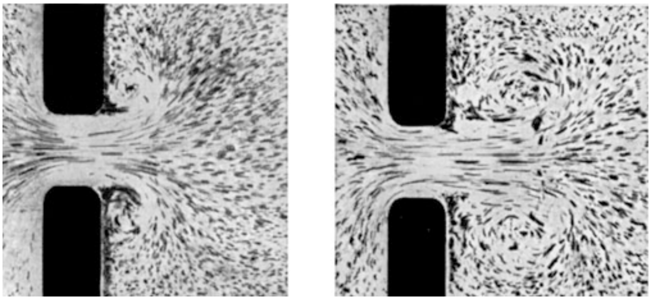


Figure 10.26 Initial vortices at the beginning of stream formation

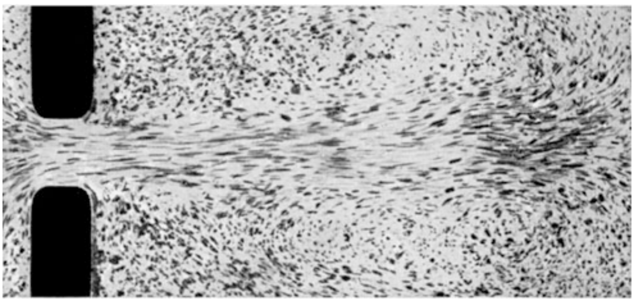
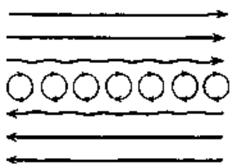


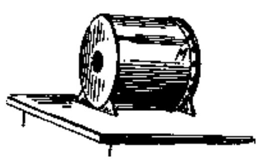
Figure 10.27 The stream of liquid, delimited by separation layers

Figure 10.28 The definition of the separation surface between two parallel flowing liquids with different velocities. In the text, the velocity in one direction is zero.



vortices with closed cores; in the simplest case, they are circular (vortex rings). They can be demonstrated with the apparatus shown in Fig. 10.29. The bottom of a drum-shaped can consists of a stretched membrane M. The air within the drum is made visible using some sort of smoke. A tap on the membrane (drumhead) forces a stream of smoky air to emerge from the round opening at the top of the drum *for a short time*. Its outer edges immediately curve around, forming a vortex ring (a "smoke ring", as occasionally demonstrated by

Figure 10.29 Apparatus for demonstrating vortex rings – closed, initial vortices (cast-off vortices) – in air **(Video 10.4)**



Video 10.4:
"Smoke rings"
http://tiny.cc/ocgvjy

smokers with their mouths). It is futile to try to divide a vortex ring along its diameter into two semicircular pieces with free ends.

C10.22. It is notable that these vortex rings, at a sufficient distance from the boundary layer in which they were formed, can themselves be described in terms of a potential flow. This fact will prove to be particularly important for our discussion of the transverse force (the *lift* in aircraft wings, sails etc.; see Sects. 10.10 and 10.11).

Such a vortex ring is in general a rather stable object[C10.22]. It can fly through the air for several meters past the end of the flow out of the drum, and as a result of its energy[5], it can blow over a playing card, snuff out a candle flame, etc. – The vortex rings that are made visible by smoke (in the shape of inflated inner-tubes) contain the *core* of the vortex, which is enclosed within the ring. In this core, *rotation* is present (that is, curl $\boldsymbol{u} \neq 0$). The second component of the vortex, the *irrotational vortex field* which surrounds the core, extends far outwards. This is shown by the mutual influence of two vortex rings. They can for example be produced in rapid succession: The second ring then overtakes the first, reducing its diameter in the process, while the first ring expands and slows, letting the second ring pass through. This performance is repeated one or two times, each time exchanging the roles of the two rings. – Two vortex rings which are on a collision course moving towards each other along the same line slow each other down and expand as they approach.

So much for the facts. – The separation surface and vortices are formed here by the same cause as everywhere, namely by *sticking* (adhesion) of the fluid to the solid object around which it is flowing, and the resulting formation of a boundary layer. – An ideal fluid (potential flow) should flow around the walls of a bottleneck at a high velocity; every real fluid, however, will be impeded by the boundary layer which forms. This impedance will act differently at the entrance and the exit of the bottleneck. On the way *towards* the bottleneck, all parts of the stream will be accelerated. Within the bottleneck, the flow velocity attains its maximum value. The impeded boundary layer will be pulled along in the direction of flow by neighboring layers of the fluid whose flow is unimpeded. Thus, *in front of* the bottleneck, the original flow field, i.e. the potential flow field, will be maintained. *Behind* the bottleneck, in contrast, all parts of the flow field will be decelerated. There, the impeded boundary layer is no longer pulled along by the neighboring layers. It falls behind; there is no alternative except to curve around and inject itself between the wall and the flow. Thus, "the flow detaches from the walls", and the separation surface and vortices are formed.

10.9 Flow Resistance and Streamline Profiles

The processes that we have just described, i.e. the formation of vortices and separation surfaces, can lead us to an understanding of the forces which occur when nearly frictionless fluids flow around solid bodies. These are the *flow resistance* or *drag* (described in this section) and the *dynamic transverse force* (or *lift*) (in Sect. 10.10). We

[5] Bell-shaped jellyfish use the recoil from water vortices which they produce as a means of propulsion.

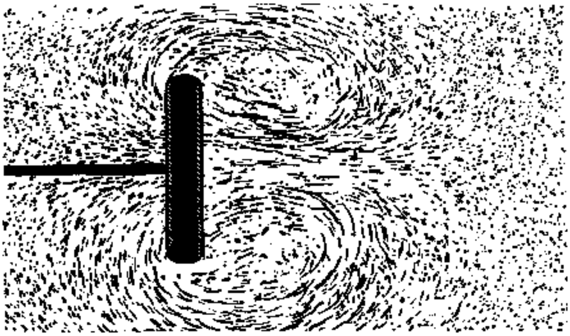


Figure 10.30 Distortion of the avoidance flow behind a plate which is oriented perpendicular to its direction of motion; observer and liquid are at rest, and the plate is moving to the left **(Video 10.3)**

Video 10.3:
"Fluid flow around obstacles"
http://tiny.cc/fcgvjy

will investigate both of them using our flow apparatus as shown in Fig. 10.8. The objects which are moved relative to the liquid and are thus surrounded by a flow are again assumed to nearly touch the glass walls of the apparatus on both sides. In both cases, therefore, we are investigating a planar flow. The results can then be extended by analogy to three-dimensional flows.

We start with the streamline patterns which we have already observed in the avoidance flow around a plate, Fig. 10.16 (potential flow). In these examples, the flow around the front and back ('leading' and 'trailing' edges) of the plate is completely symmetric. This means, according to Eq. (10.13), that the pressures and forces on the front and back sides are also symmetric. The sum of all the forces acting on the object is thus zero. Then the object could be moved within the liquid without any flow resistance at all! This, however, is in contradiction to our everyday experience; think for example of rowing a boat, or stirring your soup. – In fact, this symmetry, which is also observed in the experiments shown in Figs. 10.19 through 10.21, is destroyed very soon after the start of the motion. To demonstrate this, we take a plate which is oriented perpendicular to the flow. At the very beginning of its motion, there is a symmetric avoidance flow, which we can observe by repeated back-and-forth motions of the plate (Fig. 10.19). When the motion is continued in one direction for a longer time, the symmetric flow becomes distorted: Two large initial vortices which rotate inwards are formed (Fig. 10.30). They rapidly move off in the flow direction, and in the stationary state, we can clearly observe a *separation surface* behind the plate. It separates a region (which extends beyond the right edge of the image) from the rest of the flow (Fig. 10.31). Within this region, the liquid is rotating in a lively manner. A number of vortices are present (this is shown impressively in **Video 10.3**).

Video 10.3:
"Fluid flow around obstacles"
http://tiny.cc/fcgvjy

We now have an overview of how resistance or drag is produced by flow around objects in real fluids. It comes about through the *rotational motions* which form within the fluid at the trailing edge of

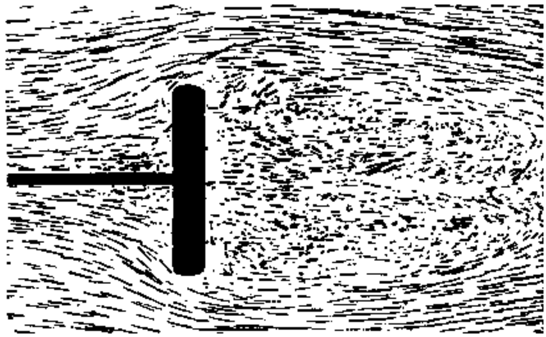


Figure 10.31 The production of flow resistance (drag) through vortices within a bell-shaped separation surface. In this experiment, observer and object are at rest, and the liquid is flowing to the right. The resistance for REYNOLDS numbers *Re* between $4 \cdot 10^3$ and 10^5 is somewhat greater than the product of the stagnation pressure and the area of the plate. Experimentally, one finds its magnitude to be $F = 1.1 \cdot \frac{1}{2}\varrho u^2 A$[C10.23].

C10.23. The comparison of Figs. 10.31 and 10.16c shows very impressively the influence of the viscosity: In both figures, the object is at rest. In Fig. 10.31, the liquid is flowing to the right; it could be flowing to the right in Fig. 10.16, also (owing to the symmetry of the streamlines, one can choose the direction at will). The lack of vortices in Fig. 10.16 is a special feature of the flow chosen for this model experiment (HELE-SHAW flow).

the object. New regions of the liquid are continually being brought into rotation. Creating the rotational motion in these vortices, providing their kinetic energy, requires that work be performed. The force which performs this work is the counter-force to the flow resistance. *The resistance experienced by an object which is surrounded by an avoidance flow is caused by rotational or vortex motions on its trailing edge.* That is the surprising experimental finding.

The resistance or drag felt by bodies within a flow is often used for technical applications. As an example, we could mention the parachute (it reduces the sinking velocity of a person from about 55 m/s to around 5.5 m/s, cf. Fig. 5.20); other examples include the oars of rowboats and the paddle wheels of steamboats. A further example is provided by *resistance rotors* as wind turbines; they have an S-shaped profile (Savonius rotors), or hemispherical shells mounted on spokes, as in an ‘anemometer’. (The air resistance of the concave side of the shell is four times larger than that of the convex side.)

In other cases, the resistance is a disturbing factor. Then it must be reduced as far as possible by careful design of the shape of the object. Only the small *frictional resistance* then remains; it arises in the boundary layer between the object and the fluid. Nature has provided us with many prototypes here. Their common characteristic is their *streamlined profile*, as in Fig. 10.32. A streamlined object of similar shape can be placed in a water flow at high velocity, and no vortices will be formed. A sphere of about the same diameter at the same flow velocity will produce strong vortex formation soon after the flow is initiated. Streamlined profiles play an important role in nature and in technology.

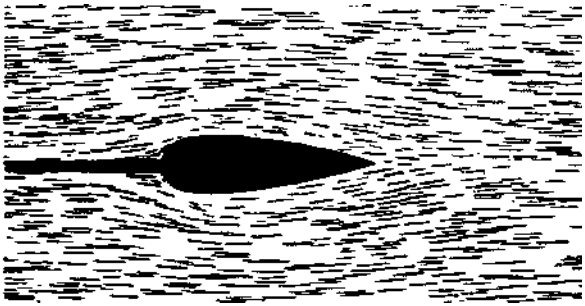


Figure 10.32 A streamlined profile (in the flow apparatus from Fig. 10.8); observers and object are at rest, the liquid is flowing to the right (see also **Video 10.3**, in which the object is moving)

10.10 The Dynamic Transverse Force or Lift

In general, the direction of the undisturbed flow is not identical with a symmetry axis of a solid body within the flow. An example is shown in Fig. 10.16b and in the right-hand part of Fig. 10.19. Then observations teach us a new result: In addition to the resistance or drag $\boldsymbol{F}_\mathrm{r}$ *in* the direction of the unperturbed flow, there is a second force which is *perpendicular* to the flow, as shown in Fig. 10.33. It is called the *transverse force* or *lift*, $\boldsymbol{F}_\mathrm{a}$. The resultant vector of these two forces is the total force $\boldsymbol{F}$ which acts on the body within the flow[6].

The transverse force cannot be completely isolated and investigated by itself. However, we can make the resistance $\boldsymbol{F}_\mathrm{r}$ very small compared to the simultaneously-present transverse force $\boldsymbol{F}_\mathrm{a}$. To this end, one has to use objects with the profile of an airfoil or a wing, e.g. as

Video 10.3:
"Fluid flow around obstacles"
http://tiny.cc/fcgvjy
One can see here very clearly the influence of the direction of flow on the formation of vortices in the boundary layer: When the streamline profile is moved upwards (i.e. to the left in Fig. 10.32, liquid at rest), no vortices are formed (apart from those which are produced at the thick end of the object at the end of the motion, when the displaced liquid flows back around the object). When however the streamline profile is moved downwards, vortices immediately form at its thicker end.

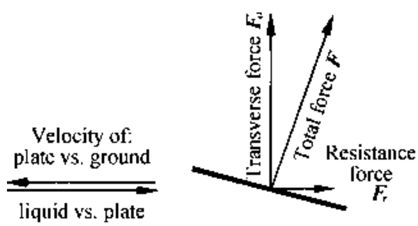


Figure 10.33 The transverse force (lift) and resistance force (drag) on a tilted plate which is moving to the left. (The resultant total force on thin plates is nearly perpendicular to the plane of the plate.)

[6] For rough calculations, one can keep in mind two useful approximations: Transverse force $F_\mathrm{a} = \frac{1}{3}\varrho u^2 A$, and resistance force $F_\mathrm{r} \approx$ 1 to 10 % of F_a (A = area of the plate or airfoil).

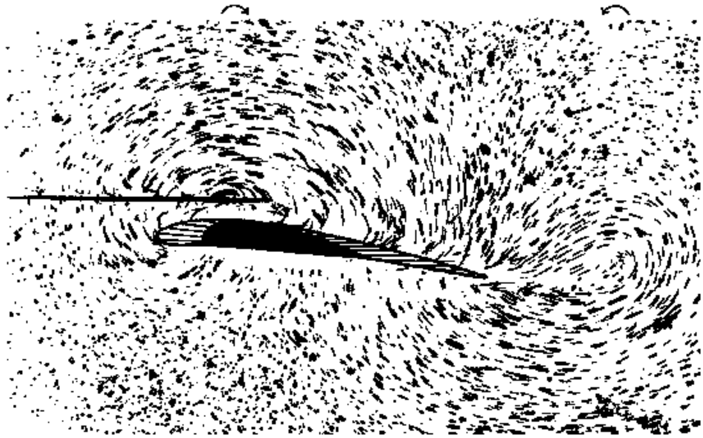


Figure 10.34 The formation of an initial vortex from the avoidance flow around an airfoil or hydrofoil which is moving in the shaded region (compare with Fig. 10.19). The fluid and the observer (camera) are at rest, the airfoil is moving to the left[C10.24] **(Video 10.3)**

C10.24. The rod which projects into the picture from the left is only a holder for the airfoil. The connecting rod perpendicular to it is not visible owing to the motion.

Video 10.3:
"Fluid flow around obstacles"
http://tiny.cc/fcgvjy

in Fig. 10.34. Furthermore, the object must either be "infinitely" long or bounded by planes (as in our flow apparatus, Fig. 10.8).

Such an airfoil or hydrofoil shows the *origin of the transverse force* (lift) rather clearly. We begin with a comparison of Figs. 10.16b and 10.34: When the motion is initiated, a long-lived initial vortex is formed as in Fig. 10.34 only by the initial avoidance flow from behind and below. The sense of rotation of this vortex is marked at the upper right with an arrow ↶. This vortex moves off with the flow. From the initial avoidance flow *in front* and below, an *irrotational vortex field* is formed instead of an initial vortex; it circles the airfoil (as its core) in a clockwise sense (↷). It has the same direction above the airfoil as the flowing fluid; below, in contrast, the two flows are oppositely directed. As a result, the fluid flows faster above and more slowly below. This produces a region of reduced static pressure above the airfoil, so that a *dynamic transverse force* or *lift* $\boldsymbol{F}_\mathrm{a}$ acts perpendicular to the direction of the unperturbed flow.

The observation of the avoidance flow is hampered by the washed-out contours of the airfoil. Thus, one usually draws the whole flow, i.e. the avoidance flow and the parallel flow, as in Fig. 10.35. At first, a potential flow forms as seen in Fig. 10.35, upper part. This forms the irrotational vortex field sketched in the center part of Fig. 10.35. The potential flows are superposed and lead to the flow field shown in the bottom part of Fig. 10.35.

The airfoil or hydrofoil can be replaced by a *rotating cylinder*. The formation of the vortex field follows the same pattern as around an air- or hydrofoil. At first, an initial vortex forms at the trailing edge and is carried off with the flow. Later, the flow pattern shown in Fig. 10.36 is established. With the directions of motion as shown, the

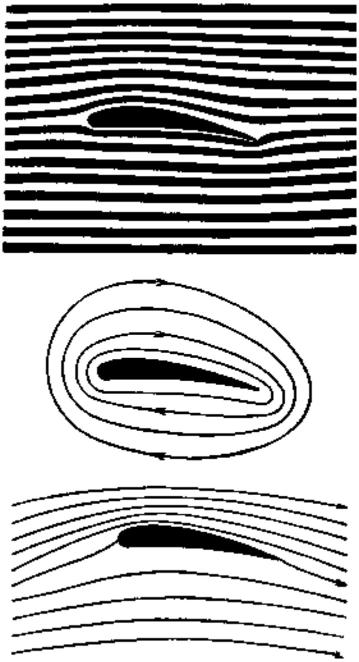


Figure 10.35 The origin of lift on an airfoil or hydrofoil. *Top:* Potential flow without a vortex field (model experiment); the streamlines observed in the model experiment are shown in the ideal case of a frictionless fluid **(Video 10.1)**. *Center:* The irrotational vector field which is produced by the initial vortices in a real fluid ($\eta > 0$) (schematic). *Bottom:* The superposition of the two flow fields. The irrotational vortex field cannot be observed by itself, but it can be clearly discerned in Fig. 10.34 (see also **Video 10.3**, "Fluid flow around obstacles")

Video 10.1:
"Model experiments for streamlines"
http://tiny.cc/8wqujy

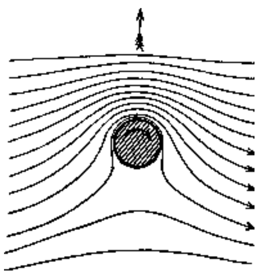


Figure 10.36 Streamline pattern around a rotating cylinder. On its upper side, the surface of the cylinder has the same direction of motion as the avoidance flow; this prevents the formation of an initial vortex there. On its lower side, the two motions are oppositely directed, which favors the formation of an initial vortex. The difference in the pressures above and below the cylinder gives rise to the lift force.

cylinder experiences a transverse force or lift in the direction of the feathered arrow.

To demonstrate this effect, we use a light cardboard cylinder about the size of a rolled-up dinner napkin (Fig. 10.37).
Its ends are closed off with disks which are somewhat larger in diameter than the cylinder. A flat cloth ribbon is rolled up on the cylinder and is attached to a handle like a whip. When the handle is jerked to the side horizontally, the cylinder is accelerated across the table top, but at the same time it begins to rotate because the ribbon unrolls. Instead of flying off along a falling parabola, it rises up sharply and follows a looping path.

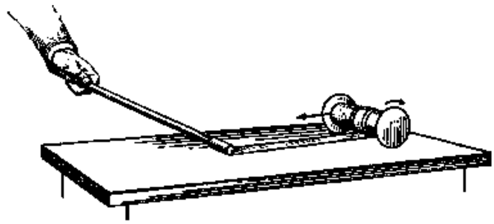


Figure 10.37 The dynamic transverse force or lift on a rotating cylinder (MAGNUS effect) **(Video 10.5)** (GUSTAV MAGNUS, 1802–1870, Physics Institute, University of Berlin).

Video 10.5:
"MAGNUS effect"
http://tiny.cc/dcgvjy

In both cases, for air- or hydrofoils as well as for a rotating cylinder, the relation discovered independently by M. W. KUTTA and N. J. JOUKOWSKI yields the magnitude of the dynamic transverse force F_a:

$$F_a = \varrho u \Gamma l \tag{10.24}$$

(ϱ = density; u = velocity of the parallel flow, in Fig. 10.37 for example the velocity of the cylinder on the table top; l = length of the airfoil or the rotating cylinder; Γ = vortex strength (circulation) of the vortex which surrounds the airfoil or the rotating cylinder of length l)[C10.25].

C10.25. The MAGNUS effect, using a vertical rotating cylinder, was suggested as a propulsion method for ships by A. FLETTNER (1885–1961).

In nature and in technology, a *planar* flow around an airfoil, hydrofoil or a rotating cylinder is never observed. The ends of the airfoil or cylinder are not bounded on both sides by large planar surfaces, like the glass walls of our flow apparatus in Fig. 10.8. One can also not construct infinitely long wings or cylinders. – Their finite lengths however give rise to a new effect. The vector field which circles the wing or cylinder (potential vortex) produces not only the transverse force which provides lift, but also a *drag* which opposes the motion. It is called the *induced drag*. We give a brief description of how it comes about (Fig. 10.38): At the two ends of a wing (airfoil), the high-pressure regions from the bottom of the wing come into contact with the low-pressure regions from the top surface. Air flows from below to above, producing vortices at both wingtips. Together with the vortex field around the wing as core and the initial vortex, they form a single closed vortex line. Its length continually increases, as more and more air is set into rotation at the wingtips. The work required to produce this vortex line must be generated by a force, and its counter-force is the "induced drag". – Without this drag, an aircraft could maintain a constant altitude in frictionless air without any propulsion.

Summary of Sects. 10.9 and 10.10: There are two ways besides *friction* in a boundary layer to generate forces between a nearly frictionless fluid and a solid body. First of all, by producing vortices at the trailing edge of the body: Like friction within the boundary layer, this creates a *resistance* (drag) opposed to the direction of motion.

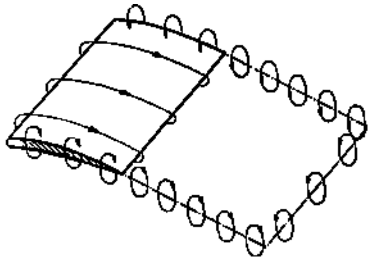


Figure 10.38 The origin of induced drag: The small curved arrows indicate only the sense of rotation and not the limits of the flow field (in reality, they should be spirals). The air flows upwards to the sides in a wide region *along-side* the area passed over by the wing. This is the reason why many migratory birds, for example ducks and geese, prefer to fly side by side in a staggered configuration, forming a "wedge" or "string". Then, except for the lead bird, every bird is flying in an upwards-flowing region of air, so that it requires less power to produce the lift needed to maintain its altitude. Only the bird at the point of the formation lacks this assistance, so that it must be relieved from time to time.

Second, by producing an irrotational vortex field with the body as its core: This gives rise to the dynamic *transverse force* (lift, transverse to the direction of the unperturbed flow); and in addition, even when the body has a good air- or hydrofoil profile, as a result of its finite length also an *induced drag* (opposing the direction of motion within the fluid, like every resistance force).

10.11 Applications of the Transverse Force

The transverse force which acts on rotating bodies (the lift force in Fig. 10.37) is used for example in sports. Example: A "cut" tennis ball, i.e. a ball hit with a glancing blow, flies further than a ball which is not rotating, because the transverse force compensates the weight of the ball. – For bodies which have the profile of a wing (airfoils, hydrofoils, sails), there are numerous applications of the transverse force: In *airfoils*, the component of the transverse force which is *perpendicular* to the direction of flight is used as *lift*; in *sails*, the component which is *in* the direction of motion drives the ship. – Examples:

1. If the lift acting on an *aircraft* is equal and opposite to its weight, it will fly horizontally, obeying the scheme indicated in Fig. 10.33. Its forward velocity is usually maintained by an engine. The essentials were already summarized in Sect. 5.11 – If the engine is shut off, the induced drag (a result of the finite length of the wings) and the frictional losses in the boundary layer will consume its kinetic energy.

These losses must be compensated from the stored potential energy; i.e. the aircraft must slowly glide to the ground. The angle of glide of its flight path is determined by the ratio of the lift to the drag (F_a/F_r; cf. Fig. 10.33). Therefore, that quotient is called the *glide ratio*.

Gliding is the mode of flight of sailplanes, and it is practiced by glider pilots and some birds. – In gliding flight, there are two methods of gaining altitude:

a) By gliding in a rising column of air (a 'thermal');

Examples
A seagull suspended in the air stream that is moving upwards at an angle behind the stern of a ship; a bird of prey circling in the rising, warm air above and around a tall chimney.

b) by making use of the vertical gradient of a horizontal wind velocity. In the boundary layer of the air above the surface of the earth (land or sea), the horizontal wind velocity increases with increasing altitude.

Example
The albatross glides along a sloping flight path downwards in the direction of the wind, collecting kinetic energy. Near to the surface of the sea, it loops around and faces the wind, rising steeply higher, since it can use its stored energy to penetrate the layers of increasing wind speed, thereby increasing the lift acting on its wings. Once it has reached a certain altitude, it turns again to follow the wind in a downwards glide, and so forth.

2. A *child's kite* is held against the wind by a string from the ground; this permits the air to flow around it, without carrying it away in a horizontal direction. – The profile of a kite is a rather poor approximation to a good airfoil, but it suffices.

3. A ship can move readily in the direction of its long axis, but it is hard to move it sideways. Like the kite-string, this transverse resistance to motion can hold the ship against the wind, so that its sails are surrounded by an air flow. This makes it possible to *sail* even in a direction which is not the same as that in which the wind is blowing. In Fig. 10.39, we see a ship sailing "close-hauled" (also called "working to windward"); i.e. the wind is coming obliquely from the

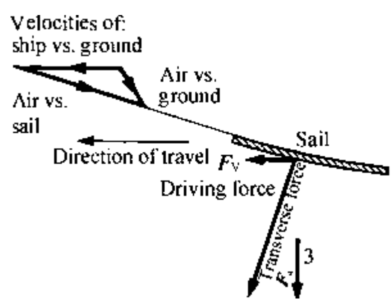


Figure 10.39 Sailing "close-hauled": *Arrow 3* shows the direction of the sidewards drift.

forward direction. The velocity of the ship must be subtracted vectorially from the velocity of the wind relative to the ground; as a result, the wind is flowing "flatly", i.e. at a small angle of attack, around the sails. The component $\boldsymbol{F}_v$ of the transverse force *in* the direction of travel drives the ship forward. The component perpendicular to the direction of travel leads to a small sidewards drift.

4. The rotor blade of a *wind turbine as motor*. Figure 10.40 shows the cross-section of a rotor blade. The wind is blowing perpendicular to the plane of the rotor, but hits the blades at a small angle of attack. The component F_v of the transverse force which is parallel to the plane of the rotor drives its rotation. The orbital velocity of the blades at some distance from the rotor hub may exceed the wind velocity by a large factor.

In order to make this intuitively clear, put a flat wedge on a smooth horizontal surface and press on its flank vertically with the point of a pencil. This will slide the wedge horizontally by a distance which is greater than the vertical motion of the pencil. As a toy, one can build a windmill with blades of a symmetric profile, for example a semi-cylindrical cross-section (as shown in Fig. 10.41). A starting push gives the blade (at the position of the cross-section shown in the figure) a velocity $\boldsymbol{u}$ relative to the ground. The direction of $\boldsymbol{u}$ determines whether the initial vortex is cast off the left side or the right side of the blade.
This in turn determines the sense of rotation of the circulation and thus of the whole rotor. The blade is subject to an air flow under a small angle of attack. All else occurs as shown in Fig. 10.40.

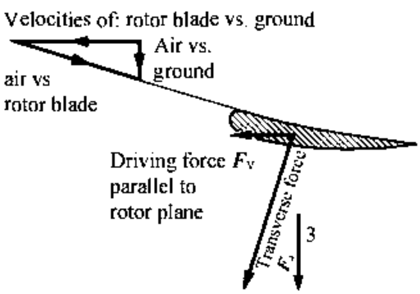


Figure 10.40 The principles of operation of a wind-turbine blade. The component of the transverse force in the direction of *arrow 3* puts a load on the vertical axis of the support structure.

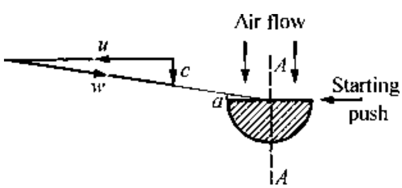


Figure 10.41 A toy windmill with two blades which have a symmetric cross-section. They consist of half-cylindrical rods which can rotate around the *A*-*A* axis. When they are given a push in the direction indicated, the initial vortex is formed at the right and the windmill rotates clockwise as seen looking in the direction of the air flow (wind).

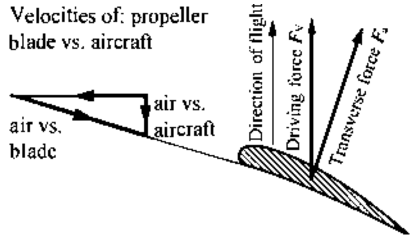


Figure 10.42 The principles of operation of an aircraft propeller blade. The propellers or screws of ships operate in analogous fashion

5. *The aircraft propeller*[C10.26]. The sketch shown in Fig. 10.42 again represents the cross-section of a propeller blade. The component of the transverse force perpendicular to the plane of the propeller provides the driving force or thrust which maintains the flight velocity of the aircraft. This is seen differently by an observer sitting in the aircraft; he or she would say: The propeller is a fan which is blowing an *air stream* backwards. Accelerating the airstream requires a force, and the counter-force acts on the aircraft and drives it forward.

C10.26. In the first edition (1930), POHL explained at this point: “The propeller of an aircraft or motor ship does *not* bore into the fluid like a corkscrew. Its blades are simply rotating airfoils (hydrofoils).” One should keep this fact in mind in discussing the other examples in this section, as well!

Of course, the air stream could be produced by an enclosed fan instead of a free-standing propeller. Modern turbojet engines dispense with reciprocating pistons, and this represents a considerable technical advance[C10.27].

C10.27. The first experiments with jet engines for aircraft were carried out in 1935 by HANS V. OHAIN (1911–1998) in Göttingen. He was at that time a doctoral student and assistant in the research group of R.W. POHL (see e.g. *The Jet Age*, W.J. Boyne and D.S. Lopez, eds. Smithsonian Institution, Washington, DC (1979), p. 25). See also the Wikipedia article on R.W. POHL, in particular Reference 17.

Exercises

10.1 Two balls of radii R_1 and R_2 and densities ϱ_1 and ϱ_2 are sinking in petroleum ether with the constant velocities u_1 and u_2. Determine the viscosity η and the density ϱ of the petroleum ether. (Sect. 10.3)

10.2 A cylindrical container of radius R is filled with water up to a depth h. The water is flowing out through a circular opening in the bottom of the cylinder, with a volume current of $\mathrm{d}V/\mathrm{d}t = 1\,\mathrm{cm}^3/\mathrm{s}$. How large is the radius r of the opening? (Sect. 10.5)

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_10) contains supplementary material, which is available to authorized users.

Vibrations and Waves

II

Vibrations

11

11.1 Preliminary Remarks

Knowledge of vibrations and waves has its origin in their close connection to our sense of hearing and to music. The human body possesses an extremely sensitive detector of mechanical vibrations and waves over an astonishingly broad frequency range (from roughly $\nu = 20\,\mathrm{Hz}$ to $20\,000\,\mathrm{Hz}$)[1]. Today, general topics of the physics of vibrations and waves are given priority in introductory courses, while acoustics as such is less emphasized. The material in this and the next chapter has been chosen and organized with this in mind[C11.1].

C11.1. Acoustics in the specific sense will be treated in the following chapter beginning with Sect. 12.24. The terms 'oscillations' and 'vibrations' both refer to periodic motions. The latter is less general and is usually applied to mechanical and acoustic motions, the subject of this chapter. 'Oscillations' may also refer to periodic phenomena in non-material systems, e.g. electric and magnetic fields (see Vol. 2). Here, we use both terms, 'oscillations' when more generality is implied, as e.g. the *harmonic oscillator*.

11.2 Producing Undamped Vibrations

Thus far, we have treated only the sinusoidal vibrations of simple pendulums under a linear force law. The scheme of such pendulum vibrations was presented in Figs. 4.13 and 4.14. The vibrations of these simple pendulums were produced by an impulse directed to the massive or inertial element of the pendulum. They are *damped*, i.e. their amplitudes decrease with time, and the energy which they obtained from the initial "pulsed excitation" is gradually lost, mainly through the unavoidable friction between the pendulum and its environment.

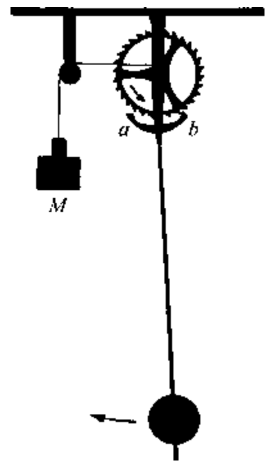
However, for many physical, technical and musical purposes, one requires *undamped* vibrations, whose amplitude remains constant over time. The production of such undamped vibrations calls for the continuous replacement of the energy losses mentioned above. The techniques which have been invented to accomplish this can be summarized by the keywords *feedback* or *auto control*: The pendulum itself activates a mechanism which accelerates its motion at the right moment in the sense of its momentary direction of vibration.

The *classical paradigm for all control through feedback is the pendulum clock*, Fig. 11.1. It replaces lost kinetic energy by tapping the stored potential energy of a load of mass M that has been lifted to a certain height. The energy transfer is accomplished by a "crown wheel", with teeth which are cut asymmetrically, together with an "escapement" attached to the pendulum. This escapement permits the pendulum to control the stepwise rotation of the crown wheel and

[1] The unit of frequency is 1 hertz $= 1\,\mathrm{s}^{-1}$ (abbreviated Hz)

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_11

Figure 11.1 Auto-controlling of a gravity pendulum through feedback via an escapement and crown wheel



the transfer of energy from the descending load M. In the position shown in the figure, a tooth of the crown wheel is pressing against the inner flank of the right hook of the escapement, b, and thus accelerates the pendulum in the direction of the arrow, to the left. As soon as the pendulum passes the midpoint of its swing, the tooth will slide off the hook b and immediately thereafter, the other hook a will again engage another tooth of the crown wheel. This tooth is now pressing against the upper flank of the hook a, so that the pendulum is accelerated to the right, and so forth.

Positive feedback of this kind can be practically implemented in a great variety of ways, often by purely mechanical means. For example, one can produce a periodic connection of a vibrating system to a source of energy by means of "sticking" or "hooking" of two bodies which are momentarily at rest relative to each other (frictional vibrations).

Demonstration

In Fig. 11.2, we see a gravity pendulum about the size of an average clock pendulum, in a side view. It is connected to a shaft of about 4 mm diameter

Figure 11.2 Auto-controlling of a gravity pendulum via friction with a rotating axle (frictional vibrations). The length of the pendulum is around 30 cm, and its mass is 200 g. The leather in the clamps on the axle has to be rubbed with resin, like a violin bow.

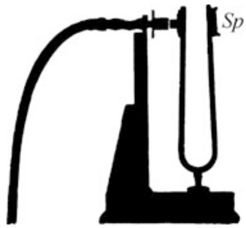


Figure 11.3 Pneumatic feedback control of a tuning fork **(Video 11.1)**

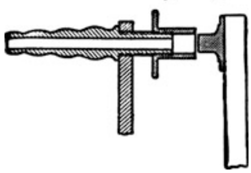


Figure 11.4 Auto-controlling of a tuning fork by pneumatic feedback **(Video 11.1)**

Video 11.1:
"Vibrations of a tuning fork"
http://tiny.cc/idgvjy
The video shows the tuning fork as in the figure. Note the adjustment of the feedback by turning the knurled screw on the cylinder b in Fig. 11.4.

by two stuffed leather clamps. When the shaft is rotated, the pendulum is pulled forwards with it. The clamps initially stick to the shaft ("sticking friction"). At a certain deflection of the pendulum, the torque resulting from the pendulum's weight becomes too great and the friction of the clamps is overcome; they slip back along the shaft ("sliding friction"). The pendulum swings back, reverses, and then during its next swing forward, at a certain moment its velocity relative to the shaft is zero – the pendulum clamps and the shaft are moving synchronously. The clamps now again stick to the shaft, and the pendulum is again accelerated up to the point where it breaks free. Its next swing begins, with the same amplitude as the first, and so forth.

Feedback control via airflows (pneumatic control) is also widely used. Figure 11.3 shows an example, used to excite the vibrations of a tuning fork.

The essentials are shown in Fig. 11.4 as a cross-sectional view. A piston a is fitted smoothly into the cylinder b, but not touching its walls. The cylinder is connected to a source of compressed air by a tube. The air pressure forces the piston out of its resting position in the cylinder, pushing the tine of the tuning fork to the right. When the piston is outside the cylinder, a ring-shaped gap opens between the piston and the wall of the cylinder, allowing the air to escape along closely compressed streamlines. From BERNOULLI's equation (Eq. (10.13)), the resulting static pressure of the air is reduced and the piston is pulled back in. The natural frequency of the tuning fork controls this process through the mechanical coupling of the piston's motion to the vibrations of the fork.

The use of *electrical* devices to control *mechanical* vibrations has long since gained importance. The oldest example is the electric doorbell, familiar to every school child today (Fig. 11.5). A pendulum with an iron rod ('clapper') vibrates in front of one pole of an electromagnet M. The rod also carries the contact spring of an interrupter switch S.

In descriptions of the operation of a doorbell, the decisive point is often misunderstood. When the interrupter contact is closed, the iron rod of the

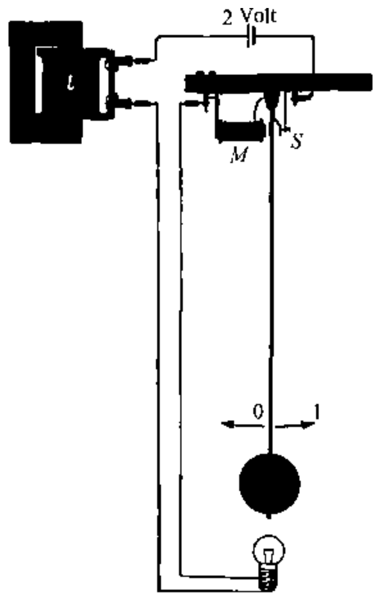


Figure 11.5 Auto-controlling of a gravity pendulum by an electromagnet

pendulum (the 'clapper') is accelerated towards the magnet. This acceleration occurs not only during the half-swing $1 \to 0$, but also during the half-swing $0 \to 1$; but along the path $0 \to 1$, the acceleration is in the wrong direction. It is opposite to the motion of the swinging pendulum, retarding the pendulum's motion and reducing its energy. As a result, an additional condition must necessarily be fulfilled: The gain in energy along the path $1 \to 0$ must be greater than the energy loss along the path $0 \to 1$. Only the difference between these two quantities of energy is useful for driving the vibrations of the pendulum. Practically, this means that the current in the electromagnet along the path $0 \to 1$ must be less on average than that along the path $1 \to 0$. The current in the electromagnet must therefore increase with time from $0 \to 1$.

Technically, this increase in the current is produced by the self-induction of the electric circuit. – For the demonstration, we use a gravity pendulum (as shown in Fig. 11.5) which swings *slowly* ($\nu = 2$ Hz), and put an auxiliary coil L (a "choke") with a large inductance into the circuit (see Vol. 2, Chap. 10). A small lamp under the rest position of the pendulum allows us to follow the slow increase of the current clearly (Fig. 11.6): The lamp begins to glow during each swing only when the pendulum is reversing at its maximum deflection 1.

Figure 11.7 shows a trace of the motion of the clapper of a doorbell without the bell gong. The clapper rod was vibrating in front of a slit and its motion was recorded photographically using the light passing through (cf. Fig. 1.9)[C11.2]. The time variation of the vibrations in this case reveals clear deviations from a simple sine curve; the curves seem to be pointed. *Every feedback mechanism perturbs the sinusoidal form of the vibrations* (distortion!). Elimination of the damping is bought at the price of abandoning strictly sinusoidal vibrations. But the distortion can be reduced by careful design, to

C11.2. See also Figs. 11.22 and 11.23 in Sect. 11.5.

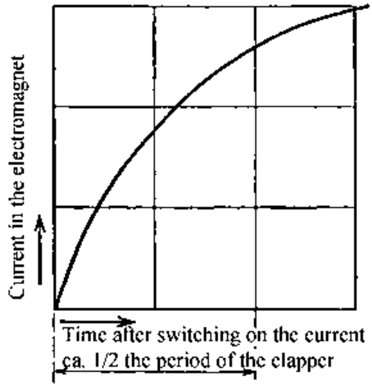


Figure 11.6 The current as a function of time in the feedback circuit of Fig. 11.5 (schematic of a doorbell)

Figure 11.7 The time graph of the non-sinusoidal vibrations of a doorbell clapper

a much greater extent than in this intentionally exaggerated demonstration experiment.

Many modern electrical feedback mechanisms for mechanical vibrations make use of triodes (electron tubes or transistors). Examples can be found in Vol. 2, Chap. 11.

11.3 The Synthesis of Non-Sinusoidal Periodic Processes and Structures from Sine Curves

The deflection, velocity etc. of most periodic processes are not strictly sinusoidal. Likewise, most periodic structures do not have strictly sinusoidal profiles. Nevertheless, sine curves or sine functions play an important role in physics: It is possible to *synthesize, i.e. both to generate and to represent* non-sinusoidal, periodic curves by making use of simple sine functions. We demonstrate this first by using vibrations which we generate *kinematically*, referring to the well-known relation between a circular orbit and a sine curve (Sects. 1.7 and 4.3). We begin with the superposition of two sinusoidal vibrations of differing frequencies. We move a rod along a circular path in front of a slit and present the time sequence of the slit images as separate pictures spread out in space (using a polygon mirror in the optical path). We can see the rod and the slit in the upper part of Fig. 11.8. The rod is attached at its two ends to the circumferences of two disks *I* and *II* so that it can rotate with them;

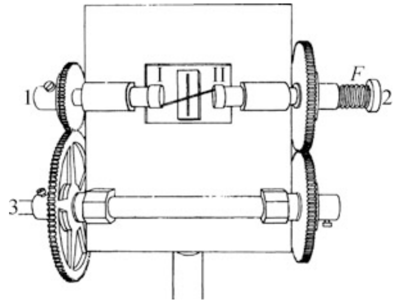


Figure 11.8 Demonstration apparatus for the superposition of two sinusoidal vibrations. The two shafts 1 and 2 are rotated via their gears by an electric motor which drives shaft 3[C11.3].

C11.3. Today, we can of course demonstrate the superposition of oscillations in a much simpler way by using electronic oscillators or frequency synthesizers. Mechanical devices may however have an advantage in that their operation can be understood intuitively in a "palpable" manner!

they are driven by an electric motor. The gears permit a fixed, integer *frequency ratio* to be established between the rotation rates of the disks, and furthermore, any desired *phase difference* between the two vibrations can be chosen.

> To adjust the phase difference, the upper gear on the right, which is held by the spring F, can be pulled out and turned by the desired angle before being engaged again to the lower gear.

The slit can be slid horizontally within the window. This allows the *amplitude ratio* of the two vibrations to be adjusted to the desired value.

Oscillations S whose frequencies are in the ratios of whole numbers will be denoted in the following by integer indices, i.e. S_1, S_2, $S_3 \ldots$ We use the same indices for their amplitudes A. – Now for some examples:

In Fig. 11.9, we see the traces of two sinusoidal oscillations S_1 and S_5, i.e. oscillations whose frequencies are in the ratio 1 : 5. For the ratio of the amplitudes $A_1 : A_5$, we have chosen a value of about 3 : 1. The bottom trace shows their superposition: The trace of the combined

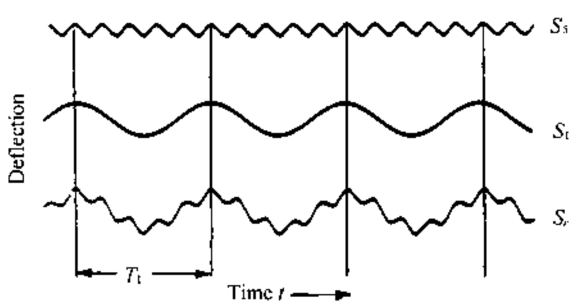


Figure 11.9 The superposition of two sinusoidal oscillations S_1 and S_5, whose frequencies are in the ratio 1 : 5 and whose amplitudes have a ratio $A_1 : A_5 \approx 3 : 1$. This picture as well as the following figures 11.10 and 11.11 are photographic images, made using the apparatus shown in Fig. 11.8.

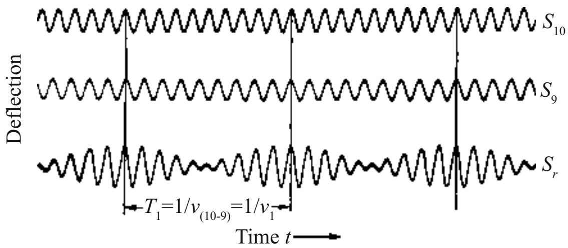


Figure 11.10 The superposition of two sinusoidal oscillations S_{10} and S_9, i.e. two oscillations whose frequencies are in the ratio 10 : 9 and whose amplitudes are approximately equal. The resulting curve S_r exhibits *beats*.

oscillation S_r resembles a sine curve which was drawn by someone with a very shaky hand.

In Fig. 11.10, upper part, we show two sinusoidal oscillations S_9 and S_{10} with the frequency ratio $9:10$, of nearly equal amplitudes, $A_9 \approx A_{10}$. Their superposition is shown as the bottom trace, S_r. It again resembles a sine curve, but with a periodically-variable amplitude. Oscillations of this form are called *beats*. The beat frequency, often denoted as ν_B [C11.4], is equal to the difference $\Delta\nu$ of the two individual frequencies. In this example, the oscillation amplitude goes to zero at each beat minimum. At the time of the minimum, the two equal amplitudes of the two sinusoidal oscillations are opposite; their phase difference is 180°. At the time of the beat maxima, in contrast, the two amplitudes add with a phase difference of zero to give a value twice that of the individual amplitudes. With two component oscillations of unequal amplitudes, the minima of the beat curve would not go to zero.

C11.4. Note that this beat frequency is twice as large as the difference frequency derived from the trigonometric sum formula:
$$\sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cdot\sin\frac{\alpha-\beta}{2}.$$

In Fig. 11.11, upper part, we see the traces of two oscillations S_1 and S_2 with an amplitude ratio of $A_1 : A_2 \approx 3:2$. Their superposition gives a curve S_r which is symmetric around the time axis.

In Fig. 11.11, lower part, we see the same oscillations S_1 and S_2 as in the upper image, but now the oscillation S_2 begins at a time $t = 0$ with a phase angle of 90° (its maximum), while oscillation S_1 has a deflection of zero (phase angle 0°). The resulting superposed oscillation S_r looks quite different, in spite of having the same amplitudes and frequencies. It is not symmetric around the time axis. In this example, we see clearly the influence of the *phase difference* on the form of the resulting curve.

Summarizing the superposition of just two sine curves: In Figs. 11.9 through 11.11, we could "synthesize" some non-sinusoidal curves using two sine functions, i.e. we could generate the curves as well as describing them.

In the non-sinusoidal oscillations in Figs. 11.9 through 11.11, a certain oscillation pattern repeats itself in every detail after a period T_1. The reciprocal $1/T_1$ is called the *fundamental frequency* ν_1 of the

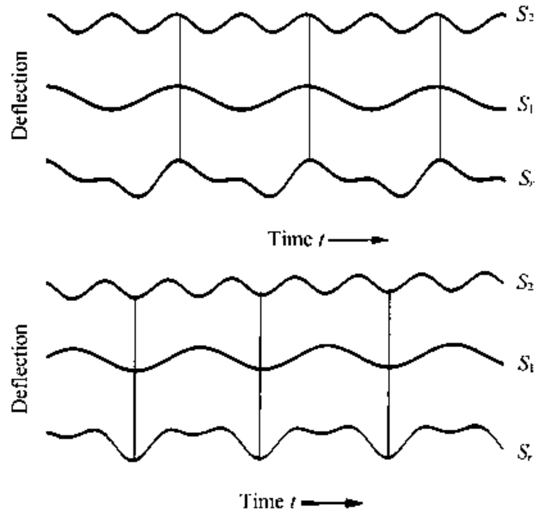


Figure 11.11 The superposition of two oscillations S_1 and S_2 whose frequencies are in the ratio 1 : 2. The comparison of the two resulting curves S_r shows the influence of the phases on the form of the superposition curve.

non-sinusoidal oscillation process. The frequencies of the two component oscillations are integer multiples of this fundamental frequency.

In a corresponding manner, by adding more and more component sine functions, we can "synthesize" ever more complex curves. The amplitudes and phases of the components can be appropriately chosen. Their frequencies must always be integer multiples of the fundamental frequency of the periodic curve which we want to synthesize. – Two examples:

In Fig. 11.12, top, we synthesized a beat curve from two component functions S_9 and S_{10}. In the bottom curve, we have added a third component $S_1 = S_{(10-9)}$ to this beat curve[C11.5]. The frequency of the third component is thus taken to be equal to the difference (beat frequency) of the two other frequencies. Furthermore, its positive maxima are chosen to fall together with those of the beat curve. – The addition of such a difference oscillation transforms the original beat curve, which was symmetrical around the abscissa, into an asymmetrical curve. The magnitude of the asymmetry depends in a clear-cut way on the amplitude of the difference oscillation employed.

C11.5. An apparatus for the mechanical superposition of *three* vibrations was described by G. BEUERMANN (*Praxis der Naturwissenschaften* **27**, 227 (1978)).

For our second example, we wish to synthesize the oscillation shown as the upper trace in Fig. 11.13, a periodic sequence of "rectangles" or "boxes", by superposing sine curves. This can be done with a moderate degree of approximation by graphically adding just three sine curves, S_1, S_3, and S_5. We obtain the curve labelled S_r. The approximation can be improved as much as desired by adding additional sine curves, S_7, S_9 . . .

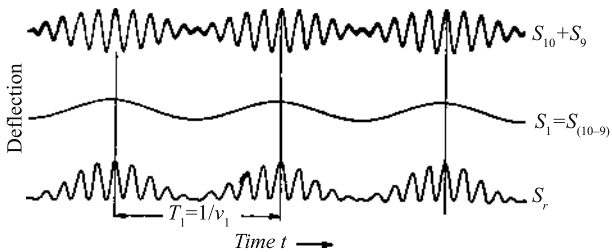


Figure 11.12 The asymmetric oscillations S_r resulting from the superposition of two sine functions S_{10} and S_9 with their difference-frequency curve $S_1 = S_{(10-9)}$. The frequency of this latter curve is thus equal to the difference of the other two frequencies.

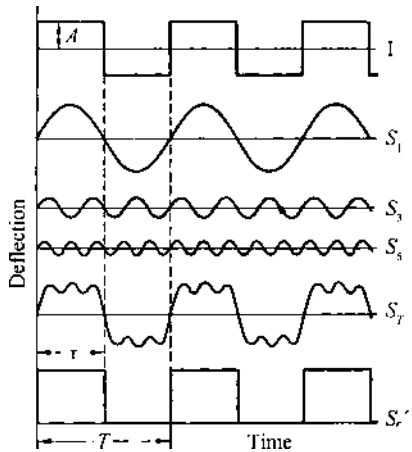


Figure 11.13 Synthesis of a rectangular oscillation I by superposing three sine curves, S_1, S_3, and S_5; S_r is the resulting curve. The curve S_r' shows curve I shifted upwards by A, so that it 'rests on' the axis of the abscissa (cf. Fig. 11.14).

The oscillation curve I sketched in Fig. 11.13 can thus be described approximately using the three sine curves shown below it. In analytic form, this description is as follows:

$$x = \frac{4A}{\pi}\left(\sin\omega t + \frac{1}{3}\sin 3\omega t + \frac{1}{5}\sin 5\omega t + \cdots\right) \tag{11.1}$$

($x =$ deflection, $A =$ amplitude of the rectangular function, $\omega = 2\pi/T$) .

This example of "FOURIER *analysis*", i.e. the representation of a complex periodic curve in terms of a sum of sine curves whose angular frequencies are integer multiples of the fundamental frequency ω, can only hint at the great importance of this method. FOURIER analysis can of course also be applied to spatially-varying structures; we need only change the labels on the coordinate axes. The time coordinate t is replaced by a length coordinate, and $\omega = 2\pi/T$ by $k = 2\pi/D$, where D, the length period, corresponds to the time period T and reflects the spatial periodicity of the structure. Furthermore, FOURIER analysis can also be applied to propagating or travelling waves. For example, a box-shaped wave travelling to the right, in the positive z direction, with a velocity u (cf. Fig. 11.13), can be described by Eq. (11.1); we need only replace ωt by $(\omega t - kz)$, where $\omega/k = (2\pi/T)/(2\pi/D) = D/T = u$.

11.4 The Spectral Representation of Complex Oscillatory Processes

When the oscillatory processes are very complicated, one often dispenses with a representation of the time dependence of the oscillation and uses instead its *spectral representation*.

A spectrum is plotted against the *frequencies* of its component oscillations as abscissa. The ordinate values, called spectral lines, indicate the *amplitudes* of the individual component oscillations by their lengths. Thus, in Fig. 11.14, we see the spectrum belonging to the oscillation curve S'_{r} from Fig. 11.13. It is a *line spectrum*, the simplest representation of an oscillatory process. A spectrum contains less information than the complete representation of the time dependence of the oscillation: *A spectrum contains no information about the phases.* Knowledge of the phases is to be sure indispensable for tracing the oscillation curve (e.g. Fig. 11.11); but we do not require this knowledge for many physically important tasks associated with non-sinusoidal oscillations.

In Fig. 11.13, we are dealing with a special case. There, we chose $\tau/T = 1/2$. If the ratio of τ/T becomes smaller, the number of component sine functions required increases. As an example, in Fig. 11.15 we have taken $\tau/T = 1/12$.

In Fig. 11.15, the frequency spectrum of this rectangular oscillation is illustrated with its first 20 spectral lines. If one adds up the first 10 of these component oscillations, the result is the periodic curve III; the sharp upper

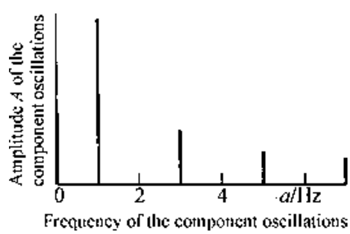


Figure 11.14 The frequency spectrum of the rectangular oscillation curve S'_{r} shown at the bottom of Fig. 11.13; a is an arbitrary number, e.g. 10^3, which has been divided out of the component frequencies as a common factor. – In the representation of the oscillation curve I at the top of Fig. 11.13, which is symmetric around the axis of the abscissa, the spectral line at the frequency 'zero' would not occur.

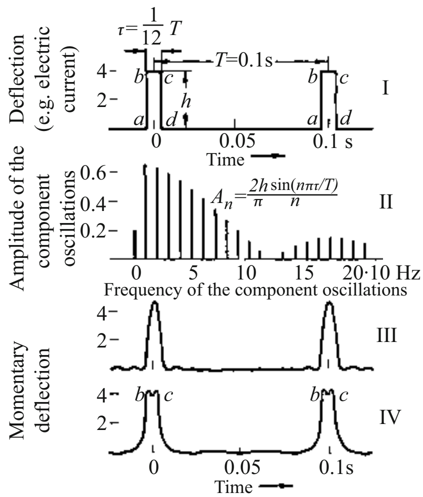


Figure 11.15 Curve I: A rectangular oscillation or pulse train in which the time width τ of the pulses (e.g. electric current pulses) is much smaller than their repetition period T (i.e. the pulse duty factor is small). Curve II shows the first 20 spectral lines of the associated line spectrum. A spectral line at the frequency 'zero' means that there is a "constant" offset (e.g. a constant value of direct current). Curve III is the resultant superposition of the first 10 component oscillations (spectral lines), and curve IV shows the superposition of the first 20 components.

corners b and c are still missing. In curve IV, the next 10 spectral lines are added in. This has at least begun to reproduce the upper corners b and c. To reproduce the lower corners a and d, a still greater number of additional spectral lines must be included. The same holds quite generally for those parts of curves that contain straight segments that are steeply inclined to the time axis, e.g. the steeply falling segments on one side of a sawtooth profile.

We show as examples two more spectra of important oscillation phenomena.

1. *Frequency spectra of damped sinusoidal oscillations with periodic pulsed excitation.* We take a numerical example for brevity: Some sort of oscillator is to be excited to *undamped* sinusoidal oscillations of frequency $\nu = 400\,\text{Hz}$. After its initial excitation, a sine curve of constant amplitude and unlimited length results. The spectrum consists of only a single spectral line at the frequency 400 Hz.

Now, this oscillator is somehow damped. As a result, after a *one-time pulsed excitation*, it exhibits an oscillation of finite length whose amplitude decays with time (Fig. 11.16, part g). Above this curve, we see the oscillation of the same system with *periodically-repeated pulsed excitation*. In part e, a new excitation pulse is applied after each 8 oscillations; in part c, after each 5 oscillations; and in part a, after each 2 oscillations.

Part II

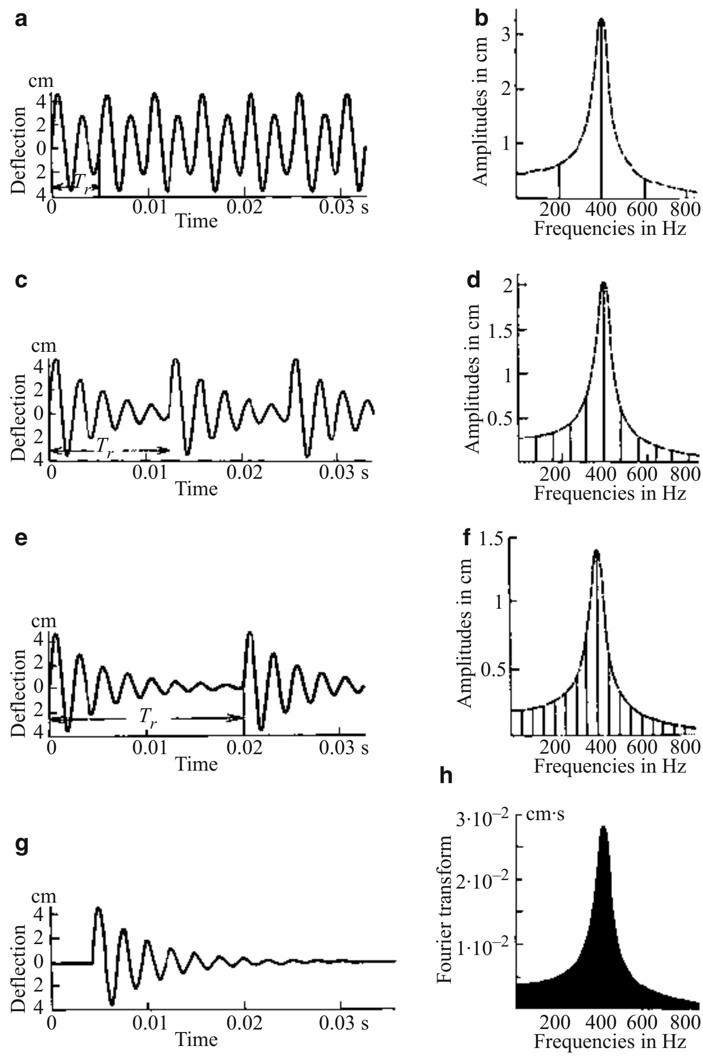


Figure 11.16 The same damped sinusoidal oscillation as a function of time for pulsed excitation at different pulse frequencies; **a** pulsed after each 2 oscillations, or pulse frequency 200 Hz, **c** pulsed after each 5 oscillations, or pulse frequency 80 Hz, **e** pulsed after each 8 oscillations, or pulse frequency 50 Hz, **g** only a single excitation pulse; **b**, **d**, **f** corresponding frequency spectra (FOURIER transforms) of the oscillation curves to the left (note the scales of the ordinates); **h** continuous spectrum (FOURIER *integral*) of the damped oscillation at the left, with only one excitation pulse

Alongside each of these three time functions, we find the associated spectrum. None of them is similar to the simple spectrum of the undamped oscillation, that is with only a single spectral line at the frequency 400 Hz. Along with this original frequency of 400 Hz, we see a whole series of additional spectral lines. In each of the three spectra, the lowest frequency is that corresponding to the pulse rate, or for short, the *pulse frequency*. In the three spectra, from above, it is 200, 80 and 50 Hz. The pulse frequency is the fundamental frequency ν_1 of the three non-sinusoidal oscillations. All the other spectral frequencies must be integral multiples of the pulse frequency used for that particular spectrum. Therefore, the spectral lines occur only occasionally at the same frequencies in the spectra for different pulse frequencies. But, they are – and this is important – always in the same frequency *range*. It is called the *formant range* (the dashed curves in Fig. 11.16b, d, and f).

With decreasing pulse frequency, the required number of frequency components or spectral lines in the spectral representation continues to increase. We require an increasingly large number of sinusoidal oscillations in order to represent the broad gap regions between the damped oscillations by mutual compensation of their amplitudes. In the limit of a single excitation pulse (zero pulse frequency), we arrive at

2. A *continuous spectrum for a damped oscillation with a one-time pulsed excitation.* In Fig. 11.16, part g, we see the time dependence of the damped, decaying oscillation after a single excitation pulse; and in part h, its frequency spectrum. The spectral lines now occur with an infinite density; they fill the region under the envelope curve (dashed in the upper images) continuously. The region under the curve is therefore shown as a black area in part h. Instead of a spectrum with individual, separate lines, we now have a *continuous spectrum*.

These important relationships have been treated only descriptively here. Their analytic derivation is dealt with in detail in every reasonably complete mathematics course under the topic of FOURIER analysis[C11.6].

C11.6. For example in H.J. Pain, *The Physics of Vibrations and Waves*, John Wiley, New York, 4th ed. (1993), Chap. 9.

11.5 Elastic Transverse Vibrations of Linear Solid Bodies Under Tensile Stress

Systems capable of vibrations, or *oscillators*, have in our treatment up to now been reduced to a simple *schematic* form, consisting of an object with inertia to store kinetic energy, and an elastic object (spring) to store potential energy. The most clear-cut form of this scheme was a massive ball held between two extended helical springs (Fig. 4.13). We will henceforth refer to this 'ball-and-spring pendulum' as an *elementary oscillator* (A system which oscillates with a pure sine function is called a *harmonic oscillator*). This simple scheme was thus far sufficient to represent the majority of the oscillators which we have considered, although at times only by stretching

Figure 11.17 Transverse vibrations of two coupled elementary oscillators; both massive balls are moving in phase (stopped-motion images)

Figure 11.18 Transverse vibrations of two coupled elementary oscillators; the balls vibrate with a 180° phase shift, i.e. in opposite directions (stopped-motion images)

the point somewhat. It is however by no means sufficient to describe all the cases which occur in nature. Often, a separate localization of the inertial element and the elastic element is not reasonable. After all, every object of whatever form can vibrate; we know this from our daily experience. We have thus arrived at the problem of the *proper elastic vibrations* (normal modes) of arbitrary bodies.

A vibration of the elementary oscillator along the axis of its springs is called a *longitudinal vibration*; a vibration perpendicular to the spring axis is a *transverse vibration*. We will start by investigating transverse vibrations.

In Figs. 11.17 and 11.18, two such elementary oscillators are connected together or *coupled*. This system can vibrate in two ways: In the first case, both massive balls vibrate *synchronously* or *in phase*. In Fig. 11.17, two instantaneous ('stopped-motion' or 'snapshot') images of these vibrations are drawn. In the second case, the two balls vibrate *in opposite directions*, or *phase shifted by 180°*. Here again, we show two stopped-motion images (Fig. 11.18).

The vibration frequencies are different in the two cases. Using a stopwatch, we observe in Fig. 11.18 a higher frequency than in Fig. 11.17. For two coupled elementary oscillators, we thus find two transverse *normal modes* of vibration with the corresponding (proper or eigen-) frequencies ν_1 and ν_2.

In an analogous manner, we see in Fig. 11.19 three coupled elementary oscillators. Now, three different transverse vibration modes are found; all three are shown as stopped-motion images in the figure. They can be demonstrated experimentally without difficulties. With *three* coupled elementary oscillators, we thus obtain *three natural or proper frequencies* (often called 'eigenfrequencies'), corresponding to three normal modes of vibration.

We could continue indefinitely in this way, repeatedly adding one more coupled oscillator to our chain. For a chain of N coupled elementary oscillators, we obtain N normal modes. In the limit of large N, we arrive at *continuous* linear objects. For such an object, we can thus expect a practically unlimited number of transverse normal

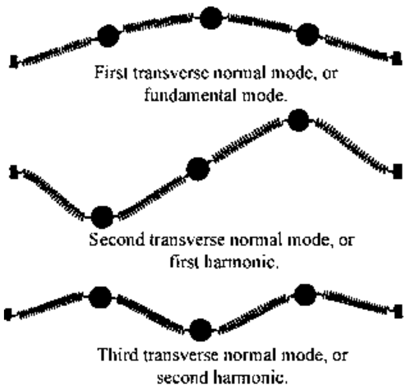


Figure 11.19 The three possible transverse vibrations of three coupled elementary oscillators (stopped-motion images)[C11.7]

C11.7. One must be careful with the nomenclature: The first normal mode is also called the *fundamental vibration*, and the second is called the *first harmonic*, etc.!

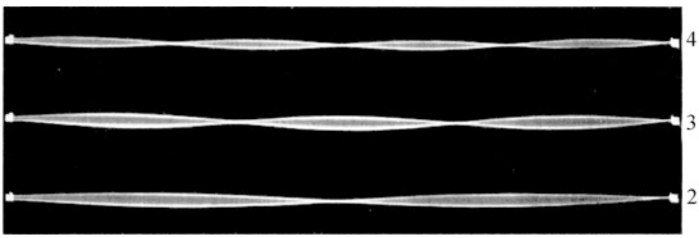


Figure 11.20 Photographic (time exposure) images (side view) of the second to the fourth transverse normal modes of a stretched rubber elastic band; the band appears light in front of a dark background. Where it shows a gray color, the velocity of motion perpendicular to its long axis is maximal. To excite a normal mode of frequency ν, one end of the band is moved periodically by a motor, either *transversally* to its long axis at the frequency ν, or *along* its long axis at the frequency 2ν. In the latter case, the excitation is technically termed *parametric*, because the tension of the band (as a parameter) is periodically varied. (The tension attains its maximum value *twice* during *one* vibration period of the normal modes.) **(Video 11.2**, see also **Video 1)**

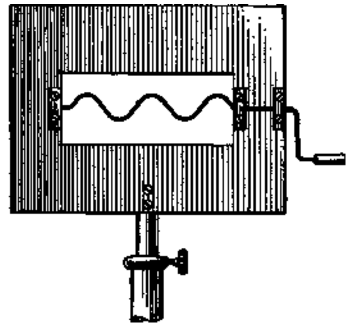
Video 11.2:
"Transverse normal modes of a stretched rubber band"
http://tiny.cc/negvjy
The video shows some transverse normal modes as standing waves, up to the eighth normal mode.

Video 1:
"R.W. Pohl Lecturing"
http://tiny.cc/fpqujy

modes. As a first example, we show the undamped transverse vibrations of a horizontally-stretched rubber elastic band. Figure 11.20 shows a side view as time-exposure photographs of its second to fourth normal mode vibrations. In each of these examples, we see *three* quantities *periodically* distributed along the band, namely the transverse *deflection*, the transverse *velocity*, and the *slope* of the band relative to its rest position. All three quantities show *nodes* and *maxima*. At their nodes, each of the three quantities remains zero at all times. At their maxima, the three quantities exhibit their largest values. The deflection maxima and the velocity maxima are found at the same positions; likewise the nodes of these quantities. The maxima of the slopes, however, are at the positions of the nodes of the deflection and the velocity; for example at the two ends of the band.

For a purely kinematic visualization of transverse normal modes, a wire bent to a sine-curve shape with a crank at one end is quite satisfactory (Fig. 11.21). The wire is positioned in front of the projector lamp and

Figure 11.21 Visualization of transverse normal modes or standing waves



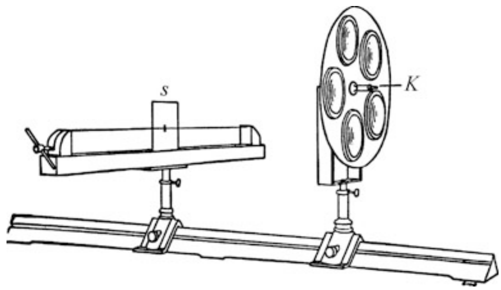


Figure 11.22 Shadow projections of the vibration curves of one point on a vibrating string, using a rotating lens wheel **(Video 11.3)**

Video 11.3:
"Transverse vibrations of a string"
http://tiny.cc/odgvjy
The string is plucked or bowed.

rotated around its long axis. The projected image then shows the individual instantaneous graphs of the vibration (often called the "vibration phases" for short), one after another. When the crank is turned quickly, we can readily observe the transition to the time-exposure photos as seen in Fig. 11.20. This primitive setup is rather useful.

We return once more to Fig. 11.20 and imagine that the band which is vibrating in its fourth normal mode is struck in the plane of the page. Then the whole band begins to vibrate in addition in its first normal mode, and the two normal modes occur *simultaneously*. Such a simultaneous occurrence of several normal modes is often found in stringed musical instruments. In Fig. 11.22, we see a string which is stretched horizontally. Its vibrations can be excited by a violin bow.

With a slit which is mounted perpendicular to the string, we can select a single "point" along the string and record its motion photographically. Figure 11.23 shows some examples: A single point along the string, e.g. in Fig. 11.22 its midpoint, can be seen to move transversally to the string in a way which is by no means a simple sine function. Instead, one sees as a rule rather complex, non-sinusoidal vibration curves. They result from the superposition of a large number of normal modes (proper vibrations). The appearance of a *single* normal mode can be achieved only by very selective bowing and even then only approximately. In general, the strings of musical instruments exhibit a very complex vibration spectrum[C11.8].

C11.8. See e.g. the frequency spectrum of a violin tone in Fig. 12.86.

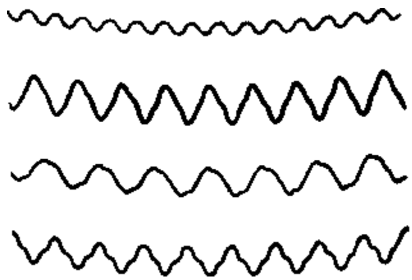


Figure 11.23 The time curves of the deflection of one "point" on a violin string which is undergoing transverse vibrations at multiple frequencies **(Video 11.3)**

Video 11.3:
"Transverse vibrations of a string"
http://tiny.cc/odgvjy

For demonstration experiments, instead of moving the projection lens along a horizontal straight line (as in Fig. 1.9), one can use the *lens wheel* shown in Fig. 11.22. When it is rotated, the individual lenses are brought into the optical path one after another. The wheel is rotated by hand using the knurled knob K. The abscissa of the projected images obtained (the time axis) is slightly curved (compare Fig. 11.23; this is a harmless flaw).

11.6 Elastic Longitudinal and Torsional Vibrations of Stressed Linear Solid Bodies

At the beginning of Sect. 11.5, we defined the *longitudinal vibrations* of an elementary oscillator as a vibration of the massive component of the oscillator *parallel to* the long axis of its helical springs.

In Fig. 11.24, we illustrate the two longitudinal vibration modes of two coupled elementary oscillators. At left, both massive balls are moving in the same direction, synchronized "in phase". At the right, they are vibrating in opposite directions, "phase shifted by 180°". We continue adding additional elementary oscillators to make a chain, and find N normal modes with N oscillators. Thus we again arrive in the limit of large N at a linear object with a practically unlimited number of longitudinal normal modes of vibration. As an exam-

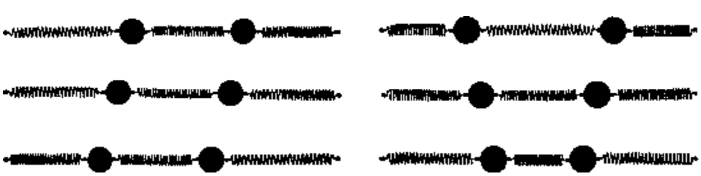


Figure 11.24 Three stopped-motion images each of the longitudinal vibrations of two coupled ball-and-spring oscillators; top and bottom at the times of their maximum deflections, center image at passage through their rest positions. On the left, both massive balls are vibrating with the same phase, on the right with a phase shift of 180°.

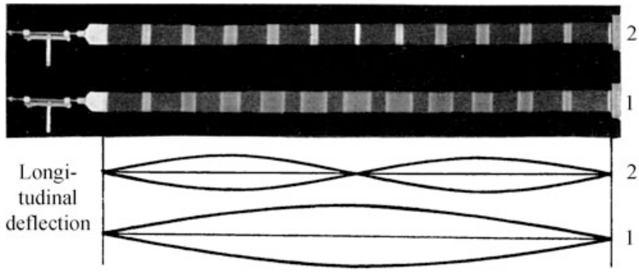


Figure 11.25 *Top:* The first and second longitudinal normal modes of a rubber band with white cross stripes on a black background, 1 m long; time-exposure photos. Where the white stripes are visible only as a gray blur, the vibration velocity along the band is high. *Bottom:* graphic representations of the distribution of maximum deflection (and of the longitudinal velocities) along the band. To excite the normal vibration modes, the left end of the band was pulled periodically *in the direction* of its long axis at the frequency ν of the normal mode by a motor. (In **Video 1**, the author demonstrates this experiment himself.)

Video 1:
"R.W. POHL Lecturing"
http://tiny.cc/fpqujy

ple, we show the undamped longitudinal vibrations of a horizontally-stretched black rubber band with white cross stripes.

Figure 11.25 shows photographic (*time exposure*) images of the first and the second longitudinal normal modes. In both examples, we immediately see two quantities periodically distributed along the band, namely the longitudinal *deflections* and the longitudinal *velocities*. The maxima of the deflection and the velocity fall at the same points, as do their nodes. The third quantity is the elastic deformation[C11.9] (stretching and compression), which is also periodically distributed along the band. The periodic distribution of the elastic deformation causes a periodic variation ΔN_l in the *density of stripes* N_l along the band. We define the density of stripes N_l as the quotient

C11.9. This deformation corresponds to the slope in the case of transverse vibrations (Fig. 11.20).

$$N_l = \frac{\text{Number of stripes in a segment } \Delta l}{\text{Length } \Delta l} = \frac{1}{\text{Spacing of the stripes}}. \tag{11.2}$$

The two *stopped-motion images* in Fig. 11.26 show the variations ΔN_l of the density of stripes N_l along the length l of the band for the first longitudinal normal mode, nearly at the phases of the maximum deflections. The maxima of these variations are at the ends of the band. They are thus at the points where the deflection and the velocity have their nodes (Fig. 11.25).

In addition to the transverse and longitudinal vibrations of linear solid bodies, there are also *torsional oscillations*. They can be conveniently demonstrated using a braided rubber band a few centimeters wide which is stretched horizontally. Figure 11.27 shows time-exposure images for three normal modes.

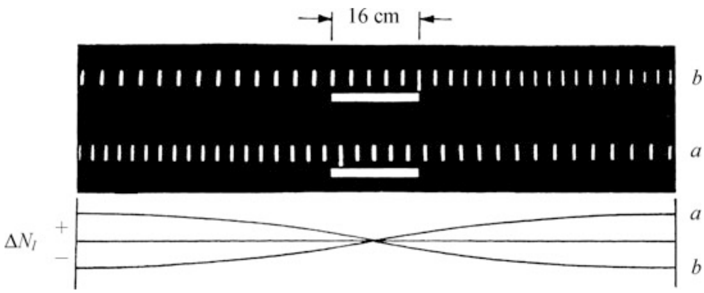


Figure 11.26 Photographic high-speed images (ca. 10^{-5} s) of the first longitudinal normal mode of a striped rubber band nearly at the phases of its maximum longitudinal deflections (to be observed stroboscopically!). The maximum deflection in the center is ± 8 cm. The somewhat longer vertical streak allows us to discern the vibration phases. Due to overstretching the band, the images are only qualitatively accurate[C11.10].

Figure 11.27 The third to the fifth torsional oscillation modes of a 1 m long and 3 cm wide stretched rubber band. The holder at one end is turned back and forth by an eccentric on an axis parallel to the long axis of the band; angles of a few degrees are sufficient.

C11.10. The connection between the variations ΔN_l shown in Fig. 11.26 and the quantity N_l defined in Eq. (11.2) is given by $\Delta N_l = N_l - N_{l,\mathrm{mean}}$, where $1/N_{l,\mathrm{mean}}$ is the mean spacing of the stripes, averaged over the total length of the band. Since the maxima of ΔN_l are at the ends of the band, we find for the nth normal mode that $(n+1)$ maxima exist for ΔN_l. Compare Fig. 11.31 in the following section.

11.7 Elastic Vibrations in Columns of Liquids and Gases

As before, we treat liquids and gases together (as 'fluids'). Our experiments will be carried out for the most part using air.

In the interior of liquids and gases (their difference: formation of surfaces), no transverse or torsional vibrations are possible, but rather only *longitudinal vibrations*. This is a direct result of the free motions of all the liquid or gas particles past each other[2].

As in the case of solid bodies, we will first treat *linear* arrangements of liquids and gases. Linearly-bounded liquid or gas columns can be constrained in *tubes*.

Gas columns can easily be excited to normal-mode vibrations. For example, as a demonstration, one can close off one end of a cardboard tube around 1 m long and several centimeters in diameter with

[2] 'Particles' in the sense of small volume elements, not individual molecules.

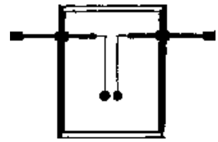


Figure 11.28 Aerodynamic detection of the back-and-forth air flows along the axis of an organ pipe. (The balls could also be hung one behind the other instead of side by side. Then the air flow in both directions produces a mutual repulsion of the balls.)

a rubber membrane. By plucking or tapping the membrane, we can excite this "air column" to loud, readily audible, but rapidly decaying normal-mode vibrations. Or else we could put a solid cover over one end and pull a removable cover rapidly off the other end. These longitudinal vibrations take place in a manner basically similar to those of an elastic rubber band as described in Sect. 11.6. Imagine that the air column is divided up into thin layers, where each layer takes the place of one stripe on the rubber band.

Between the nodes of the vibrations, these layers flow back and forth. Their motion can be made visible with small *dust particles* suspended in the air column. These can be observed microscopically and thus used to measure the maximal deflections to both sides ("amplitudes"). – For demonstration experiments with a large audience, the back-and-forth air flows at the maxima of the flow velocity can be shown using aerodynamic forces which are independent of the direction of motion.

Example

In the interior of a tube with a square cross-section, we hang two small pith-balls on thin threads (Fig. 11.28). Two windows allow us to observe the balls as a projected shadow image. The line connecting the two balls is first placed perpendicular to the long axis of the tube. Then for a flow along the tube axis, the streamline pattern shown in Fig. 10.17 is found. The streamlines are pressed together between the two balls; the balls must *approach each other* when the column vibrates or is blown as an organ pipe. This is indeed the case.

The nodes of this longitudinal flow can be shown by scattering a fine powder on the inside bottom of the tube. The powder particles come to rest at the nodes and form KUNDT's dust figures.

We show the normal-mode vibrations at the frequency of $\nu \approx 3 \cdot 10^4$ Hz (Fig. 11.29). Excitation is provided by a flue pipe which is blown directly in front of the open end of the tube (Fig. 11.35).

The periodic distribution of the flow velocity maxima can be demonstrated with RUBENS' flame tube (Fig. 11.30).

The flame tube is several meters long and is filled with natural gas. On its upper side, there is a series of burner openings running the length of the tube. One end of the tube is closed, the other carries a membrane which can be excited by some means to give undamped vibrations. Their frequency has to be equal to one of the natural frequencies of the gas column in the tube. – When the vibrating gas is flowing back and forth in the tube,

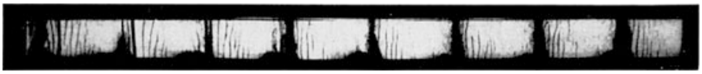


Figure 11.29 KUNDT's dust figures. During the vibrations, the dust forms a fine fog which hangs perpendicular to the axis of the tube like curtains. They move slowly along the axis. They demonstrate that the flow within the longitudinally-vibrating gas column is associated with complex side effects ("second-order effects"). These come about through the formation of a boundary layer between the walls of the tube and the flowing parts of the gas column.

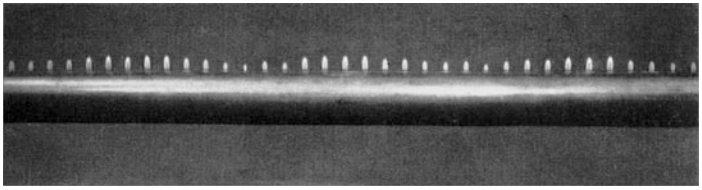


Figure 11.30 RUBENS' flame tube shows the distribution of the flow velocities in a longitudinally-vibrating gas column. The heights of the flames are constant over time and their maxima lie above the maxima of the flow velocity[C11.11]. (**Video 1** shows W. SPERBER igniting the RUBENS' flame tube. http://tiny.cc/fpqujy)

a *boundary layer* forms along its walls; cf. Sect. 10.2. In this layer, the static pressure above the velocity maxima increases by an amount which is *constant over time*, and above the nodes, it correspondingly decreases[3]. This causes the periodic distribution of the flame heights, which is constant over time. An application of these phenomena is shown in Fig. 12.46.

In the case of longitudinal vibrations of an elastic rubber band, the *stripe density* was distributed periodically along the length of the band. The maxima of the stripe density were located where the axial motion exhibits *nodes* (compare Fig. 11.26). Precisely the same is true for a longitudinally-vibrating gas or liquid column; but now, instead of the stripe density, we observe the *number density* N_V of the molecules (Sect. 13.1). For a longitudinally-vibrating column of gas, we thus obtain the distribution sketched in Fig. 11.31. Three phases are shown for the fourth normal mode in a tube which is closed at both ends. Gray shading indicates the normal number density, a lighter shading is a reduced density, and a darker shading is an in-

[3] Rotations of the gas within the boundary layer (Sect. 10.2 and Eq. (10.18)) cause the formation of two stationary vortex rings between each velocity maximum and the two neighboring nodes. The symmetry axis of these rings is oriented along the long axis of the tube. The sense of rotation in two neighboring vortex rings is opposite. At the walls of the tube, the stationary flows from the vortices approach each other in *those* sections of the tube where the velocity maxima of the longitudinally-vibrating column of gas are located. There, the static pressure in the boundary layer is higher. In *those* sections of the tube where the nodes are located, the stationary flows oppose each other. There, the static pressure in the boundary layer is lower.

C11.11. H. Rubens and O. Krigar-Menzel, *Ann. der Physik* **17**, 149 (1905). These authors also report that this effect is observed only at low sound amplitudes. At higher amplitudes, the maximal flame heights appear at the nodes of the motion. These maxima then vibrate at the frequency of the standing sound wave in the tube. RUBENS and KRIGAR-MENZEL explain the spatial vibrations at low sound amplitudes by effects in the boundary layer, although not in as much detail as given here. – In a similar experiment at the 3rd Physics Institute in Göttingen (Dr. D. RONNEBERGER), standing sound waves (at a frequency of 2800 Hz) in air in a plexiglass tube are shown. The tube contains a small amount of water along its bottom. In this experiment, pressure amplitudes of the order of $10\,\mathrm{N/m^2}$ (≈ 1 mm H_2O) are observed at the nodes, corresponding to the case of high sound amplitudes in RUBENS and KRIGAR-MENZEL's experiments.

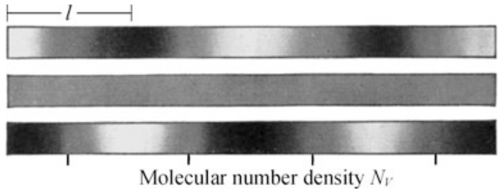


Figure 11.31 Three snapshot images of the distribution of the number density N_V along a closed gas column vibrating at its fourth natural frequency. Top and bottom, at the times when the density variation has its greatest amplitudes; center, at a time in between when the density distribution is uniform. The length l of the section shown in the top and bottom images corresponds to the images b and a in Fig. 11.26, i.e. a half wavelength[C11.12].

C11.12. In complete analogy to the longitudinal vibrations of a rubber band, in which ΔN_l vibrates (cf. Fig. 11.26), in a gas column, the number density N_V vibrates (and with it also the local static pressure) around a mean value. Note that correspondingly, in the n-th normal mode for N_V and for ΔN_l, the number of maxima is $(n+1)$. In a tube which is open at both ends, in which the gas column is vibrating at one of its natural frequencies, there are nodes of the number density (and the pressure) at the ends of the tube, and maxima of the flow velocity there (see **Exercise 11.2**).

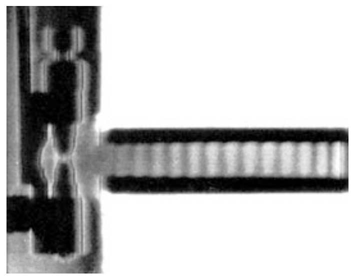


Figure 11.32 The periodic distribution of the number density N_V in a longitudinally-vibrating gas column excited by a flue pipe as in Fig. 11.35, photographed in a dark field with TOEPLER's schlieren method (Vol. 2, Sect. 21.11, Fig. 21.26). In the figure, above and below are reversed (The distance between the condenser and the wire-shaped slit in the exit diaphragm, b in Fig. 21.26, was 4.8 m).

creased density. This distribution, shown schematically in Fig. 11.31, can be demonstrated experimentally, most effectively with a *schlieren* method. We give an example in Fig. 11.32. It shows longitudinal normal-mode vibrations of an air column excited by a small flue pipe. The frequency is in this case about $4 \cdot 10^4$ Hz (corresponding to a wavelength of $\lambda \approx 8$ mm). Sound waves with frequencies above ca. $2 \cdot 10^4$ Hz are called *ultrasound* or *ultrasonic*.

Technically, normal modes of gas columns play an important role in the construction of pipes, whistles and wind instruments of all kinds. The usual types of design are – at least superficially – well known. Their function is rather complicated in detail and has been only qualitatively elucidated. In the case of the flue pipe, there is a periodic decay of the air flow past the lip (*labium*) into individual vortices. The air flow is controlled by the vibrations of the column of air in the pipe. A similar situation holds for the reed and the air column in a reed pipe. This mechanism causes deviations from the sinusoidal form in the pipe vibrations. Figures 11.33 and 11.34 show a still rather simple (nearly sinusoidal) pipe vibration and its line spectrum.

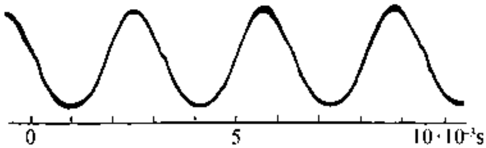


Figure 11.33 A nearly sinusoidal vibration in a pipe (image by F. TRENDELENBURG)

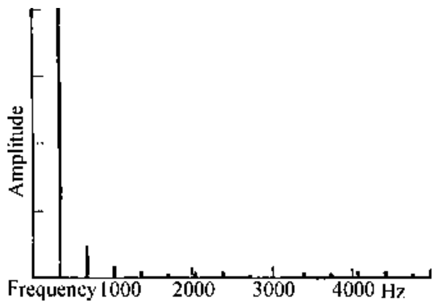


Figure 11.34 The spectrum of the pipe vibration shown in Fig. 11.33

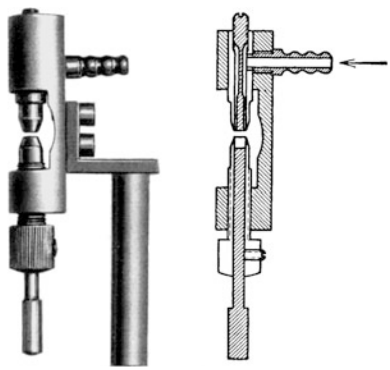


Figure 11.35 A flue pipe for frequencies from about 10^4 to $6 \cdot 10^4$ Hz. The lip slit and the lip are designed to be rotationally symmetric. The actual volume of the pipe represents only a very rough approximation to a linear column of air.

Flue pipes which produce high-frequency tones are an important tool for physics (Fig. 11.35). The right-hand image in the figure shows construction details of the pipe (essentially a whistle operated with compressed air).

11.8 Normal Modes of Stiff Linear Bodies

In considering the normal modes of solid, linear (one-dimensional) bodies, we have thus far treated only limp objects which had to be stretched by external forces, for example a rubber band. Demonstration of the normal modes of *stiff* linear objects, such as rods made of glass or metal, is more difficult. Such rods have to be held on knife-edge supports under two of the nodes, or hung there from thin cords.

Figure 11.36 Bending vibrations of a flat steel bar (87 cm long). The excitation was provided by a small electromagnet under the left end of the rod, which carried an alternating current at a frequency of 252 Hz. The nodes are made visible by scattering sand on the rod.

Video 11.4: "Longitudinal vibrations of a helical spring"
http://tiny.cc/ydgvjy
The video provides a supplement to this section, in the form of longitudinal normal modes or standing waves on a small helical spring; they are readily seen as a projected shadow image. This experiment was included in the book up to the 11th edition.

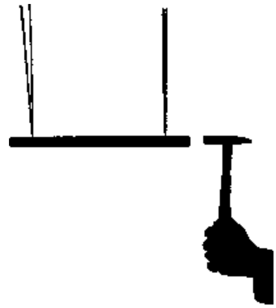


Figure 11.37 Longitudinal vibrations of a rod hanging from a cord (rod length $l = 25$ cm), fundamental frequency $\nu = c/2l$ (c = sound velocity in the rod)

Figure 11.36 illustrates *transverse* vibrations for such a case, using a flat steel bar. These oscillations are called *bending vibrations*.

Figure 11.37 shows a demonstration of *longitudinal* vibrations **(Video 11.4)** of a short, cylindrical steel rod. They are excited by striking one end of the rod. This impulse excitation yields a damped vibration. The cross-section of the rod remains unchanged at the nodes, but at the maxima, an expansion alternates periodically with a contraction. We can hear a tone which decays within a few seconds.

The frequencies of the first normal modes of these bodies obey the following relations:

For transverse vibrations a rod resting on both its ends[C11.13]

$$\nu_1 = 0.453\,\frac{h}{l^2}\sqrt{\frac{E}{\varrho}}\,, \tag{11.3}$$

for comparison: a stretched rope

$$\nu_1 = \frac{1}{2l}\sqrt{\frac{\sigma}{\varrho}}\,. \tag{11.4}$$

For longitudinal vibrations stretched rope or a rod whose ends are either fixed or both free[C11.14]

$$\nu_1 = \frac{1}{2l}\sqrt{\frac{E}{\varrho}}\,. \tag{11.5}$$

C11.13. The modes of vibration of a rod resting on its two ends (but not clamped there) are the same as those of a stretched rope, i.e. sinusoidal with nodes at the ends. But note that in this vibration mode, the first natural frequency is inversely proportional to l^2. Bending vibrations (see also Fig. 11.36) and their natural frequencies are rather complicated. For details, see e.g. S. Timoshenko, D.H. Young, and W. Weaver, *Vibration Problems in Engineering*, John Wiley, N.Y., 4th ed. (1974), Chap. 5.

C11.14. In the case of a stretched rope and the rod fixed at both ends, the nodes of the motion are held at the ends. With a free-standing rod, the node is at its center. Think of the vibrations of a column of air in pipes that are closed at both ends or open at both ends (organ pipes!). For longitudinal vibrations, see also Sect. 12.17.

For torsional oscillations held fixed at both ends (no rotation) or a free cylindrical rod[C11.15]

$$\nu_1 = \frac{1}{2l}\sqrt{\frac{G}{\varrho}}\,. \tag{11.6}$$

(l = length, ϱ = density, σ = tensile stress, h = thickness of the rod, E = elastic modulus, G = shear modulus, cf. Table 8.1.)

C11.15. When the ends of the rod are fixed, the nodes of the motion are held there; for a free-standing rod, the node is at its center.

The vibrations of short crystal rods, for example quartz crystals, have acquired considerable technological importance in recent years; in particular in the broad area of communications technology and for watches and clocks, they serve as a "microscopic pendulum".

The longitudinal vibrations of crystal rods are also suitable for generating standing waves in columns of liquid. These vibrations are excited electrically; to make the wave crests and troughs visible, a *schlieren* image with bright-field illumination can be employed, as shown e.g. in Fig. 11.32.

11.9 Normal Modes of 2-Dimensional and 3-Dimensional Bodies. Thermal Vibrations

We can be rather brief here. In 2 and 3 dimensions, we can also trace the occurrence of normal modes of vibration to the coupling of many elementary oscillators. But this would be, apart from a few exceptions, an exceedingly challenging business mathematically[C11.16]. In the majority of all practically important cases, one remains dependent on experimental observations. To detect the nodal lines, usually the accumulation of dust or sand scattered on the surface is employed. Figure 11.38 shows the nodal lines of a square and a circular metal plate in various modes of vibration.

C11.16. It can be solved numerically today using the methods of so-called Finite-Element Analysis (carried out by fast computers).

If, instead of scattered sand, one makes use of optical methods (refraction of polarized light), then considerably more complex normal-mode vibrations can be observed (see Vol. 2, Sect. 24.9, 'Strain Birefringence'). Figure 11.39 gives two examples for short glass cylinders with circular cross-sections.

Bowing of the plates leads to a bell shape. The vibrations of these relatively simple forms are already unpleasantly complicated. In the simplest case, a glass as seen from above vibrates according to the scheme shown in Fig. 11.40. At K, we see the crossover points of four nodal lines which take the form of "meridians". We can imagine that the simplest vibrations of our skull capsules, which harbor our hearing organs in their walls, must take a roughly similar form.

In the range of extremely high frequencies up to the order of 10^{13} Hz, all solid bodies, regardless of their form, possess an enormous number of elastic normal modes. Oscillations at such frequencies make

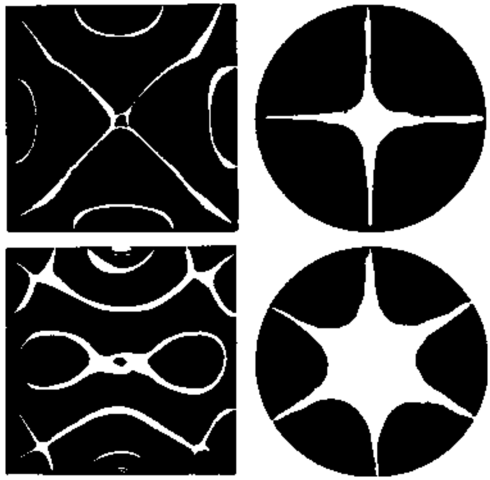


Figure 11.38 CHLADNI sound patterns (positive photographic images)[C11.17]

C11.17. The first publication of this research contained drawings of over 100 normal-mode vibrations of round and square plates (E.F.F. Chladni, "*Entdeckungen über die Theorie des Klanges*", Weidmanns Erben und Reich, Leipzig (1787)).

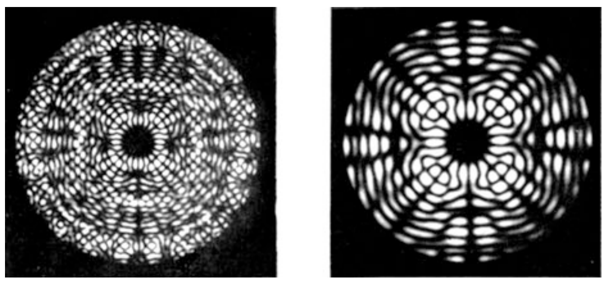
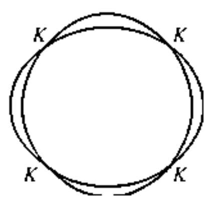


Figure 11.39 Glass cylinders vibrating at high frequency, observed with linearly-polarized light between two NICOL prisms. Their diameters are 30 and 44 mm and the excitation frequency is 1.54 MHz (images taken by L. BERGMANN)[C11.18]

C11.18. L. Bergmann, "*Der Ultraschall*", Verlag S. Hirzel, Stuttgart, 6th ed. (1954), pp. 636ff.

Figure 11.40 Simple vibrations of a wineglass, seen from above (schematic drawing)



up the disordered "thermal motions" (Chap. 16) in solid bodies or crystals. At the highest end of this frequency range, the individual atoms or molecules in the crystal lattice vibrate in the manner visualized roughly in Figs. 11.18 and 11.24 (right).

In considering the normal modes of gas-filled volumes, the natural frequencies and normal modes of air-containing spherical or bottle-shaped vessels with short, open necks are especially worth mentioning. These are "HELMHOLTZ resonators", which are important in metrology. They represent sound sources of a convenient form with well-defined fundamental frequencies[C11.19] **(Video 11.5)**.

C11.19. See e.g. T.B. Greenslade, "Experiments with Helmholtz Resonators", *Physics Teach.*, Vol. 34 (1996), p. 228.

For architects, the normal modes of large dwelling and meeting rooms are important. They are excited when people speak, sing or play music.

Video 11.5:
"HELMHOLTZ resonators"
http://tiny.cc/0cgvjy
The video was filmed in the historic collection of the 1st Physics Institute in Göttingen.

11.10 Forced Oscillations

Following a pulsed excitation, every oscillator vibrates at one or more of its *natural frequencies*. But one can also cause every oscillator to vibrate at other, arbitrary frequencies which are not equal to any of its natural frequencies. In this case, the oscillator is undergoing *forced oscillations*. Such forced oscillations play an extraordinarily important role in many areas of physics[C11.20].

C11.20. The torsional pendulum described in this section and used for demonstrating forced oscillations, known as "POHL's pendulum", can be found today in every collection of demonstration experiments and in every elementary physics lab course. See also Vol. 2, Sects. 11.7 and 26.2. A detailed mathematical treatment can be found in H.J. Pain, "*The Physics of Vibrations and Waves*", John Wiley, New York, 5th ed. (1999), Chap. 2.

To treat forced oscillations, we must first deal with the concept of *damping* of an oscillator more precisely than we have done so far. Due to unavoidable losses of energy or intentional energy outputs, the amplitude of every oscillator decays with time following an excitation pulse. The graph of the oscillation as a function of time is given by curves of the type illustrated in Fig. 11.42a (deflection α). In many cases, these curves follow a simple law for sinusoidally oscillating systems: The ratio of two maximum deflections or amplitudes on the same side of the curve remains constant all along the time curve. It is termed the *damping ratio* K. Its natural logarithm is called the *logarithmic decrement* Λ[C11.21]. The damping constants and the logarithmic decrements are given in Fig. 11.42a.

The reciprocal of the logarithmic decrement is equal to the number of oscillations after which the *amplitude* of the deflection has decreased to $1/\mathrm{e} = 37\,\%$ of its original value.

C11.21. This means that the decrease of the amplitude α is exponential:
$\alpha(t) = \alpha_0 \cdot e^{-\delta t}$,
with the logarithmic decrement Λ, defined by
$\delta T = \Lambda = \ln K$
(T = oscillation period). This follows mathematically from the fundamental equation for rotational motion (equation of motion, Eq. (6.7) in Table 6.1 for a frictional torque proportional to the angular velocity (see **Exercise 11.5**).

Keeping these definitions in mind, we now want to elucidate the essentials of forced oscillation using a demonstration experiment which provides clear and detailed examples. This experiment makes use of torsional oscillations at a rather low frequency: as usual, the slower the sequence of events, the clearer our observations. Figure 11.41 shows a suitable torsional oscillator (torsional pendulum). Its inertial element is a flat copper wheel with three spokes, whose hub is mounted on an axle D. A *spiral spring* is also attached to the axle, with its other end mounted on a lever at A; the lever can rotate back and forth around D. Via the connecting rod S and a motor-eccentric arrangement, the lever carrying the end A of the spring can be moved sinusoidally at a chosen frequency (the motor is geared down to a low rotational speed and can be further controlled by a rheostat) and variable amplitude. In this way, a sinusoidal torque of constant amplitude

Video 11.6:
"Free and forced oscillations of a torsional pendulum (POHL's pendulum)"
http://tiny.cc/uegvjy

Video 1:
"R.W. POHL Lecturing"
http://tiny.cc/fpqujy

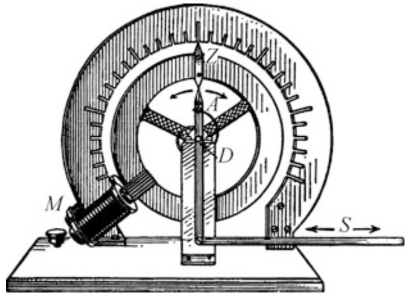


Figure 11.41 A torsional pendulum for demonstrating forced oscillations, with a torsion constant of $D^* = 2.3\,\mathrm{N/rad}$, moment of inertia $\Theta = 3.3 \cdot 10^{-3}\,\mathrm{kg\,m^2}$, and eigenfrequency $\nu_e = 0.42\,\mathrm{Hz}$. The spiral spring has several turns (not shown in the drawing). (**Video 11.6**; in **Video 1**, POHL himself demonstrates his torsional pendulum.)

and variable frequency can be applied to the axle D[C11.22]. The deflection of the torsional pendulum can be read off the angular scale outside the wheel. Below left at M is an arrangement which allows the damping of the wheel's motion to be adjusted within wide limits.

C11.22. By way of explanation: The total torque M acting on the axle D is determined by the angular deflection of the spiral spring. This deflection is $\beta - \alpha$, where $\beta = \beta_0 \sin(2\pi\nu t)$ is the angle that describes the motion of the lever around the axle, and α is the momentary angular deflection of the pendulum wheel. We then find for M: $M(t) = D^*(\beta(t) - \alpha(t))$ (D^* is the angular elastic constant (torsion constant) of the spiral spring). In the mathematical treatment of the torsional pendulum, this is the torque which enters into the equation of motion (Eq. 6.7). We find $\Theta\dot{\omega} + D^*\alpha(t) = D^*\beta(t)$ (Θ is the moment of inertia of the pendulum wheel). On the right-hand side is the additional sinusoidal torque generated by the lever acting at A, as described in the text. For $\beta = 0$, i.e. the rest position of the lever, we obtain the differential equation which describes the motion of a free torsional pendulum (without damping).

The damping mechanism operates by inducing eddy currents in the copper wheel. A small electromagnet has its pole pieces on each side of the rim of the wheel. The oscillating wheel moves through the magnetic field produced by the magnet, without touching the pole pieces. The current in the magnet coils can be varied to control the damping. See Vol. 2, Chap. 8.

Before beginning the actual demonstration, we first determine the natural frequency or *eigenfrequency* ν_e and the *damping ratio* K for the oscillator. For both measurements, the wheel is deflected with the lever in its rest position and its reversal points (maximum deflection) are registered.

Numerical example
The oscillation period $T_e = 2.39\,\mathrm{s}$; thus the eigenfrequency $\nu_e = 1/T_e = 0.42\,\mathrm{Hz}$. The ratio of two consecutive amplitudes is found to be nearly constant at 1.285; this is the damping ratio K. To illustrate this, the sequence of amplitudes to the left and the right at intervals of 1.2 s are graphed in Fig. 11.42a, and their endpoints are connected with a drawn-in line.

Now for the actual experiment: We set the "exciter" in motion – here the lever attached to the spring – and record its frequency ν, wait for the stationary state of oscillation to be established, and then record the forced amplitude α_0. Corresponding pairs of values of α_0 and ν are listed in Fig. 11.42b for four different damping ratios. The curves A, B, and C are somewhat asymmetric Gauss curves. These are the 'deflection curves for forced oscillations'.

When the damping is weak, *but only then*, the maximum amplitudes α_0 in the frequency range near the eigenfrequency of the pendulum are especially large compared to the amplitudes at other frequencies.

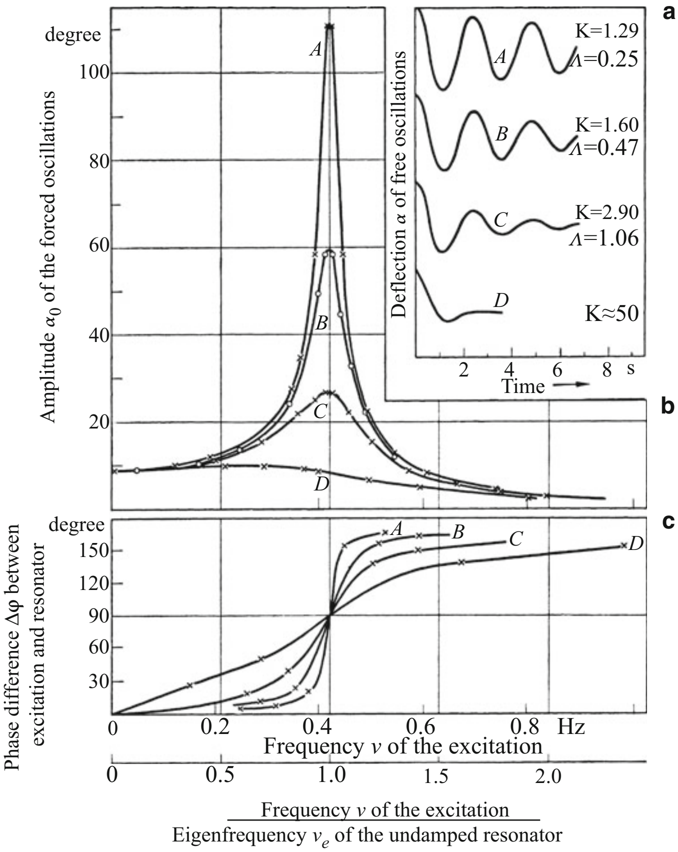


Figure 11.42 **a** The time dependence of free oscillations of the torsional pendulum in Fig. 11.41 with various values of the damping; curve *D* was recorded with a camera. The small increase in the period with increasing damping (at $\Lambda = 1$ only 4 %) is negligible compared to the precision of the present measurements. This allows us to refer to only one eigenfrequency in the text, $\nu_e = \omega_e/2\pi$, independently of Λ. **b** The deflection amplitudes of forced oscillations of the torsional pendulum at a constant excitation amplitude, as a function of the excitation frequency and the damping of the resonator. With increasing damping, the maximum value of α_0 at the resonance frequency ν_e decreases visibly. At an excitation frequency of zero, α_0 is simply determined by the two endpoints of the end of the spring (*A* in Fig. 11.41). **c** The influence of the excitation frequency and the damping on the phase difference $\Delta\varphi$ between the excitation and the resonator. The excitation is always ahead in phase. The measured points were taken from 'stopped-motion' photographic images. At the resonance frequency ν_e, the phase difference $\Delta\varphi = 90°$, independently of the damping. Note the optical illusion at the crossover point of the curves (see also Sect. 1.1). **(Video 11.6)**

Video 11.6:
"Free and forced oscillations of a torsional pendulum (Pohl's pendulum)"
http://tiny.cc/uegvjy

The ratio

$$V = \frac{\text{Amplitude at the eigenfrequency } \nu_\mathrm{e}}{\text{Amplitude at excitation frequency zero}}\,, \tag{11.7}$$

is called the *enhancement*; for curve A, it attains a value of 12.5. This special case ($\nu = \nu_\mathrm{e}$) is termed *resonance*. Referring to this term, any oscillator operated in the mode of forced oscillations may be called a *resonator*, and the curves A, B, C are *resonance curves*.

At large values of the damping (curve D), peculiar features are seen: The maximum of the resonance curve is barely visible, and it is shifted towards lower frequencies.

The resonance curves which we have observed experimentally for a particular example – our torsional pendulum – hold in fact quite generally. That is the reason why we have included two captions on the frequency axis in Fig. 11.42, parts b and c; the lower axis is independent of the particular numbers obtained with our demonstration apparatus. It gives the frequency of the excitation as a fraction of the eigenfrequency of the undamped resonator, whatever its nature may be. Thus, the curves apply not only to any kind of mechanical or acoustic forced oscillations, but also to those of electrical and optical systems with vastly differing resonance frequencies (see also Vol. 2, Sects. 11.7 and 26.2).

Given the universal applicability of these curves for forced oscillations of the most diverse systems and amplitudes (lengths, angles, pressures, electric currents, voltages, field strengths, etc.), one should keep in mind a simple general picture of how they come about. To this end, we consider another experimental observation: It concerns the influence of the excitation frequency on the *phase shift* by which the amplitude of the excitation (the end of the spring at A in our demonstration) precedes the amplitude of the resonator (the pointer Z in the demonstration). We thus have to record simultaneously both the position of the end of the spring at A and that of the pointer Z for the torsional pendulum in Fig. 11.41. Figure 11.42c shows the results.

At very low frequencies, the pointer Z and the excitation A move together, with zero phase shift, and both reverse their motions at the same time. As the excitation frequency increases, the amplitude of the excitation begins to move *ahead* of the amplitude of the torsional pendulum (the resonator) more and more.

At resonance, the phase shift is 90°. This means that *all along the path of the pendulum, the exciter tenses the spring in such a way that it always accelerates the motion of the torsional pendulum*. At the left-hand maximum rotation of the torsion wheel, the end of the spring at A is just moving out of its rest position to the right; the exciter thus produces a torque to the right[C11.23]. This torque reaches its maximum value (end of the spring A at its right-hand extreme value) when the torsion wheel is just passing through its rest position. The

C11.23. This is the torque $D^*\beta$; see Comment C11.22.

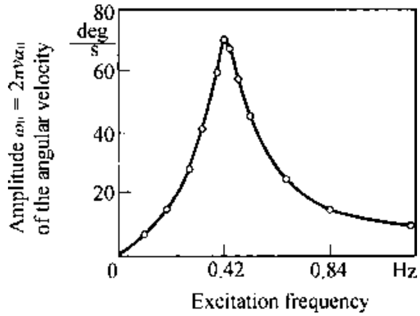


Figure 11.43 The resonance curve of the angular velocity of the torsion wheel as it passes its rest position ($\alpha = 0$, its maximum angular velocity ω_0), for curve C from Fig. 11.42b. At resonance, the maximum value of the angular velocity is $\omega_0 = 71$ deg/s $= 1.23$/s. This resonance curve has its maximum at ν_e, independently of the damping. This is also true of the resonance curve of the *energy*, since the latter is proportional to the square of the velocity.

torque is zero (the end of the spring again at its rest position) just at the moment when the wheel is reversing at its right-hand maximum rotation. For the oscillation of the torsion wheel from the right to the left, we have the same situation with a reversed sign. At resonance, the external torque, which leads the deflection α by a phase angle of 90°, is thus feeding energy into the torsional pendulum along its whole back-and-forth path. Without the losses due to damping, the oscillation amplitude at resonance would increase without limit ("resonance catastrophe").

Figure 11.43 gives the maximum angular velocity ω_0 of the torsion wheel for curve C in Fig. 11.42b, i.e. the amplitude of the angular velocity with which the wheel passes through its rest position. It is related to the maximum deflection α_0, i.e. the amplitude of the deflection, by the simple equation[C11.24]

$$\omega_0 = \alpha_0 \cdot 2\pi\nu \tag{11.8}$$

(ν = excitation frequency) .

The pendulum goes through its rest position with the maximum value of its kinetic energy E_0. It is given by

$$E_0 = \frac{1}{2}\Theta\omega_0^2 \tag{6.5}$$

(Θ is the moment of inertia of the torsion wheel).

The dependence of the energy amplitude E_0 on the excitation frequency is shown in Fig. 11.44. A graph of this kind is called the *energy resonance curve*[C11.25]. The kinetic energy E_0 vanishes at an excitation frequency of zero, in contrast to the maximum angular deflection α_0, i.e. a pendulum at rest stores no kinetic energy.

C11.24. Here, it is important to distinguish between the time-dependent angular velocity $\omega = \dfrac{d\alpha}{dt}$ and the constant angular frequency $2\pi\nu$ with which the resonator is being excited (i.e. the frequency of the external sinusoidal force). In the stationary state, for the deflection we have $\alpha(t) = \alpha_0 \sin(2\pi\nu t)$ and for the time-dependent angular velocity $\dfrac{d\alpha}{d\,t} = \alpha_0(2\pi\nu)\cos(2\pi\nu t)$, whose amplitude is denoted by ω_0 in Eq. (11.8).

C11.25. Away from the rest position, the total energy of the pendulum is given by the sum of its kinetic and potential energies. The resonance curve shown in Fig. 11.44 it thus identical to the curve for the *total energy* of the resonator.

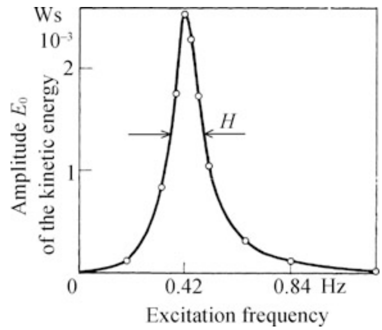


Figure 11.44 The resonance curve of the kinetic energy of the pendulum when the torsion wheel is passing through its rest position. If the excitation is turned off, this energy will be dissipated within about 8 s, as seen in curve C of Fig. 11.42a. The half-width H of the resonance curve is the frequency range at whose limits the amplitude of the kinetic energy is only half as great as its maximum at the (resonant) frequency ν_e. To a good approximation, $H = \Lambda\nu_e/\pi$.

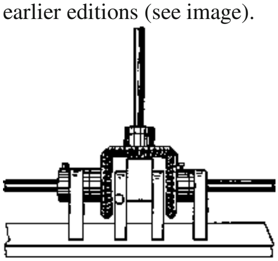
In considering the various, diverse applications of forced oscillations, one should keep in mind just which quantities are being plotted in the corresponding resonance curves.

11.11 Energy Transfer Stimulated by Resonance

In the previous section, we demonstrated forced oscillations. The excitation was provided by a lever which was moved periodically by an electric motor equipped with an eccentric. – Now, we put a mechanical phase shifter between the motor and the eccentric[C11.26], initially set at zero. Then we adjust the excitation frequency ν to be equal to the eigenfrequency of the resonator (resonance condition). The excitation then leads the deflection of the resonator with a phase shift of $\varphi = 90°$. The energy of the resonator is increased from its starting value of zero to its maximum value E_{max}.

Now we come to the new effect: We increase the phase shift from 90° to 270° without interrupting the operation of the pendulum or the excitation. As a result, the motion of the resonator along its back-and-forth path is continually retarded, and its stored energy is released to the excitation system, until the pendulum comes to rest. This process is called *stimulated energy transfer*. It plays an important role in optics and in electrodynamics[C11.27].

C11.26. POHL refers here to the phase shifter described in earlier editions (see image).



The phase shift of 180° required here could also be simply produced by shutting off the motion of the eccentric for half an oscillation period, e.g. from the maximum deflection of the pendulum on the one side to its maximum deflection on the other.

C11.27. For example in the LASER (Light Amplification by Stimulated Emission of Radiation); see e.g. R. W. Pohl, *Optik und Atomphysik*, Springer Verlag, Berlin, 13th ed. (1976), Chap. 14, Sect. 15; or Simon Hooker and Colin Webb, *Laser Physics*, Oxford University Press (2010), Chap. 2.

11.12 The Importance of Resonance for the Detection of Pure Sinusoidal Oscillations. Spectral Apparatus

According to the considerations in Sect. 11.10, the forced oscillations of an oscillator or resonator can attain very large amplitudes even when the periodic forces acting on them are quite small. This requires that the oscillator be only weakly damped, and its eigenfrequency must be as close as possible to the frequency of the excitation force. There are many demonstration experiments which show how amazingly large amplitudes can be produced in this way (see also **Video 11.7**). We will give another example here, using the forced bending oscillations of a leaf spring (Fig. 11.45). We have already employed its forced oscillations in Sect. 1.8, Fig. 1.8 to explain stroboscopic time measurements. The excitation was produced by an asymmetric axle passing perpendicularly through the holder of the spring.

Video 11.7:
"Forced oscillations with a pocket watch"
http://tiny.cc/zegvjy
This experiment was included by POHL up to the 11th edition (1947). The video shows how a pocket watch pendulum (with an eigenfrequency of 1.95 Hz) is excited to forced oscillations by its own escapement, which oscillates at a frequency of 2.00 Hz when the watch is at rest. The swinging of the watch reduces the frequency of the escapement to 1.95 Hz. Thus, the swinging watch produces a periodic excitation at its own resonance frequency, and this generates a large amplitude.

A series of such leaf springs or reeds mounted on a common base can function as a useful measuring instrument, a *vibrating-reed frequency meter* (Figs. 11.46 and 11.47). The oscillations being investigated are either coupled mechanically to the base (for example by attaching it to the frame of a vibrating machine) or more conveniently by using an electromagnet. Frequency meters of this kind are typical of *spectral apparatus*. They decompose an arbitrarily complex oscillation without regard to its phases into a spectrum of pure sine-wave components. From this spectrum, one can read off the frequencies of the individual sinusoidal components and their amplitudes. This can be most simply demonstrated using an electrical system. We offer two examples:

1. We refer to Fig. 11.13 and pass a rectangular or chopped direct current through the electromagnet of the vibrating-reed frequency meter. Figure 11.48a shows the arrangement, which needs no further explanation. The rotating commutator revolves once in the time $T = (1/20)$ s, and the current is switched on for a period of (1/40) s

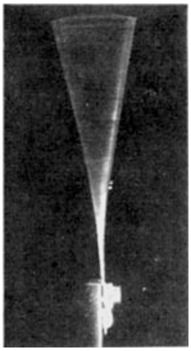


Figure 11.45 A leaf spring or reed which is being excited to forced vibrations (compare Fig. 1.13), time exposure

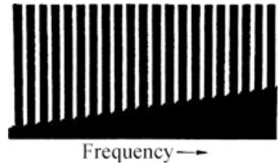


Figure 11.46 Sketch of a vibrating-reed frequency meter. The leaf springs have suitably graded resonance frequencies and each has a white, square knob at its tip (see Fig. 11.47).

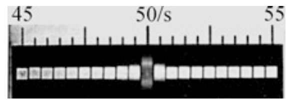


Figure 11.47 Section of the scale of a technical vibrating-reed frequency meter which shows the frequency of household alternating current (50 Hz)

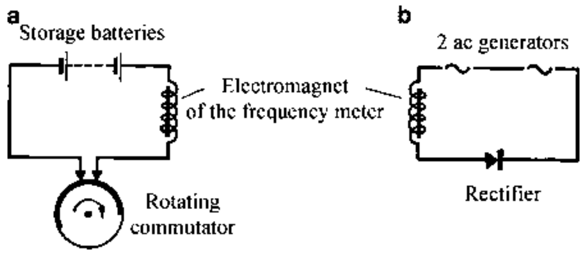


Figure 11.48 **a** Production of a rectangular direct-current waveform using a rotating commutator; **b** Production of beats with two sine-wave alternating-current signals, and cutting off the beat waveform with a rectifier

during each revolution. The spectral apparatus indicates the frequencies $\nu_1 = 20$ Hz and $\nu_3 = 60$ Hz. The next-higher frequency $\nu_5 = 100$ Hz can just be discerned in the spectrum.

2. We now refer to Fig. 11.12 and replace the direct-current source of Fig. 11.48a (storage batteries) by two alternating-current generators connected in series, with the frequencies $\nu_7 = 70$ Hz and $\nu_4 = 40$ Hz (Fig. 11.48b). The spectral apparatus indicates these two frequencies. – Then we add a rectifier to the circuit and cut off the beat curve on one side, so that it resembles the curve S_r in Fig. 11.12. Immediately, in addition to the two frequencies ν_7 and ν_4, the difference frequency $\nu_7 - \nu_4 = 30$ Hz is also indicated.

Difference oscillations or difference frequencies occur in general when some sort of "nonlinear element" is involved in the transmission of the oscillations; that is, when for example a force and the resulting deformation, or (as with the rectifier) voltage and current are not proportional to each other. A difference frequency occurs only rarely alone. Usually, other so-called *combination frequencies* are also present. They can be computed from the scheme $\nu_c = a\nu_1 \pm b\nu_2$ (a and b are small integers).

Experiments of this kind are quite important; they show that a non-sinusoidal oscillation process behaves like a physical mixture of its individual component oscillations. Every individual component can excite the leaf springs of the vibrating-reed frequency meter to forced vibrations, independently of all the others. This physical independence of the individual component oscillations plays a large role in all applications of forced oscillations. The next section gives an important example.

11.13 The Importance of Forced Oscillations for Distortion-Free Recording of Non-Sinusoidal Oscillations

In the majority of cases, our sensory organs suffice for simply *detecting* mechanical vibrations. Our bodies can for example register vibrations of the ground (ν ca. 10 Hz) already at a horizontal amplitude of only 3 μm. Our fingertips can feel vibration amplitudes of around 0.5 μm (at $\nu = 50$ Hz) with a gentle touch. We will give numerical details of the enormous sensitivity of our ears later in Sects. 12.28–12.30. In general, however, one is not satisfied simply with detecting vibrations; rather, a *true-to-form or distortion-free recording of their time dependence* is often desirable.

In every recording method, the vibrations in question produce motions in some sort of *probe instrument* (levers, mirrors, microphones, sensory organs). These motions are recorded continuously, for example (after electronic amplification) by an oscilloscope. The probe instrument is excited to forced oscillations; it is a resonator with the eigenfrequency ν_e.

For a flawless recording which correctly reproduces the time dependence of the process, two error sources must be avoided: First, the recording system must not give preference to certain frequency components of the oscillation due to their frequencies ν. Second, the phases of the individual components must not be shifted relative to one another. – Both requirements are fulfilled as long as ν is not greater than ca. $0.7\,\nu_e$ and the *damping ratio* K of the recording system[4] is ≈ 50.

Justification: From curve D in Fig. 11.42b, for the same excitation amplitudes, the amplitudes of the forced oscillations up to $\nu \approx 0.7\,\nu_e$ are *independent* of the frequency ν. – According to curve D in part c

[4] This statement does not hold for *seismographs* is $\approx$ 50. In their case, the eigenfrequency of the instrument (the "resonator") must be *low* compared to the frequencies of the seismic waves. This is for two reasons: First, the ground serves as an accelerated frame of reference; it generates inertial forces; these cause the ground to serve as the *excitation source* for forced oscillations. Second, the scale of the instrument (the resonator) is also part of the motions of the excitation source. As a result, at the same excitation amplitudes, the deflections shown by the instrument (the resonator) are small at low excitation frequencies and large at higher frequencies, since the resonator can follow the excitation source (the ground) less and less as the frequency increases, in contrast to curve D in Fig. 11.42b.

of Fig. 11.42, the phase angle $\Delta\varphi = 2\pi t/T = 2\pi t\nu$ up to $\nu \approx 0.7\,\nu_e$ is *proportional* to ν. Then, $2\pi t$ = const and thus t = const. This means that the components at all frequencies are delayed by the *same* time t, i.e. they are recorded *without* phase shifts relative to one another.

The recording of undistorted oscillation curves is rather challenging, but it is a task which is essential for many purposes in science. For acoustic/musical applications, for example for the production of musical recordings and their playback, the requirements are fortunately less stringent (compare Sect. 12.28).

11.14 The Amplification of Oscillations

For recording oscillations and numerous other tasks of vibration technology, *amplification* of the oscillations plays an important role: The essential characteristic of amplification is always that the oscillations themselves *control* the input of energy from an external source. This is accomplished today almost exclusively using electronics. Nevertheless, it is useful to comprehend the essentials of amplification with a clear-cut, purely mechanical example. For this mechanical amplification, the energy is not taken from an *electric current*, but instead from a current of *water*. This water stream is *controlled* by the vibrations which are to be amplified. This can be demonstrated successfully with the relatively primitive setup shown in Fig. 11.49. A jet of water from a glass nozzle flows nearly horizontally onto a strongly damped, stretched membrane (e.g. a tamborine). The jet forms a smooth filament; such a filamentary stream is a very *labile* object (Fig. 10.6). Tiny motions of the nozzle cause changes in the cross-section downstream, or interruptions of the flow. The changing impulse of the stream of water hitting the membrane excites the latter to rapidly decaying vibrations which can be heard some distance away. Thus, a light tap on the nozzle will be amplified to a loud throb from the membrane. Similarly, the vibrations of a small tuning fork held against the water nozzle, or the ticking of a watch, become clearly audible all over a large auditorium.

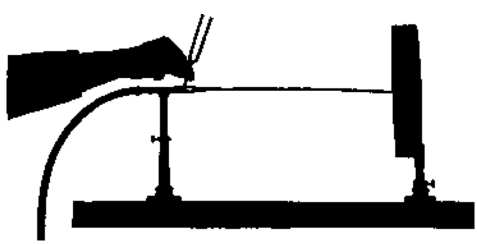


Figure 11.49 A stream of water as an acoustic amplifier. (In **Video 1**, POHL demonstrates the amplification, although silently, using a tuning fork which he presses against the water nozzle.)

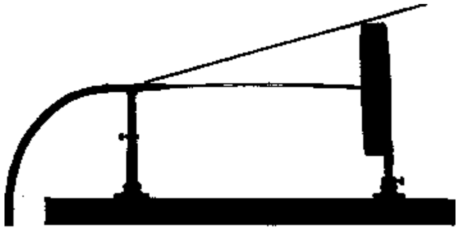


Figure 11.50 A water amplifier regulated by auto control of undamped vibrations (feedback). (This experiment is also demonstrated by the author in **Video 1** (again silently).)

Video 1:
"R.W. Pohl Lecturing"
http://tiny.cc/fpqujy

Electronic amplifiers are used technologically on a large scale to *generate undamped electrical oscillations*. Our mechanical amplifier can accomplish the same task by generating undamped mechanical vibrations. One need only set up a *feedback loop* between the resonator – here the membrane – and the glass nozzle. The vibrations of the membrane must be mechanically transmitted back to the nozzle; then the membrane *controls* the decay of the water stream with the rhythm of its own eigenfrequency. It suffices to lay a metal rod on the membrane and the nozzle, as in Fig. 11.50. Immediately, loud-sounding undamped vibrations are heard. Their frequency can be varied as desired, by adjusting the tension of the membrane and thereby changing its eigenfrequency.

11.15 Two Coupled Oscillators and Their Forced Oscillations

We now wish to investigate some of the properties of coupled pendulums, for simplicity considering only two pendulums, each of which has the same natural frequency. We distinguish three different types of coupling mechanisms[C11.28]:

1. Acceleration coupling (Fig. 11.51a). One pendulum is hanging from the other. It is located in the accelerated frame of reference of the first pendulum and is thus subject to inertial forces.

2. Force coupling (Fig. 11.51b). The two pendulums are connected to each other by an elastic element (spring).

C11.28. A mathematical treatment of the examples discussed here is given in Müller-Pouillet's "*Lehrbuch der Physik*", 11th edition, Vieweg und Sohn, Braunschweig (1929), Vol. 1, p. 55.

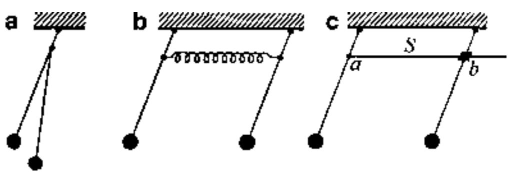


Figure 11.51 **a** Acceleration coupling; **b** Force coupling; **c** Frictional coupling. In the case of frictional coupling, no beats are observed. The first pendulum excites the second so that it begins to swing, and from then on, both swing together with the same amplitudes and phases **(Video 11.8)** (see **Exercise 11.7**).

Video 11.8:
"Coupled pendulums: Force coupling"
http://tiny.cc/pfgvjy
The video shows the normal mode vibrations of two coupled gravity pendulums (Fig. 11.51b) as well as their beats; and also the beating of a spring-torsional pendulum and a double gravity pendulum (Fig. 11.53).

Video 11.9:
"Acceleration coupling and chaotic oscillations"
http://tiny.cc/8cgvjy
The two individual vibrations of two coupled gravity pendulums (Fig. 11.52) as well as their beats are observed (**Exercise 11.6**). Two similar gravity pendulums coupled in a different manner exhibit a new form of vibrations at large amplitudes, so-called chaotic vibrations (see "Chaotic Vibrations, an Introduction for Applied Scientists and Engineers", Francis C. Moon, John Wiley, New York (1987)).

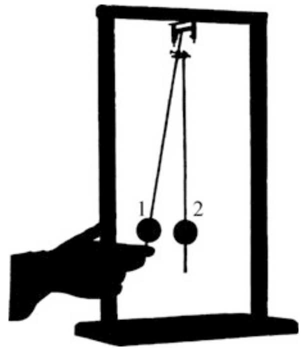


Figure 11.52 Two coupled gravity pendulums (**Video 11.9**)

3. Frictional coupling (Fig. 11.51c). A part of one pendulum, e.g. the connecting rod *S* which can rotate about *a*, rubs against a part of the other pendulum, e.g. the rotatable collar *b*.

In the following, we consider only the first two cases, that is acceleration coupling and force coupling. After the pendulums are coupled, the two eigenfrequencies which we have already seen in Sect. 11.5 are observed. The lower frequency ν_1 applies to synchronous swinging of the two pendulums, and the higher frequency ν_2 corresponds to swinging in opposite directions (analogously to the vibrations shown in Figs. 11.17, 11.18, and 11.24).

Now we come to a new observation: Initially, we move only *one* of the two pendulums (No. 1) away from its rest position and then release it (Fig. 11.52). Something surprising happens – pendulum No. 1 gradually passes all of its energy to the initially immobile pendulum No. 2 and excites it to a large vibration amplitude. Pendulum 1 itself comes to rest during this process. Then the cycle begins anew, with the roles of the two pendulums reversed (**Exercise 11.6**).

This process of energy transfer can be described in two ways: First, as *beats* of the two superposed eigenfrequencies ν_1 and ν_2[C11.29]. Second, as *forced vibrations at resonance*. Pendulum No. 1, which was initially released at its maximum deflection, serves as the excitation source for pendulum No. 2 as resonator, and leads it by a phase angle of 90°. It accelerates No. 2 along its whole swing with the correct sign. It is itself braked by the counter force according to *actio* = *reactio*. We are dealing here with forced vibrations with a strong negative feedback from the resonator to the excitation source.

C11.29. The effect seen in this example holds quite generally for coupled pendulums: All the vibrations can be described in terms of their normal modes (eigenfrequencies).

We offer two more examples of coupled oscillations:

In Fig. 11.53, two bifilar gravity pendulums with the same vibration period are hung from one another. The upper ball has a much greater mass than the lower one. If the upper one is given a small, hardly noticeable push, the lower pendulum, with its smaller mass, exhibits beats of a large amplitude. The beats of the larger-mass pendulum can barely be seen.

Figure 11.53 Two coupled gravity pendulums, hung in a bifilar arrangement, with inertial elements of very different mass. They can swing in a plane perpendicular to the page. **(Video 11.8)**

Video 11.8:
"Coupled pendulums, coupled by force"
http://tiny.cc/pfgvjy

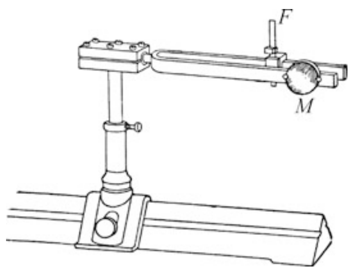


Figure 11.54 A tuning fork with an attached strongly damped leaf spring (MAX WIEN's demonstration, *Ann. d. Phys.* **61**, 151 (1897)) **(Video 11.10)**

Video 11.10:
"Coupled oscillations with damping"
http://tiny.cc/5egvjy

In the following examples, one of the two pendulum vibrations can be damped. A strongly-damped leaf spring (reed) is attached as a little 'rider' to a tuning fork. The arrangement is shown in Fig. 11.54. The damping of the reed is produced in the usual way by its rubber mounting. The reed and the tuning fork each have the same natural frequency.

First, we hold the reed with a fingertip to keep it from vibrating. Then the tuning fork's tone gradually dies away after it is struck once, over a time of around a minute. We can make its vibration readily visible from some distance away by using the mirror M. We then repeat the experiment, leaving the reed free to vibrate. After being struck, the tuning fork stops vibrating very quickly, after only a few seconds. Its energy of vibration, which is transferred to the coupled reed, has been dissipated as heat in the reed's rubber holder. Instead of the long beats heard with an undamped oscillator, now we can register only a few. When the spring is properly dimensioned, the energy will be completely absorbed by the end of the first beat.

So much for the *free* oscillations of two coupled pendulums. In technical applications, the *forced* oscillations of two coupled pendulums play a significant role. We limit our considerations to a single example, the reduction of rolling motions of ships in heavy seas[C11.30].

C11.30. Another example of such an application are the "vibration dampers" in tall buildings. These are large pendulums mounted within the buildings, whose vibration frequency is the same as the eigenfrequency of the structure, thus forming a system of coupled pendulums. When the building is excited to vibration by the wind, i.e. it begins to sway, energy can be transferred to the pendulum and damped there by friction. See Ch. Ucke and H.-J. Schlichting, "Schwingende Puppen und Wolkenkratzer", *Physik in Unserer Zeit* 3/2008(39), p. 139; or B. Breukelman and T. Haskett, "Good Vibrations", *Civil Engineering*, ASCE 71/2001(12), p. 55.

Video 11.11: "Antirolling tank"
http://tiny.cc/jfgvjy
Varying the throttle opening H controls the frequency of the oscillating water column so that it can be adjusted to the roll frequency of the ship's hull. The throttle provides the damping.

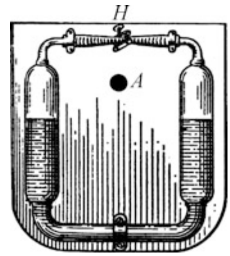


Figure 11.55 Model of an antirolling tank **(Video 11.11)**

We can think of the tuning fork in Fig. 11.54 as a steamship, and the leaf spring as a strongly damped pendulum built into the ship's hull. Furthermore, we imagine the pulsed excitation of the tuning fork to be the periodic action of the waves on the ship; then we can already recognize the principle. The strongly-damped pendulum is implemented technically as a water column in a large U-tube.

The model illustrated in Fig. 11.55 shows an *antirolling tank* of this type, attached as a pendulum to a board with the profile of a ship's hull, which can rotate (roll) around A. The two arms of the U-tube are connected above through an air pipe with a throttle valve H. When this valve is closed, the water column cannot oscillate. The board, that is the model ship, performs around 20 oscillations (rolls) after being tipped to 40°. By opening the valve, we can release the water column and at the same time provide damping of its oscillations. Now, the model stops rolling when tipped to 40° after only two or three oscillations, and returns to its rest position.

11.16 Damped and Undamped Wobble Oscillations

Spring pendulums and gravity pendulums (simple oscillators) exhibit a clear-cut, unified scheme for the occurrence of their oscillations:

Sinusoidal oscillations: **"These phenomena are given preferential treatment in every physics textbook, owing to their simplicity".**

There is a periodic exchange of potential and kinetic energy. In the ideal limit of vanishing losses, this exchange can continue indefinitely and sinusoidally at a frequency which is independent of the amplitude of the oscillations. These phenomena are given preferential treatment in every physics textbook, owing to their simplicity. – A number of oscillation or vibration phenomena however do not fit into this simple scheme, for example the "wobble oscillations" which are frequently observed in everyday life. In Fig. 11.56 (left), we see a wooden column which is standing with its two linear "feet" on a flat baseplate. A small impulse in the direction of the arrow lifts the right foot, tipping the column to the left (angle α), and thereby excites the wobble oscillations, a periodic exchange of kinetic and potential energy (the latter alternating between two metastable positions, tipped either to the left or to the right around the A–A axis). Even a superficial observation allows us to grasp the defining characteristic of wobble oscillations, namely the dependence of their frequency on

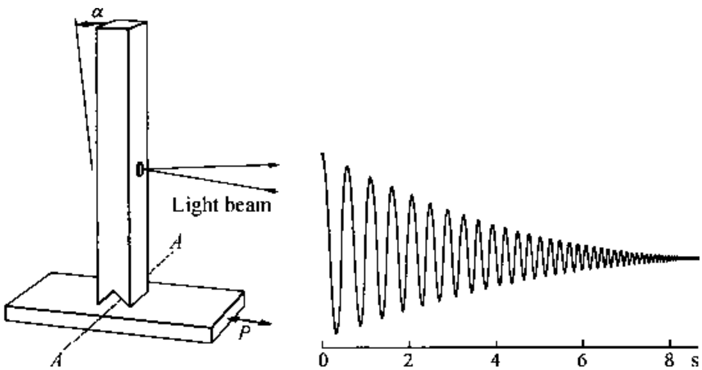


Figure 11.56 *Left*: Demonstration of wobble oscillations using a rectangular wooden column on a steel plate. The column is about 30 cm high and stands on two long "feet". *Right*: After an initial push, the oscillations of the column were recorded photographically using a light beam. **(Video 11.12)** – Forced wobble oscillations can be readily excited by employing inertial forces: the baseplate is periodically moved back and forth in the direction of the double arrow *P* by a motor with an eccentric.

their amplitude. The smaller the angular amplitude α_0, the higher the oscillation frequency becomes (compare Fig. 11.56, right).

Wobble oscillations can also be maintained with a constant amplitude, that is without damping. A first example of *external* control, i.e. *forced* oscillations, is described in the caption of Fig. 11.56. A second possibility is briefly illustrated in Fig. 11.57: *A* and *B* are two index fingers which are placed on the edge of a table *T* at a certain distance. A metal bar *S* is resting on them. Then *A* and *B* are moved up and down periodically, but with opposite phases as "exciters".
Externally controlled, i.e. *forced* wobble oscillations exhibit a notable peculiarity: Each excitation frequency corresponds to a certain amplitude, both of the resonator and also of the exciter. Therefore, there is no danger that the column in Fig. 11.56 will tip over, as long as the excitation frequency does not fall below a certain lower limiting value[5].
Wobble oscillations with *auto feedback* can also be illustrated using the example in Fig. 11.57. Now, *A* and *B* denote two sheets of lead attached

Video 11.12:
"Wobble oscillations"
http://tiny.cc/qcgvjy
In the video, besides the wooden column that tips back and forth around its centerline *A*-*A*, a metal disk is also shown (covered with plastic to enhance the visual effect). The disk is held at an angle on a flat glass plate and given a spin around its axis. It rolls around its circumference on the glass, at a continually increasing frequency, for about a minute, until finally coming to rest on the plate.

[5] In the St. Gumbertus church in Ansbach, Germany, the church towers are not attached rigidly to the nave. When their bells are rung, they are excited to wobble oscillations with a predetermined amplitude of 20 cm. This presented no danger to the rest of the structure (E. Mollwo). Nevertheless, this example of forced wobble oscillations was unfortunately eliminated by applying a rather simple structural change: The vertical plane in which the bell swings was rotated by 90°. Another example of wobble oscillations (or "rattling") is the "Rattleback" or "Celtic wobble stone" (see: https://en.m.wikipedia.org/wiki/Rattleback). This is a small, slightly asymmetric, boat-shaped piece of glass, plastic, stone or wood. When given a spin around its vertical center axis, it stops spinning after some time, and its kinetic energy is transferred to a different mode: it begins to wobble or "rattle" around a body axis. After some time, the energy is transferred back to a rotational mode and it again starts to spin, but in the reverse direction. This may be repeated one or two times. For a video demonstration, see e.g. https://m.youtube.com/watch?v=69Xm762qE8o.

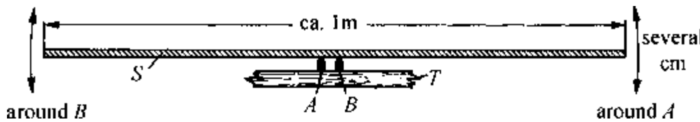


Figure 11.57 Producing wobble oscillations of constant amplitude, either with external excitation or by using thermal feedback (TREVELYAN's rocker)

to a metal block (at a spacing of ca. 6 mm), and S is a *hot* metal bar, made e.g. of copper. After an initial push, the bar touches first the left and then the right lead sheet, each time causing it to bulge more deeply downwards, which accelerates the rod's motion with the correct phase ("TREVELYAN's rocker" or thermophone)[C11.31].

C11.31. A. Trevelyan, *Trans. Roy. Soc. Edinburgh*, Vol. 12 (1834), p. 137. An illustration can be seen in the same volume in front of page 429. See also I.M. Freeman, "What is Trevelyans Rocker?", *The Phys. Teacher*, Vol. 12 (1974), p. 382; and https://www.youtube.com/watch?v=U23iwbVX-Dk for a video demonstration.

11.17 Relaxation (or Toggle) Oscillations

Wobble oscillations can continue indefinitely after an initial pulse excitation, at least in the ideal limit of negligible losses. – Another important group of oscillations, however, cannot occur at all without a continuous input of energy: This is the group of relaxation oscillations or "toggle oscillations". Relaxation oscillations occur whenever there is a delay time (relaxation time) between the storage of potential energy and its conversion into kinetic energy in an oscillator. At the end of this delay time, the stored energy is released; however, the released energy is not used again to refill the storage element; rather, a new quantity of energy is taken from the continuously-available external source.

Figure 11.58 shows a simple example. A liquid container with the profile of a non-equilateral triangle is mounted as a seesaw. It can rock back and forth between two resting points a and b. A steady stream of water fills the container and shifts its center of gravity to the left. At a certain point, the 'seesaw' becomes unstable and tips to the left, emptying the contents of the container, which then tips back to the right. The cycle begins anew.

Figure 11.59 shows an analogous setup in the form of an electrical circuit. A condenser is slowly charged from a current source through a control resistor with a high value, until the ignition voltage of a glow-discharge lamp in parallel with the condenser is reached. Then the lamp flashes on and discharges the condenser quickly. Another, similar example of electrical relaxation oscillations at a low frequency, $\nu \approx 0.1$ Hz, are the spark discharges which can be observed with an influence machine (Wimshurst machine) equipped with a Leyden jar.

Relaxation oscillations are distinguished from ordinary oscillations or vibrations (periodic inter-conversion of potential and kinetic energy) by two essential characteristics:

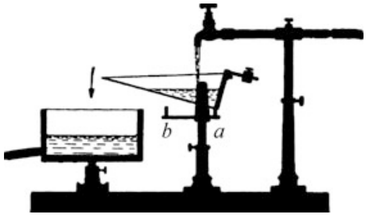


Figure 11.58 Mechanical relaxation oscillations. *a* and *b* are resting points for the "seesaw" container.

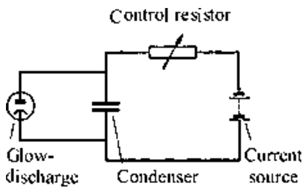


Figure 11.59 Electrical relaxation oscillations

1. Their *amplitudes* cannot be changed without modifying the construction of the oscillator (e.g. the container in Fig. 11.58); only their *frequencies* can be varied (in Fig. 11.58 by opening or closing the water faucet, in Fig. 11.59 by adjusting the variable control resistor).

2. Relaxation oscillations can be readily controlled by an auxiliary oscillation of small amplitude. As a result, they can easily be synchronized.

In everyday life, relaxation or toggle oscillations are extremely important. We can observe them for example in the creaking of a door or the screeching of chalk on a blackboard, and in the operation of a pneumatic hammer. In particular, relaxation oscillations play an especially important role in the lives of many organisms. The excitation of nerve cells and muscular activity are examples. The functioning of the human heart can be elucidated down to the finest details in terms of three coupled relaxation oscillations.

Unfortunately, the mathematical treatment of relaxation or toggle oscillations is complex and difficult; it is therefore not dealt with as extensively as it deserves in the literature of physics. However, the electrical demonstrations mentioned above have contributed notably to the understanding of relaxation oscillations.

Exercises

11.1 Sound waves with the frequencies $\nu_1 = 225\,\text{Hz}$ and $\nu_2 = 336\,\text{Hz}$, each containing their first two overtones at 2ν and 3ν, are sounded together. Show that two of these overtones lead to beats of frequency $\nu_B = 3\,\text{Hz}$. (Sects. 11.3 and 11.7)

11.2 Determine the length d of the shortest capillary tube whose air column can be excited to normal-mode vibrations by a tuning fork at a frequency of 520 Hz, if a) both ends are open; and b) if one end of the tube is closed (sound velocity in air: $c = 340\,\mathrm{m/s}$). (Sect. 11.7)

11.3 KUNDT's dust figures are produced by a whistle; their spacing is found to be 1.2 cm. Determine the wavelength and the frequency of the sound. (Sect. 11.7)

11.4 The longitudinal fundamental vibration of a rod of length $l = 2.4\,\mathrm{m}$ is 800 Hz. Evaluate the velocity of sound c_1 in the rod. (Sect. 11.8)

11.5 Video 11.6, "Free and forced oscillations of a torsional pendulum (POHL's pendulum)" begins with the free oscillations of the torsional pendulum (Fig. 11.42a). Find from this result the value δ^{-1} (i.e. the time constant τ; cf. C11.21.) of the exponential decay of the amplitude, and the frequency ν_e of the torsional pendulum, and derive the logarithmic decrement Λ, the damping ratio K and the half-width H of the resonance curve of the energy (Fig. 11.44). (Sect. 11.10)

11.6 In Video 11.9, "Acceleration coupling and chaotic vibrations": For the two coupled gravity pendulums ("acceleration coupling"), use a stopwatch to determine the frequencies of the symmetric and the antisymmetric normal modes of vibration and the beat frequency. Then investigate the relationship between these three frequencies, which is hinted at in Fig. 11.10 (see also Comment C11.4). (Sect. 11.15)

11.7 Figure 11.24 shows the longitudinal vibrations of two coupled ball-and-spring pendulums with massive balls m and spring constants D. a) Calculate the frequency ν_0 of one of the pendulums when the other is being held fixed, and also the frequencies ν_1 and ν_2 of the vibrations of the coupled pendulums. b) Replace the spring between the two balls by a weaker spring, i.e. $D' \ll D$, which represents a situation very similar to that shown in Fig. 11.51b (note that here again, one should observe the vibrations of one of the pendulums while the other is being held fixed). Determine the three frequencies of the longitudinal vibrations in this case. (Sect. 11.15)

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_11) contains supplementary material, which is available to authorized users.

Travelling Waves and Radiation

12

Part II

12.1 Travelling Waves

In Fig. 12.1, we are looking through a window to see a sine curve as a shadow image. It is moving with the velocity c in the z direction: This shadow image thus represents a *travelling sinusoidal wave*. Figure 12.1 shows a 'snapshot' of the wave. We distinguish between wave crests and wave troughs. The distance between two corresponding points along the wave, e.g. the crossing points α and β on the z axis, or two successive wave crests, is called the *wavelength* λ. The velocity with which such a crossing point or wave crest is moving along the z axis is called the *phase velocity* c of the wave.

We can represent such a travelling wave experimentally by pulling a sine curve made of wire at the velocity c past the window. A different arrangement is better, since it demonstrates the relation between a sinusoidal oscillation and circular motion: We put a wire which has been bent into a helical shape behind the window (Fig. 12.2). Then it is rotated around its long axis (by turning the crank), with the frequency $\nu = N/t$, that is N revolutions within a time t. During N revolutions of the helix, a point followed by the eye, e.g. the crossing

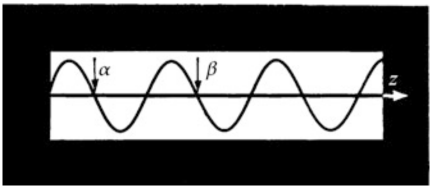


Figure 12.1 Instantaneous image ('snapshot') of a sinusoidal travelling wave

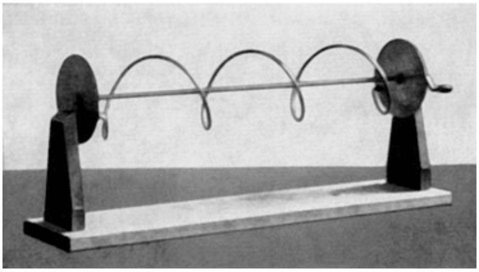


Figure 12.2 A wire which has been bent into a helix and can be rotated around its long axis

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_12

Video 12.1:
"Model of a travelling wave"
http://tiny.cc/2fgvjy
The direction in which a wave is travelling as given in the text depends on the sense in which the helix is wound (right- or left-handed), and on the sense in which the experimenter is turning the crank (clockwise or counter-clockwise). This can be clearly seen in the video.

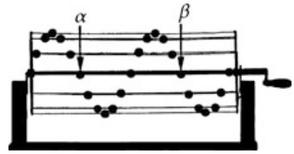


Figure 12.3 Beads arranged along a helix **(Video 12.1)**

point α in Fig. 12.1, moves along a distance $s = N\lambda$ at the phase velocity $c = s/t = N\lambda/t$ or

$$c = \nu\lambda\,. \tag{12.1}$$

This is a fundamental relationship for every wave phenomenon. – During the rotation of the helical wire, every observer sees the wave weaving along like a snake in the z direction. Nevertheless, no point on the wire is actually moving in the direction of travel z. All the points on the wire are moving circularly in planes which are perpendicular to the z axis. This can best be seen by imagining that the wire has been dissolved into a series of individual points which follow the helical form, as in Fig. 12.3 (small wooden beads on a thread). Each point undergoes a sinusoidal oscillation with the period T. At a particular time t, it has a phase of $\varphi = 2\pi t/T$. Its neighbor to the right has the same phase somewhat later: what is in fact moving to the right along the wave is its *phase*, and thus the name *phase velocity*.

For a quantitative description, we first need to examine the *vibrations* of a single point. Assume that the point α is just touching the z axis at the time $t = 0$. Then after a time t, it has reached a deflection of

$$x = x_0 \sin \omega t \tag{12.2}$$

(x_0 is the maximum possible deflection (amplitude), $\omega = 2\pi\nu$ is the circular frequency).

For the *vibrations*, the deflection x and die phase ωt depend only on the *time*. The deflection is repeated at the same *position* after each period T.

In order to arrive at a description of the *wave*, we consider the vibrations of a point at a distance z further to the right. Along its path upwards, it crosses the abscissa later than the point α. Its *vibrations* are delayed by a phase angle of

$$\varphi = 2\pi\frac{t}{T} = 2\pi\frac{ct}{cT} = 2\pi\frac{z}{\lambda}$$

relative to those of the point α. Therefore, instead of Eq. (12.2), they are described by:

$$x = x_0 \sin\left(\omega t - 2\pi\frac{z}{\lambda}\right)$$

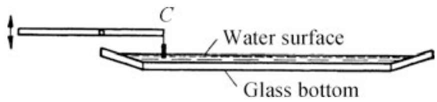


Figure 12.4 A ripple tank for observing surface wave fields. At the end of a lever is an immersion contactor C which can be moved up and down by means of a vibrator (after THOMAS YOUNG, see Comment C12.3). For circular waves, the contactor is a short, pointed rod; for linear waves, it is a horizontal bar **(Video 12.2)**

Video 12.2:
"Experiments with water waves"
http://tiny.cc/tfgvjy

or, with $\lambda = c/\nu$ and $2\pi\nu = \omega$,

$$x = x_0 \sin\omega\left(t - \frac{z}{c}\right) . \tag{12.3}$$

This is the specification of a travelling wave in terms of an equation. For the *wave*, the deflection x and the phase $\omega(t - z/c)$ depend not only on the time t, but also on the position z. The deflection repeats itself at the same *position* after each period T and at the same *time* at positions separated by $z = \lambda$.

Any state which travels at a finite velocity can form waves. To observe waves, it is expedient to begin with states which travel at small velocities c. Among these, small deformations of liquid surfaces occupy a prominent position; they form surface wave fields. These can be observed in a *ripple tank*.

In Fig. 12.4, a ripple tank is sketched in cross-section[C12.1]. For producing the waves, it contains an immersion contactor which vibrates sinusoidally up and down at the surface of the water and serves as a *wave source*. Its frequency is chosen to lie between 10 and 20 Hz; then we obtain wavelengths between 2.5 and 1.2 cm. Pointed contactors give circular waves, which travel out as concentric circles (Fig. 12.5). Linear contactors (bars) give wave crests and troughs in the form of straight lines (linear waves), as seen in Fig. 12.11.

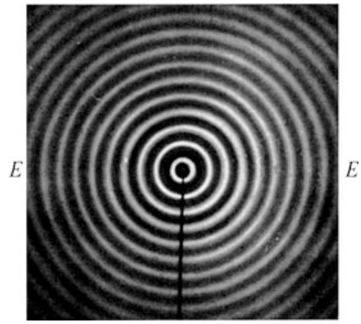
C12.1. Many of the experiments described in the following were filmed using this ripple tank **(Video 12.2)**. In order to aid in finding the individual experiments, the times (in minutes) at which they appear in the video are noted in the margin of the page (the complete video runs for somewhat more than 6 minutes).

The edges of the tank must be flat with a gentle slope, so that the waves are damped and no disturbing reflections occur there. – Illumination from below can be used to project an image of the waves in the tank onto the ceiling or, using a mirror to redirect the image, onto a wall screen. Stroboscopic illumination allows a particular phase, e.g. a certain wave crest, to be easily followed in the projected image.

12.2 The DOPPLER Effect

In Fig. 12.5, we can imagine that somewhere, there is a receiver E which responds to the oncoming waves and can count them; it could be for example the eyes of an observer. If the source and the receiver of the waves are at rest relative to the wave-carrying medium,

Figure 12.5 A two-dimensional wave field on the water surface, snapshot (exposure time 0.002 s). The wave crests and troughs are concentric circles **(Video 12.2)**



Video 12.2: "Experiments with water waves"
http://tiny.cc/tfgvjy
http://tiny.cc/tfgvjy
Circular waves (3:00). Note that the wavelength depends not only on the excitation frequency, but also on the phase velocity, which is changed when the depth of the water decreases. (This effect is used to demonstrate refraction of waves in Figs. 12.19, 12.20, 12.22, and 12.23.)

here the surface of the water, then the receiver measures the same frequency as was emitted at the source. If the source or the receiver are moving, then the DOPPLER effect is observed: During the propagation of the waves, a decrease in the spacing of the wave crests increases the frequency observed by the receiver, while an increase in their spacing reduces the frequency. – For a quantitative discussion, one must be careful when dealing with *mechanical* waves to keep the case of a moving source (Fig. 12.6) separate from that of a moving receiver[C12.2].

C12.2. In contrast to light waves, for which only the *relative* motion is of importance (see Vol. 2, Sect. 23.4).

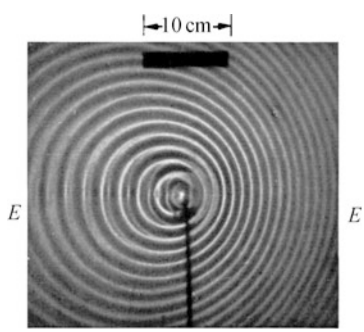
Let the relative velocity of source and receiver be u, and the phase velocity of the waves be c. Then the following observations can be made:

Receiver at rest and a moving source
Instead of the source frequency ν, the receiver observes a wave frequency of:

$$\nu' = \frac{\nu}{1 \mp \frac{u}{c}} = \nu\left(1 \pm \frac{u}{c} + \cdots\right) \tag{12.4}$$

In a time t, $N' = \nu t$ individual waves (i.e. one crest and one trough) are emitted from a moving source which is approaching the receiver. Along their path $(ct - ut)$, they are compressed; therefore, the receiver measures

Figure 12.6 Similar to Fig. 12.5, but with the wave source moving to the right. E denotes a "receiver" which is at rest. A scale 10 cm long is shown.



a wavelength of $\lambda' = (c-u)t/N' = (c-u)/\nu$. Inserting $\lambda' = c/\nu'$ yields Eq. (12.4).

Moving receiver and a source at rest

$$\nu'' = \nu\left(1 \pm \frac{u}{c}\right) \qquad (12.5)$$

If νt individual waves would arrive within the time t at a receiver which is at rest, then for a receiver which is moving towards the source by a distance ut during the time t, an additional ut/λ individual waves arrive. The moving receiver will thus detect $N'' = (\nu t + ut/\lambda)$ individual waves in the time t. It observes a frequency of $\nu'' = N''/t$, giving with $\lambda = c/\nu$ Eq. (12.5).

In Eqns. (12.4) and (12.5), the upper sign applies when the source and the receiver are approaching each other.

12.3 Interference

Using a ripple tank, THOMAS YOUNG[C12.3] discovered and named a fundamental effect in 1802; it is basic to the understanding of all wave phenomena. This is *interference*, which occurs when two waves are superposed. – In order to demonstrate this effect, we use two points which are rigidly connected together as excitation sources in the ripple tank. The result is shown as a snapshot in Fig. 12.7: The two waves add to each other, and the resulting wave field is chopped up by interference maxima and minima.

Along the symmetry direction *00*, there are travelling waves. On both sides of this symmetry axis *00*, one can observe alternating narrow, dark regions with no waves and broad regions containing waves. In the wave-free regions – called for short *interference fringes* – for every point along their center lines, the path difference, i.e. the difference of the distances to the two wave sources, is constant and is an odd multiple of $\lambda/2$. As a result, the two waves cancel each other.

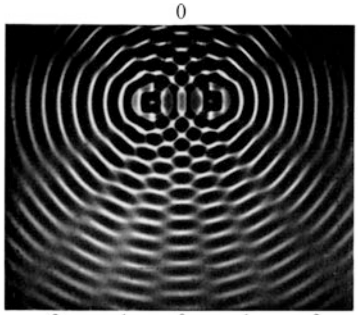
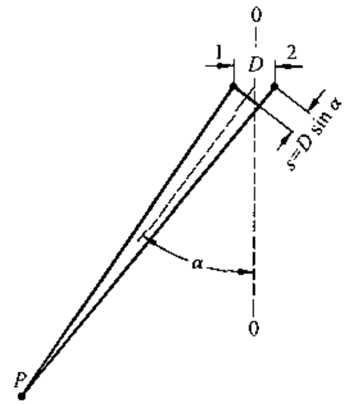


Figure 12.7 Snapshot of a two-dimensional wave field showing interference between two waves (exposure time 0.002 s). This image and all those on through Fig. 12.19 are photographic positives **(Video 12.2)**.

C12.3. THOMAS YOUNG, physician (1773–1829). An appraisal of his scientific achievements can be found in R.W. Pohl, *Physikalische Blätter* **5**, 208 (1961).

Video 12.2: "Experiments with water waves"
http://tiny.cc/tfgvjy
YOUNG's experiment (3:30) is easily reproduced, in particular with long wavelengths. One can clearly discern two directions of negative interference. When the frequency is increased (shortening the wavelength), the number of such directions increases; however, the images also become more complex. The reason for this is that the waves excited by an object which oscillates harmonically at the water surface are by no means simple sinusoidal waves with only one wavelength, as one can see in many of the experiments in this video. Their interference patterns are correspondingly complex.

Figure 12.8 The derivation of Eqns. (12.6) and (12.7)



In contrast, the waves between the interference fringes are amplified. Along the center lines of the wave-containing regions, the waves coming from the two sources add with the same phase, and thus their path differences are integral multiples of λ, as shown by the numbers in the margin of Fig. 12.7. – The curves of equal path difference, i.e. the center lines of both the wave-free and the wave-containing regions, are hyperbolas. The angle between their asymptotes and the symmetry axis *00* can be read off Fig. 12.8 at a sufficiently large distance from the sources. For the maxima (at angle α), where m is a whole number (an *ordinal number* or *index*), we find

$$\sin\alpha = \frac{m\lambda}{D}\,, \tag{12.6}$$

and for the minima

$$\sin\alpha = \frac{(2m+1)\lambda/2}{D}\,. \tag{12.7}$$

The amplitudes of neighboring wave trains have opposite signs (crest instead of trough and *vice versa*). More details are given in Vol. 2, Fig. 12.35.

12.4 Interference with Two Slightly Different Source Frequencies

Suppose that in Fig. 12.7, the two exciter points are no longer rigidly connected to each other, but instead are driven independently by two motors, at the frequencies ν and $\nu+\Delta\nu$. Then the interference fringes will *wander*. In a periodic sequence, they arrive at the same position previously occupied by the neighboring fringe (n times within a time t, or at the "beat frequency" $\nu_B = n/t = \Delta\nu$). The interference fringes exhibit fixed positions only in stopped-motion images.

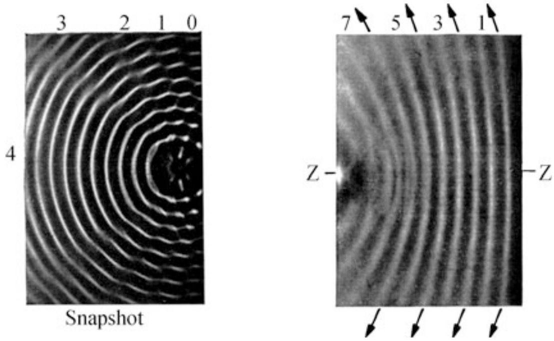


Figure 12.9 Two two-dimensional wave fields showing interference, in which the second wave was produced by reflection of the first from a wall; left: an instantaneous image, right: a time exposure, with a larger spacing of the two wave sources. The *left-hand* image corresponds to the left half of Fig. 12.7. The right margin of the picture corresponds to the surface of the reflecting wall, and to the line *0–0* in Fig. 12.7. The *arrows* in the *right-hand* image show the direction of propagation of the travelling waves in the interference field. Furthermore, the wall at the right and the wave nearest to it, with a path difference of 0, are not shown.

In many cases, it is impossible to arrive at strict synchrony between two sources and thus to avoid frequency differences $\Delta\nu$ between them. Then, one can resort to a trick: The second source is replaced by a *mirror image* of the first, guaranteeing synchrony. That is, we allow a wave to reflect from a smooth wall and observe the superposition of the reflected and the incident waves. Figure 12.9 (on the left) shows an example.

12.5 Standing Waves

In Fig. 12.9, at the left we see the interference of two waves of the same frequency as an *instantaneous image*. We now show a similar interference pattern as a *time exposure* (Fig. 12.9, right). Here, the second wave source was also produced as a mirror image of the first. *In this time exposure, we see nothing more of the progressive sequence of wave crests and troughs; we see only the hyperbolas of equal path differences.* Between the dark interference fringes, the waves travel in the directions shown by the short arrows; above the line *ZZ* upwards, and below the line *ZZ* downwards. Along the line *ZZ* (and practically also in its immediate neighborhood), we observe *standing waves* with *nodes* and *maxima*: At the maxima, crests and troughs follow each other in a periodic sequence. At the nodes, the medium (water surface) is at rest[1] Along the connecting line *Z-Z*

[1] At the nodes, a stick oriented perpendicular to the water surface would change only its angle to the vertical over time; its lower end would not move up and down.

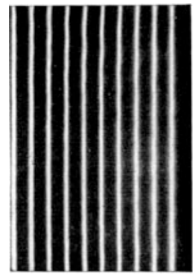


Figure 12.10 A time exposure of linear standing waves in front of a wall which is at the right. The bright regions (spacing $= \lambda/2$) are formed by the maxima when the water surface is bowed *upwards*. Therefore, they appear *periodically* at the frequency of the waves. **(Video 12.2)**

Video 12.2:
"Experiments with water waves"
http://tiny.cc/tfgvjy
Standing waves form by reflection from a wall at perpendicular incidence (1:00). Examine the images individually and observe the footnote 1.

between the two wave sources, the two waves move exactly oppositely towards one another. In a strict sense, one can speak of standing waves only in this situation, since only then is the spacing of two interference minima at its smallest value, namely $\lambda/2$. Standing waves can be produced in a simple manner experimentally on a water surface; one need only to cause linear waves to reflect with perpendicular incidence from a flat wall (Fig. 12.10).

The equation for the standing waves can be derived as follows: For the wave travelling to the right in the positive z direction, we have

$$x_r = x_0 \sin \omega \left(t - \frac{z}{c}\right) . \qquad (12.3)$$

The wave travelling to the left is described by

$$x_l = x_0 \sin \omega \left(t + \frac{z}{c}\right) . \qquad (12.8)$$

We abbreviate $\left(\omega t - \omega \frac{z}{c}\right) = \alpha$ and $\left(\omega t + \omega \frac{z}{c}\right) = \beta$, and obtain for the resulting wave function, resulting from the two oppositely-travelling waves

$$x = x_r + x_l = x_0(\sin \alpha + \sin \beta) . \qquad (12.9)$$

We then make use of the trigonometric identity

$$\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2} \qquad (12.10)$$

to obtain

$$x = 2x_0 \sin \omega t \cos \omega \frac{z}{c} \qquad (12.11)$$

or

$$x = 2x_0 \cos 2\pi \frac{z}{\lambda} \sin \omega t . \qquad (12.12)$$

This is the equation of a sinusoidal oscillation whose amplitude $2x_0 \cos 2\pi \frac{z}{\lambda}$ changes periodically along the z direction.

Standing waves on liquid surfaces can be "parametrically" excited by causing the container to oscillate in the vertical direction. The wave frequency is equal to half the excitation frequency (cf. the caption of Fig. 11.20). In this way, one can observe $\lambda < 0.1$ mm on water surfaces[C12.4].

C12.4. See W. Eisenmenger, *Physikalische Blätter* **51**, 655 (1995).

12.6 The Propagation of Travelling Waves

How do *travelling* waves propagate? Why do we speak of the *emission* of waves? – Again, we use the ripple tank to help us answer these questions. We set obstacles (pieces of wood or metal) in the path of the waves.

First of all, we let linear waves with a broad wavefront fall onto a long wall at perpendicular incidence, as in Fig. 12.11, in order to block off "half" of the waves. From symmetry arguments, we would expect the dashed line in the figure to be the boundary of the waves in this "semi-plane". Beyond this line, the "shadow" of the wall should begin. But in fact we observe something quite different. The waves pass over this geometric boundary and move along large arcs into the shadow region. (For waves, their wavelength λ plays the role of a characteristic dimension. Here, "large" thus means "large compared to λ".) This path of the waves is described in words, strangely enough, in the passive form; we say that the waves "have been diffracted". The waves which appear beyond the shadow boundary are *diffracted waves*.

In Fig. 12.12, we have made a slit in the obstacle to wave propagation by using two semi-planes, above and below, with a gap between them. Again, the dashed geometric boundaries are clearly crossed over by the waves. The connection to Fig. 12.11 is immediately evident. – In Fig. 12.13, the slit (gap) has been replaced by a rectangular obstacle of the same width. Here, the diffracted waves appear even

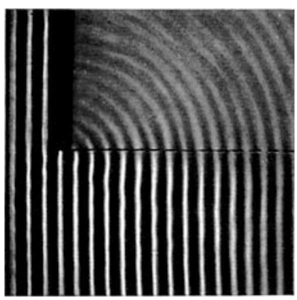


Figure 12.11 Blocking linear waves by a semi-plane. This and the following Figs. 12.12–12.20 and 12.22–12.24 are instantaneous images ('snapshots', with ca. 0.002 s exposure time). **(Video 12.2)**

Video 12.2:
"Experiments with water waves"
http://tiny.cc/tfgvjy
Diffraction (1:40). Behind a semi-plane, the waves propagate within the optical shadow region.

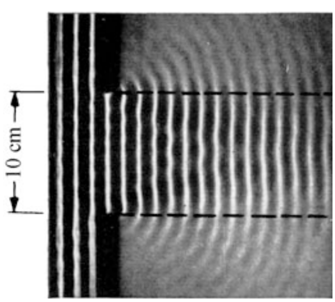
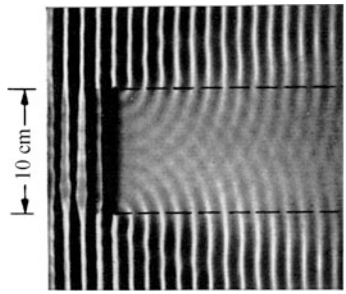


Figure 12.12 "Cutting out" linear waves by a slit

Figure 12.13 The shadow of an obstacle produced by linear waves

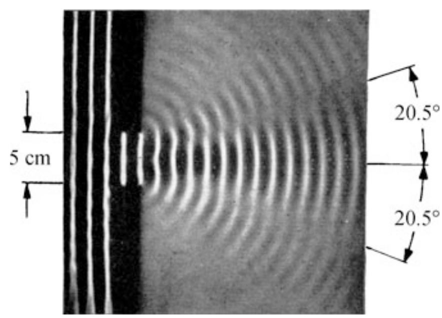


more clearly: The diffracted waves coming from the upper and lower edges of the obstacle interfere with each other, and the "shadow" of the obstacle becomes increasingly washed out at greater distances. Waves continually propagate along the axis of the shadow![C12.5]

C12.5. In optics, this is called a "POISSON spot" (Vol. 2, Sect. 21.1).

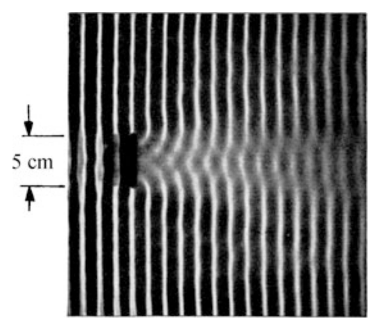
Figures 12.14 and 12.15 show the corresponding experiments for smaller dimensions. The width of the slit and of the obstacle is now only about 3λ. In this case, the geometric ray-tracing construction fails even as a rough approximation. Behind the slit, the wave train fans out broadly, and in the diffraction region, clear-cut maxima and minima can be seen on both sides. Behind the obstacle, the shadow is only indistinctly visible even at short distances; the diffracted waves are hardly weaker than on both sides in the free wave field.

Figure 12.14 Cutting off linear waves by a slit; secondary maxima in the diffraction region **(Video 12.2)**



Video 12.2: "Experiments with water waves"
http://tiny.cc/tfgvjy
Diffraction by a slit (2:00).

Figure 12.15 The very incomplete formation of a shadow behind a small obstacle in a linear wave field



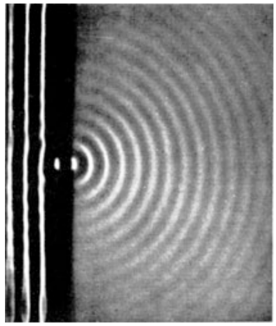


Figure 12.16 Elementary waves behind a very small opening on which linear waves are incident

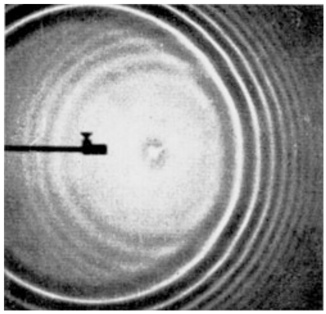


Figure 12.17 Elementary waves which are produced by scattering from a small obstacle **(Video 12.2)**

In Fig. 12.16, the width of the slit is now *smaller* than the wavelength of the waves; the incident waves which pass through it form essentially semicircular wavefronts behind the slit. – In Fig. 12.17, an obstacle of the same width was used. The incident waves take so little notice of this obstacle that we can barely see its effects against the background of continuous waves. Therefore, in Fig. 12.17 we let a wave train of limited length (produced by a brief, pulsed excitation of the wave source) pass by this small obstacle. In the stopped-motion image, this short wave train has already passed the obstacle. We see the result: The obstacle has produced a new, circular wave. – The observations in Figs. 12.16 and 12.17 can be summarized as follows: The slit in Fig. 12.16 and the obstacle in Fig. 12.17 have become *source points* for new waves. Behind the slit, these propagate as semicircular wave fronts, and around the obstacle, they are complete circular waves. Both are limiting cases of diffraction. In this limiting case, they are termed "scattered" waves, or "waves produced by scattering". The name *elementary waves* is also often used. Their amplitude decreases with decreasing width of the slit or obstacle. However, they are present even at arbitrarily small geometric dimensions. When the amplitude of the incident waves is sufficiently large, they can always be detected. Even the smallest objects make their existence known through scattering of waves ("ultramicroscopic detection").

Video 12.2:
"Experiments with water waves"
http://tiny.cc/tfgvjy
Note here the short wave train which was produced by dipping the excitation bar only once into the water surface (2:30). According to FOURIER, such a pulse consists of waves from a broad frequency spectrum (see Fig. 11.16). The waves of shorter wavelength travel faster (compare Fig. 12.74; capillary waves), and thus a characteristic wave field is formed in which the shorter waves move ahead of the longer ones ("spectral analysis by dispersion"). This phenomenon can be readily observed when swimming in a calm lake!

The results of our experiments thus far are: We can describe the propagation of waves with geometric boundaries by using simple *rays* or *bundles of rays*. However, this works only when the geometric dimensions B (width of the slit or obstacle) are large compared to the available wavelength λ ("geometric optics").

12.7 Reflection and Refraction

In Fig. 12.18, divergent waves are incident on a smooth, planar obstacle at an angle. The imperfections in its surface (scratches, bumps) are small compared to their wavelengths. The waves are reflected as if by a mirror (specular reflection). The rays which have been reversed by the mirror are sketched in. They are perpendicular to the circles representing the incident and the reflected wave fronts. For the rays drawn, the law of reflection holds: The angle of incidence equals the angle of reflection. Above the mirror, we see the superposition and interference of the incident and the reflected waves. To the right, we can discern the shadow of the mirror with its edges washed out owing to diffraction.

In shallow water, waves propagate more slowly than in deep water (see Eq. (12.36). This fact can be used to demonstrate *refraction*. In Fig. 12.19, the line *0-0* in a ripple tank separates a shallow-water region B (below) from a deep-water region A (above). Linear waves incident from the upper left at an angle α pass over this boundary. An incident and a refracted ray are sketched in, also the "normal axis" NN (perpendicular to the boundary). The transition of the waves from region A to region B is governed by the *law of refraction*:

$$\frac{\sin\alpha}{\sin\beta} = \text{const} = n_{A\to B}\,. \tag{12.13}$$

It defines the index of refraction $n_{A\to B}$ for the transition from $A \to B$[C12.6]. Using it, we find for the wavelengths

$$\lambda_B = \frac{\lambda_A}{n_{A\to B}} \tag{12.14}$$

and for the velocities of the waves:

$$c_B = \frac{c_A}{n_{A\to B}}\,. \tag{12.15}$$

C12.6. This is the "relative" index of refraction. We can write it as the ratio of two indices of refraction, each of which applies to one particular region or transparent medium: $n_{A\to B} = n_A/n_B$. If we take a fixed value for one of the regions, then the value of the index of refraction is defined for all other regions. This is done in optics by setting the index of refraction equal to 1 for the "medium" *vacuum* (see also Sect. 12.11 and, in Vol. 2, Sect. 16.3.)

Video 12.2:
"Experiments with water waves"
http://tiny.cc/tfgvjy
Reflection of plane waves by a wall at an incidence angle of 45° (1:15).

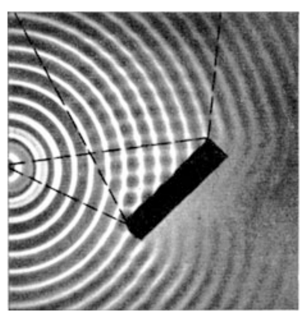


Figure 12.18 Reflection of divergent waves by a mirror at an angle of incidence of 45°. The smooth surface of the mirror appears to be distorted due to the ripples on the water (**Video 12.2**).

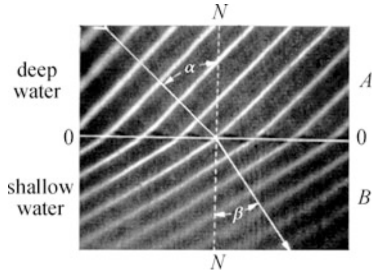


Figure 12.19 Refraction of linear waves on passing into a region of lower wave velocity: snapshot. (The waves which are reflected at the boundary *0-0* upwards and to the right are not visible; they are too weak.) (**Video 12.2**).

Video 12.2: "Experiments with water waves"
http://tiny.cc/tfgvjy
Refraction (4:20). In the video, the depth of the water in region A is 8 mm, and in B, it is 2 mm. The incident waves are refracted just like light waves which are incident on a glass plate (towards the normal to the boundary), since their phase velocity in region B is smaller than in A.

12.8 Image Formation

Figure 12.20 shows a "shallow-water lens": We place a flat, transparent, lens-shaped object into the ripple tank. Between its upper surface and that of the water, the remaining depth is only about 2 mm. This "lens" is mounted on both sides in a frame which blocks the waves. The divergent waves are delayed most when they pass through the thick center portion of the lens, and the least at its edges, corresponding to the decreasing thickness of the lens there. As a result of this delay, the curvature of the waves changes its sign. The waves converge behind the lens towards the "image point" and again diverge after passing through it. – Reflection from concave mirrors can show the same effect. We illustrate this in Fig. 12.21. This time, the wave source (the "object point") is on the left at "infinity", i.e. we use linear waves. In this case, the image point is termed the "focal point". Figures 12.20 and 12.21 are quite instructive: *Image points are in reality not the convergence points of the rays, but rather extended diffraction patterns of the lens or mirror mount*. Their diameter depends on the wavelength of the waves in question and on the diameter of the lens or mirror. The greater this diameter, the smaller the diffraction pattern, the true image point.

To represent a wave, a single line is often sufficient, namely the normal to the wavefront; it is also called the *principal ray*. Referring to this popular method of representation, one frequently calls a wave field which is bounded by parallel lines (i.e. a wave bundle) a *beam*. One thus speaks of e.g. *beams of sound* and of *light beams*.

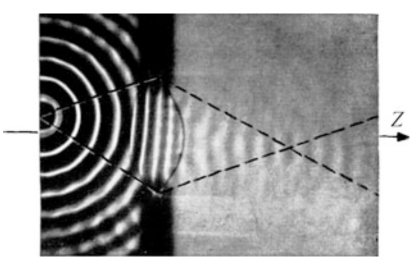


Figure 12.20 A wave source point as object point is imaged by a shallow-water lens at an "image point". The object point is intentionally not on the symmetry axis Z (**Video 12.2**)

Video 12.2: "Experiments with water waves"
http://tiny.cc/tfgvjy
In the video, the image point of an "infinitely" distant object is shown (i.e. linear wavefronts) (4:50).

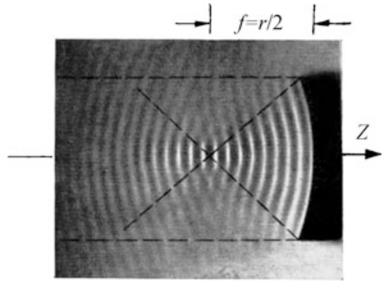


Figure 12.21 Prominence of the focal point in the interference field of short surface waves on water in front of a cylindrical concave mirror (radius of curvature r; f is the *focal length*, see also Vol. 2, Sects. 16.7 and 18.2). Time exposure, mirror diameter 14 cm. Compare Fig. 12.48

12.9 Total Reflection

Figure 12.22 shows refraction for the case that the waves are incident from below right upon the boundary surface. In this case, the angle of incidence α is *smaller* than the angle of refraction β. Experimentally, we find

$$\frac{\sin\alpha}{\sin\beta} = \text{const} = n_{B\to A} = \frac{1}{n_{A\to B}} < 1\,. \tag{12.16}$$

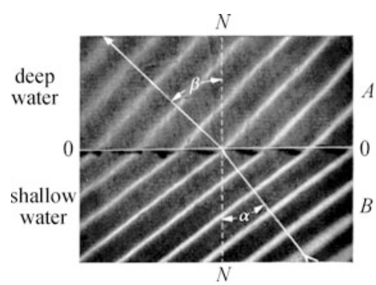
α has a maximum value of φ_{T}, defined by $\sin\beta = 1$ ($\beta = 90°$), that is

$$\sin\varphi_{\mathrm{T}} = n_{B\to A} = \frac{1}{n_{A\to B}} \tag{12.17}$$

cannot be greater than one. φ_{T} is called the *critical angle for total reflection*. At angles of incidence $\alpha > \varphi_{\mathrm{T}}$, no refracted wave can enter into the region A where the wave velocity is greater; instead, the incident waves undergo *total reflection*.

According to this formal geometrical consideration, when the critical angle φ_{T} for total reflection is exceeded, no waves at all should penetrate into the region A with its higher wave velocity (due to the fact that at an angle of incidence of φ_{T}, they are already refracted by 90°, i.e. they are propagating parallel to the boundary *0-0*); cf. Eq. (12.17).

Figure 12.22 The refraction of linear waves on passing into a region with higher wave velocity



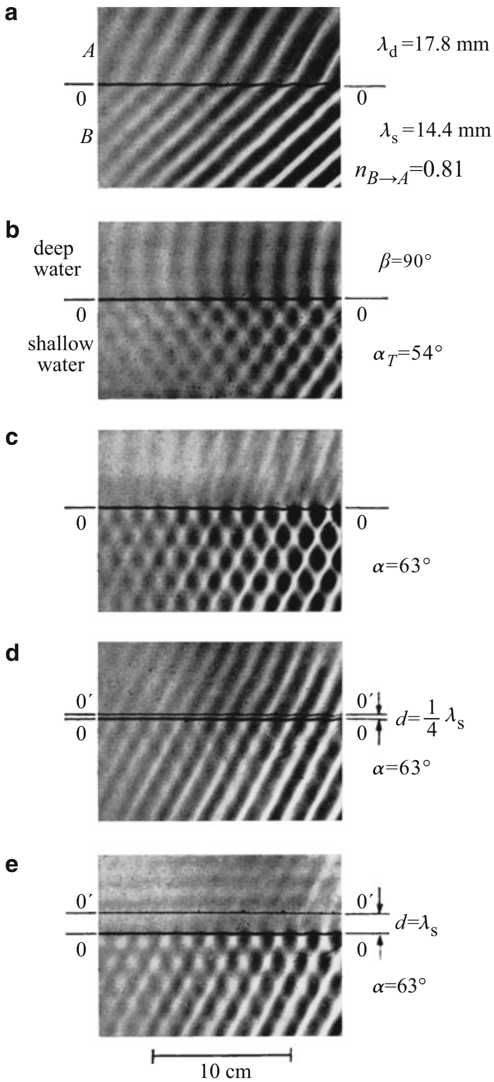


Figure 12.23 Demonstration of the total reflection of water waves. The wave propagates from below right onto the boundary *0-0*. During total reflection (images **b** und **c**), below *0-0*, sinusoidally modulated waves travel from right to left, i.e. the waves are split up by horizontal interference minima **(Video 12.2)**.

We now should consider what *in fact* happens in total reflection. To this end, we use the same setup as in Fig. 12.22; we again allow a wave *at the angle of incidence* α (here around 40°) to arrive at the boundary *from below right*, above which the region A with a higher wave velocity begins. In Fig. 12.23a, we see both refraction and reflection. The amplitudes of the reflected waves are much smaller

Video 12.2:
"Experiments with water waves"
http://tiny.cc/tfgvjy
Total reflection (5:30). The waves are produced in the shallow-water region (water depth 2 mm), and are incident on the boundary line to the deep-water region (8 mm) at an angle which is greater than the critical angle for total reflection. Besides the reflected wave, we see a transversally-damped wave, which travels along the boundary. Its depth of penetration into the deep-water region is of the order of one wavelength. When the width of the deep-water region (channel) is reduced (5:30), one can also observe a wave on the other side of the channel; this is the "tunnel effect" (see the end of this section and Vol. 2, Sect. 25.9).

than those of the incident waves. In this example, we have for the index of refraction $n_{B \to A} = \lambda_s / \lambda_d = 14.4\,\text{mm}/17.8\,\text{mm} = 0.81$. At $\sin\alpha = 0.81$ or $\alpha = 54°$, total reflection begins. In Fig. 12.23b, $\beta = 90°$. The refracted waves are perpendicular at the boundary and extend upwards as curved "diffracted" waves. The amplitude of the reflected waves is now comparable to that of the incident waves.

In Fig. 12.23c, the angle of incidence α has been further increased to 63°. We are thus in the middle of the angular range for total reflection, and there we observe the following effects:

Some waves are still propagating above the boundary (they are called *evanescent waves*). In the figure, the white wave crests are overstepping the boundary *0-0* by about 1 mm. Their propagation direction is parallel to the boundary. The amplitude of these waves decays very quickly in an upwards direction, i.e. perpendicular to their propagation direction. Thus, *these waves are strongly damped in a direction perpendicular to their direction of propagation.*

(Their continuation as curved, diffracted waves is very clearly seen. They can even distract our attention as a disturbance from the essentials of the figure. But diffraction also belongs inseparably to any bundling of waves into a beam, etc.)

The transversally-damped waves in the second region A, which according to formal-geometric considerations should be free of waves, are in fact necessary *for the appearance of total reflection*[C12.7]. The next two experiments demonstrate this. In Fig. 12.23d, the deep-water region above the boundary *0-0* has been narrowed down to a small strip or channel. Above *0′-0′*, a region of shallow water again follows it. The distance *0-0′* is equal to a quarter of a wavelength. The deep-water region is therefore narrower than the width of the transversally-damped waves in Fig. 12.23c. The result: The reflection is no longer total; some waves clearly propagate upwards across the boundary *0-0*.

And finally, the counter-experiment: In Fig. 12.23e, the distance *0-0′* has been extended up to one wavelength. The deep-water channel now offers sufficient space for the transversally-damped waves to form. This also restores the total reflection.

Summary: Total reflection can occur only when the width of the region with a smaller index of refraction (in our example *0-0′*) is not small relative to the wavelength. Otherwise, the region of smaller index of refraction does not represent an impenetrable obstacle to the waves. They are able to pass through the 'forbidden' region, although weakened, as though a path had opened through a tunnel: this is the *tunnel effect*[C12.8].

C12.7. This can be heuristically justified by the following argument: How do the incident waves "know" that the region *A* is a "forbidden zone" if they cannot penetrate it at all? Some (small) portion of the incident wave must penetrate to a limited distance (ca. one wavelength) into the region *A* as an "evanescent wave" in order to impart this information, so that the major portion of the incident wave will be ("totally") reflected at the boundary.

C12.8. This is the name of the effect in wave mechanics. A matter wave, e.g. an electron wave, can also penetrate as a damped (evanescent) wave into a forbidden region, i.e. for example into a region where its potential energy would be greater than its quantum energy. This can also occur at normal incidence. If the width *d* of the forbidden region is chosen to be sufficiently small, the matter wave can pass through the region and can be observed on its other side, in analogy to Fig. 12.23d; that is, the particles corresponding to the matter wave can "tunnel through" the potential barrier with a certain probability.

12.10 Shockwaves when the Wave Velocity Is Exceeded

Suppose that an object dips into the water surface in the ripple tank, and moves horizontally at a constant velocity u. Let this velocity be greater than the phase velocity c for propagation of surface waves on the water. Then a *shockwave* is formed, similar to the bow waves from a ship with which we are all familiar. Spatially, it corresponds to a conical wave, as produced for example in Fig. 12.87 by a bullet.

Such shockwaves can also be readily produced as a periodic sequence. Figure 12.24 shows a suitable setup. It contains a channel with boundaries *0-0* and *0′-0′*, bordered on both sides by regions of shallow water. Beyond the left margin of the picture, an oscillating excitation source produces periodic waves in the channel: Each wave crest which propagates along the channel acts like a moving object, and it produces a shockwave in the shallow-water regions on either side of the channel.

The same phenomenon can be described in quite a different manner, namely as a limiting case of refraction at an angle of incidence of $\alpha = 90°$. This is illustrated in Fig. 12.25, for the lower boundary *0-0*, to facilitate the comparison with Fig. 12.19[2]. The location of the normal *N-N* to the boundary was arbitrarily chosen, since the incident ray is propagating parallel to the boundary and not, as in Fig. 12.19,

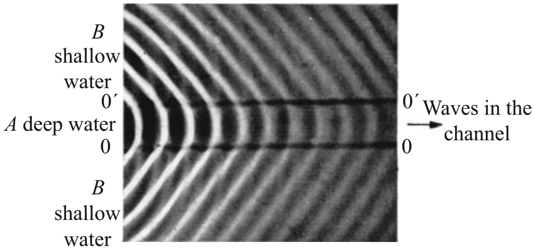
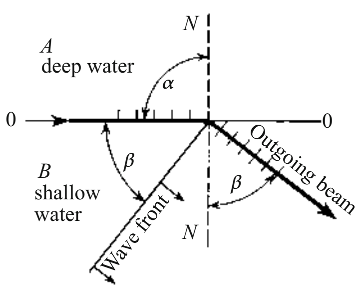


Figure 12.24 Production of a periodic sequence of shockwaves

Figure 12.25 A periodic sequence of shockwaves can be treated as a limiting case of refraction (a representation which is popular in geophysics)



[2] In Fig. 12.19, the wave was incident from the upper left at the angle α.

at an angle to it. From Eq. (12.13), it follows for $\alpha = 90°$ that the *angle of refraction* is given by $\sin\beta = 1/n_{A\to B}$. β is thus the same as when the direction of the rays is reversed (Fig. 12.22); it is then the *angle of incidence* φ_T, the critical angle for total reflection. More details will be given in Sect. 12.11.

12.11 HUYGHENS' Principle

An explanation for refraction and reflection can be provided by HUYGHENS' principle. In Fig. 12.26, let *0-0* be a reflective boundary surface; *I* is a wave crest of a linear wave incident from the upper left. It strikes the points which are arbitrarily marked on the boundary surface *one after another*. Each point can then be considered to be the source of an elementary wave of the type that we encountered in Sect. 12.6. These elementary waves are indicated by short circular arcs in the figure. Their tangent is a wave crest *II* (wave front) of the reflected wave. One of the paths which lead from the crest *I* to the crest *II* is shown as a dashed line; all such paths are traversed in the same times.

Figure 12.27 and its caption explain *refraction* at a boundary which separates two regions where the wave velocity *c* is different, in a corresponding manner.

Finally, we consider the limiting case of refraction which is treated in Fig. 12.25. To do so, we refer to Fig. 12.28. It shows the path of a single wave crest *T*-*T* (a wave front). This crest moves to the right at the velocity c_A. Its ends, which abut the walls of the channel *A*, are the source points of elementary waves which propagate as circular waves, however at the lower velocity c_B which is characteristic of

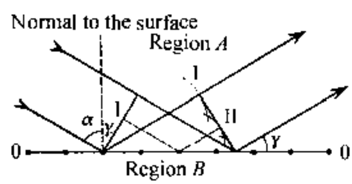


Figure 12.26 The origin of specular reflection from a plane, according to HUYGHENS' principle. The borders of the wave beam are represented as rays.

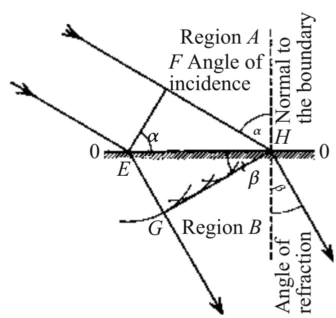


Figure 12.27 The origin of refraction according to HUYGHENS' principle. The ray paths *FH* and *EG* are traversed in the same times. Their ratio is the same as the ratio of the velocities of the waves in the two regions, i.e. $FH/EG = c_A/c_B = \sin\alpha/\sin\beta = \text{const} = n_{A\to B}$.

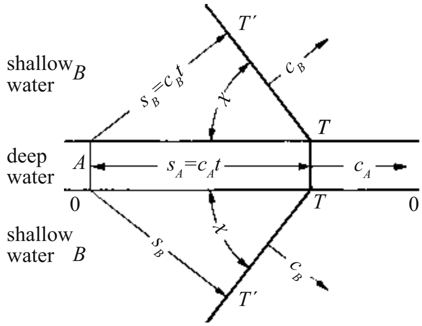


Figure 12.28 The origin of "MACH's angle" χ

the shallow-water regions B. The common tangent of all these elementary waves represents the new linear wave crest T-T'. From the sketch, one can read off the relation

$$\sin\chi = \frac{c_B}{c_A}\,, \tag{12.18}$$

and here, the angle χ is called *Mach's angle*.

Applying this result to the question raised at the beginning of Sect. 12.10 (conical shockwaves), we obtain for MACH's angle

$$\sin\chi = \frac{\text{Phase velocity } c \text{ of the waves}}{\text{Velocity } u \text{ of the object}}\,.$$

12.12 Model Experiments on Wave Propagation

In Figs. 12.26 and 12.27, neither the transverse width of the incident wave nor a structure of the boundary *0-0* was taken into account. If this simplification is not permissible, then the common tangent (*II* in Fig. 12.26, *G-H* in Fig. 12.27) of the elementary waves is no longer sufficient. One then has to consider the interference due to superposition of the elementary waves. This interference can be most clearly illustrated by *model experiments*. Initially, we will deal with the case shown in Fig. 12.12 in this way; that is, 'cutting out' linear waves by passing them through a *broad* slit.

In Fig. 12.29, the double arrows indicate a wave crest which has arrived at the opening, and its length also shows the width B of the opening. Furthermore, the system of concentric circles indicates a single elementary wave train, which originates from a single point within the opening. – This wave pattern can be transferred to a glass slide and projected onto a screen; the double arrow is drawn on the screen. Then we imagine that using additional projectors, a continuous series of similar glass slides are projected side by side on the

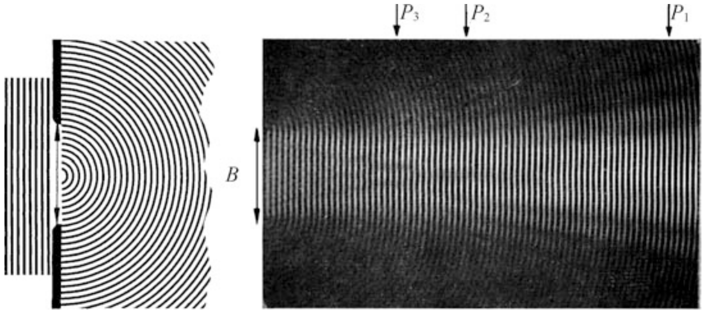
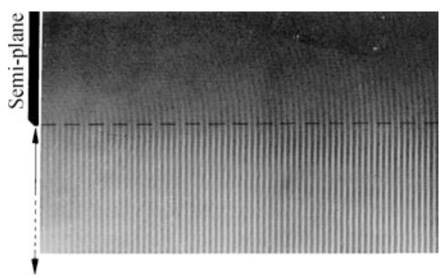


Figure 12.29 Cutting out a portion of a linear wave train by using a broad slit (FRESNEL's case of observation). At the left, the wave pattern has been transferred to a glass plate. The profile of the waves has been chosen to be not sinusoidal, but rather rectangular, to avoid losing detail in the printing process. Look at the right-hand image (and later also Fig. 12.31) along its long axis; the arrows indicate points P_1, P_2 and P_3 which one can imagine to lie on the symmetry axis of the wave field. These points will be used in Sect. 12.14 as space points for the *zone construction*.

screen. In practice, one uses a more clever arrangement: We use only the *one* glass-slide image from Fig. 12.29 (left) and quickly move its wave center up and down in the direction of the double arrow, using some sort of a mechanical setup. Neither the eye nor photographic film can separate these images which follow each other rapidly in space and in time; they register only the superposition of all the elementary wave trains. This produces the wave pattern that is reproduced in Fig. 12.29, right. It shows the structure of the wave field more clearly than the earlier Fig. 12.12. Along the beam axis, the waves are initially interrupted by nearly wave-free stripes, as indicated by the arrows P. The wave field takes on a simple form only when the distance from the slit becomes large compared to its width, for example to the right of the arrow P_1.

In the model experiment of Fig. 12.29, if we leave the upper edge of the slit as is, but move the lower edge a considerable distance downwards, then we arrive at diffraction by a semi-plane (Fig. 12.30). This corresponds to Fig. 12.11.

Figure 12.30 Model experiment for diffraction by a semi-plane



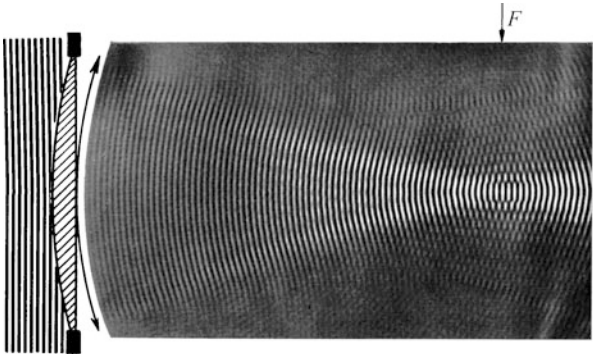


Figure 12.31 A model experiment demonstrating FRAUNHOFER diffraction by a broad slit, and showing the formation of an "image point", here a "focal point" F. In its neighborhood, the waves are linear. Compare the neighborhood of the focal points in Figs. 12.21 and 12.48.

In Fig. 12.29, the wave train after passing the slit is divergent. This case is called *Fresnel diffraction*. By adding a converging lens, one can convert the divergent waves resulting from the diffraction into convergent waves. Then one refers for short to *Fraunhofer diffraction*:

The waves propagate more slowly within the lens than in its surrounding medium. As a result, the center of the beam is delayed relative to its outer parts. The wave front becomes concave; thus, the *straight* double arrow in Fig. 12.29 should be replaced by a *circular arc*. Everything else proceeds just as before. We move the wave source point (using some sort of mechanical setup) back and forth along the curved double arrow. The result is shown as a photograph in Fig. 12.31.

FRAUNHOFER's observation method yields a diffraction pattern in the focal plane of a convergent lens which is as simple as that obtained only at very large distances from the slit in the FRESNEL case. For this reason, the FRAUNHOFER case is usually preferred.

Finally, we show two model experiments on FRESNEL diffraction from narrow slits (Fig. 12.32). Both of the diffraction patterns already exhibit a simple structure, even close to the slit; this is obtained for broad slits only at a considerable distance.

12.13 Quantitative Results for Diffraction by a Slit

To begin with, we use Fig. 12.33 to understand how the minima that can be seen on either side of the central diffraction maximum in Fig. 12.32 come about. To this end, we suppose that the observa-

Part II

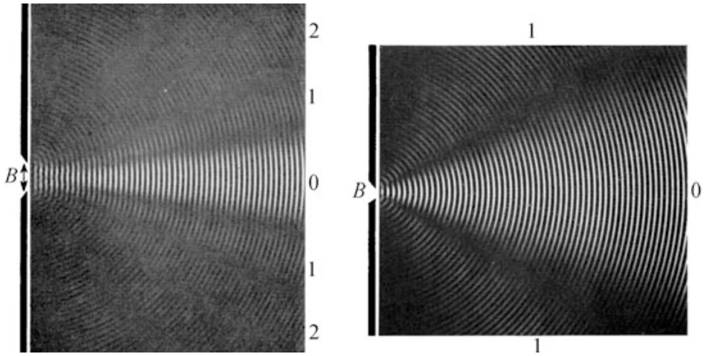


Figure 12.32 Two model experiments showing the passage of waves through narrow slits with different widths *B*. For the demonstration, we attach the glass plate with the half-waves (Fig. 12.29, left) to the end of a vibrating reed which can be made to vibrate back and forth by an electromagnetic drive like that of a doorbell.

tion point *P* is very far away from the slit, so that the two rays leading from the edges of the slit to *P* are practically parallel. Furthermore, we divide up the slit (of width *B*) into a number *N* of regions, for example $N = 12$, each of which is the same size, and number them as 1, 2, 3 etc. Each of these regions is considered to be the point of origin of an elementary wave with the same number. All *N* elementary waves meet and are superposed at the observation point *P*. Then the amplitudes of the elementary waves add to give the overall amplitude at the point *P*. The essential aspect of this addition is the different paths taken by the individual elementary waves.

Suppose that the maximum path difference *s*, between the first and the twelfth elementary wave, is equal to λ. Then the path difference between the first and the sixth, or between the second and the seventh elementary waves, etc., is in each case equal to $\lambda/2$. This means that the amplitudes of each of these pairs cancel each other; as a result, in the corresponding direction (angle α), there is no wave amplitude

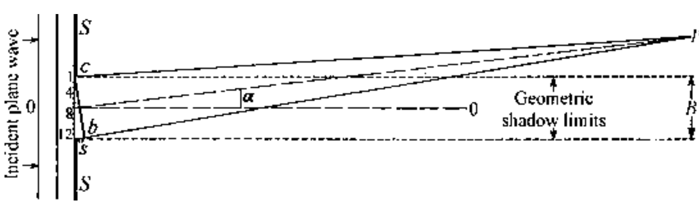


Figure 12.33 Calculating the diffraction pattern from a slit

at all at P, and we have a minimum; from Fig. 12.33, this angle α is given by

$$\sin\alpha = \frac{\lambda}{B}\,. \tag{12.19}$$

This equation agrees with the observations. It allows us to calculate λ if we know the slit width B and measure the direction of the first minimum. – Along other directions, we can carry out the addition of the elementary waves graphically. This yields the "amplitude landscape" or diffraction pattern which plays an important role for waves of all types (see Fig. 12.35). It gives the distribution of wave amplitudes for the various observation directions at some distance behind a slit of width B.

The path difference between any two of the N neighboring elementary waves is

$$\Delta\lambda = \frac{s}{N} = B\frac{\sin\alpha}{N}\,. \tag{12.20}$$

For the point P_0 on the slit's symmetry axis *0-0*, we find

$$s = 0, \quad \alpha = 0, \quad \sin\alpha = 0, \quad \Delta\lambda = 0\,.$$

Thus, all 12 amplitudes in Fig. 12.34 add here *without* a phase difference, as shown in the auxiliary figure *0*. Their sum is drawn in as a heavy arrow R_0 below the image and as the height R_0 in Fig. 12.35 above the point $\sin\alpha = 0$ on the abscissa.

For the next point P_1, we choose $s = \frac{\lambda}{3}$; then we have $\sin\alpha = \frac{\lambda}{3B}$, and the path difference of two neighboring elementary waves is $\Delta\lambda = \frac{1}{12}\,\frac{\lambda}{3}$, or, in angular units, $\Delta\varphi = \frac{1}{12}120° = 10°$.

The amplitudes of the 12 elementary waves add as shown in the auxiliary figure 1. As their sum, we obtain the arrow R_1. It is is entered in Fig. 12.35 as the result of the graphical addition above the point $\sin\alpha = \frac{\lambda}{3B}$ on the abscissa.

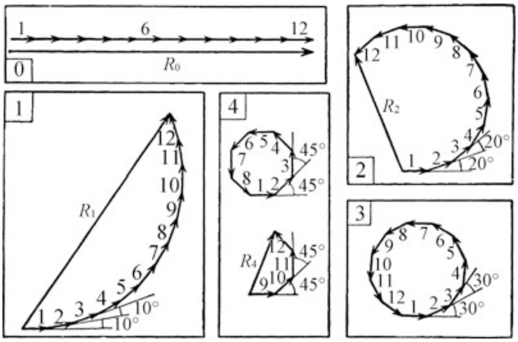


Figure 12.34 Auxiliary figures for the graphical construction of Fig. 12.35

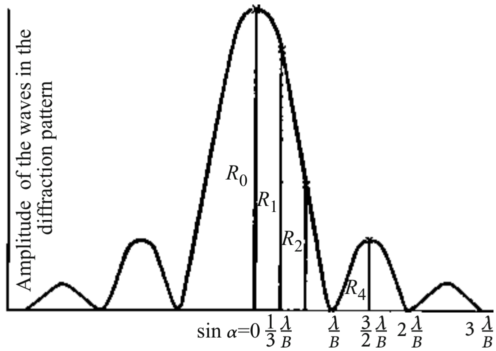


Figure 12.35 The "amplitude landscape" (diffraction pattern) resulting from the passage of a plane wave through a rectangular slit. Figure 12.34 contains the auxiliary figures needed for the construction. The radiation power (or energy current, i.e. the energy transported by the waves per unit time) of the waves is proportional to the squares of their amplitudes. Thus, for a comparison with measurements (e.g. in Fig. 12.62), the ordinate values of this diffraction pattern must be squared. Here, the signs of the amplitudes are not taken into account (see the note at the end of Sect. 12.3).

We continue in an analogous manner. For the point P_2, we choose

$$s = \frac{2}{3}\lambda, \quad \text{i.e. } \sin\alpha = \frac{2}{3}\frac{\lambda}{B}, \quad \Delta\lambda = \frac{1}{12}\frac{2}{3}\lambda, \quad \Delta\varphi = 20^\circ .$$

Auxiliary figure 2 yields the sum as the length of the arrow R_2.
For the next point, we choose

$$s = \lambda, \quad \text{i.e. } \sin\alpha = \frac{\lambda}{B}, \quad \Delta\lambda = \frac{\lambda}{12}, \quad \Delta\varphi = 30^\circ .$$

The amplitudes of the 12 elementary waves add in the auxiliary figure 3 to give a closed polygon; their sum is zero. Therefore, in Fig. 12.35 at the value $\sin\alpha = \lambda/B$ on the abscissa, we have put the corresponding point on the axis (ordinate = 0).
Finally, we take

$$s = \frac{3}{2}\lambda, \quad \text{i.e. } \sin\alpha = \frac{3}{2}\frac{\lambda}{B}, \quad \Delta\lambda = \frac{1}{12}\frac{3}{2}\lambda, \quad \Delta\varphi = 45^\circ .$$

The graphical addition is shown in auxiliary figure 4. The amplitudes of the first 8 elementary waves form a closed octagon; their sum is zero. The 9th to the 12th amplitudes yield a half-octagon and thus correspond to the arrow R_4.
For $s = 2\lambda$, or $\Delta\varphi = 60^\circ$, both the amplitudes of the elementary waves 1 to 6 and those of 7 to 12 give zero, so that the point at $\sin\alpha = 2\lambda/B$ on the abscissa in Fig. 12.35 again lies on the axis. This should be sufficient. We can readily complete Fig. 12.35, keeping in mind its symmetry on both sides of the center point (at $\sin\alpha = 0$).
In Figs. 12.14 and 12.33, we have treated the limiting case of *Fraunhofer diffraction*. The incident wave crests are practically straight lines. The points P in the observation plane are "infinitely" distant to the right from the slit, or they lie in the focal plane of a lens.

12.14 FRESNEL's Zone Construction

In Sect. 12.13, we treated a special case of a general procedure which we will now deal with; it is known as FRESNEL's zone construction. In Fig. 12.36, let S be the wave source point and P the "observation point" (or *"receiving point"*). Using P as their center, we draw a system of spherical waves with the wavelength of the radiation used (wave crests are black and wave troughs are white). Furthermore, centered on the source point S, we imagine a spherical surface of radius a. It intersects the wavefronts from P, defining ring-shaped, alternately white and black *zones*. Looking from the observation point P, we see a spherical surface sector containing a system of concentric rings, similar to that shown below in Fig. 12.38. For the radius r_m of the mth zone on the spherical sector, we find the simple geometric relation

$$r_\mathrm{m}^2 = m\lambda \frac{ab}{a+b} \tag{12.21}$$

(for the derivation, see Fig. 12.36).

The path of the waves via the mth zone is longer by $\Delta = m\lambda/2$ than along the line directly connecting the source point S and the observation point P, whose length is $(a + b)$. All the zones have approximately the same areas, namely

$$A = \pi(r_{\mathrm{m}+1}^2 - r_\mathrm{m}^2) = \pi\lambda \frac{ab}{a+b} \,. \tag{12.22}$$

Now, we add to Fig. 12.36 the *object*, either a circular hole in an opaque screen, or a circular disk; the double arrow indicates the diameter of this object. Then only a portion of the zones remains visible from the observation point P. Looking from P, we see the (spherically convex) zone areas as in Fig. 12.37. The number of "surviving"

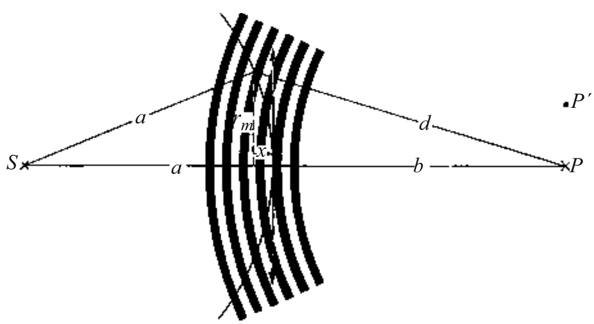


Figure 12.36 The FRESNEL zone construction. m is the index of the black and white zones, which are numbered sequentially. We have $r_\mathrm{m}^2 = a^2 - (a - x)^2$, $r_\mathrm{m}^2 = d^2 - (b + x)^2$, and $d = b + m\lambda/2$. From these three equations, we compute r_m^2 by neglecting small terms containing $\lambda^2/4$ and x^2.

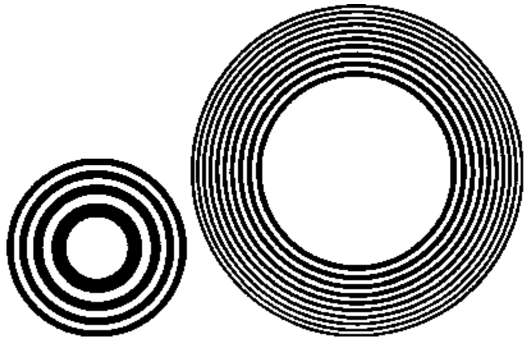


Figure 12.37 The zones which are not covered by an opaque screen with a circular opening (*left*), and by a circular opaque disk of the same size (*right*), reduced to two-thirds their size in Fig. 12.36. The right-hand image can be thought of as containing additional rings outwards, with decreasing line thicknesses.

zones changes depending on the distances a and b. Furthermore, we consider each of the remaining zones to be the source of new elementary waves. These interfere with each other. The resultant of all the elementary waves that arrive at P gives the amplitude at that point. Examples:

1. The number of zones allowed to pass by a *circular aperture* is *even*. Each pair of black and white zones practically cancels out (but not completely!). The observation point lies in a nearly wave-free region along the beam axis. This can be seen e.g. in Fig. 12.29 on the right at the observation point P_2. Here, the opening B allows only the two innermost zones to pass through (with $m = 1$ and $m = 2$); thus an *even* number. In the zone construction, we must keep in mind that in this case of an incident plane wave, the quantity a in Eq. (12.21) is very large ($\to \infty$).

2. The number of zones allowed to pass by a *circular aperture* is *odd*. The effect of the zone *left over* after forming pairs remains at full strength. The observation point lies along a region of the beam axis which contains waves. – This can be seen e.g. in Fig. 12.29, right, for the observation point P_3. Here, the aperture leaves the *three* innermost zones free (with $m = 1$, 2 and 3); thus an *odd* number.

3. An experimental investigation of the examples mentioned above, but with a point-like wave source (sound waves), will be described in Sect. 12.20, point 5 (and **Video 12.4**).

4. If we replace the circular aperture by a circular *obstacle*, then at the observation point, all the zones of higher index m combine. Whether there is one more or less is unimportant. The resultant of all the elementary waves at the observation point has practically always the same value; waves are always present there, e.g. on the center axis of the "shadow" in Figs. 12.13 and 12.15; cf. also Vol. 2, Sect. 21.1[C12.9].

C12.9. Understanding this fact is by no means trivial. A derivation can be found for example in P. Drude, *Lehrbuch der Optik*, Hirzel Publishers, 2nd ed. (1906), p. 155; or also in M. Born, *Optik*, Springer-Verlag, 3rd ed. (1933), reissued in 1972, pp. 144–147. English: *Principles of Optics*, Max Born and Emil Wolf, Pergamon Press, 4th ed. (1970), available at https://archive.org/details/PrinciplesOfOptics.

Figure 12.38 A zone plate for light waves (red filter light) and an observation point P at a distance of 2.7 m (actual size). The plate acts as a lens with more than one focal length. The longest is $f = 2.7$ m; about ten of the shorter ones can be readily observed.

5. The zone construction can also be applied to observation points that are not on the axis of symmetry. Imagine the zone surface to be mounted on an arm which can be swung around ($a + b$ in Fig. 12.36). Its pivot point is at the source point of the waves, and its free end at the observation point. Then a sideways motion of the observation point from P to P' moves the whole zone surface at once; thus the zones that are free to pass *through* the aperture or *alongside* the disk (fixed double arrow in Fig. 12.36!) are different from before. The resultant of all their elementary waves yields the maxima and minima outside the center of the image.

6. At large values of a, the zone surface becomes nearly planar. Then the zone image of a circular aperture can be transferred to a glass plate (e.g. photographically) without serious errors. The black rings are made opaque and the white rings transparent. Such a *zone plate* can be used to form images. As an example, in Fig. 12.38, a zone plate for light waves ($\lambda = 0.6\,\mu$m) and an observation point at a distance of 2.7 m are shown in original size. Their focal length b is given by Eq. (12.21). Additional focal points are to be found at $b/3$, $b/5$, etc. (see also Vol. 2, Sect. 21.8).

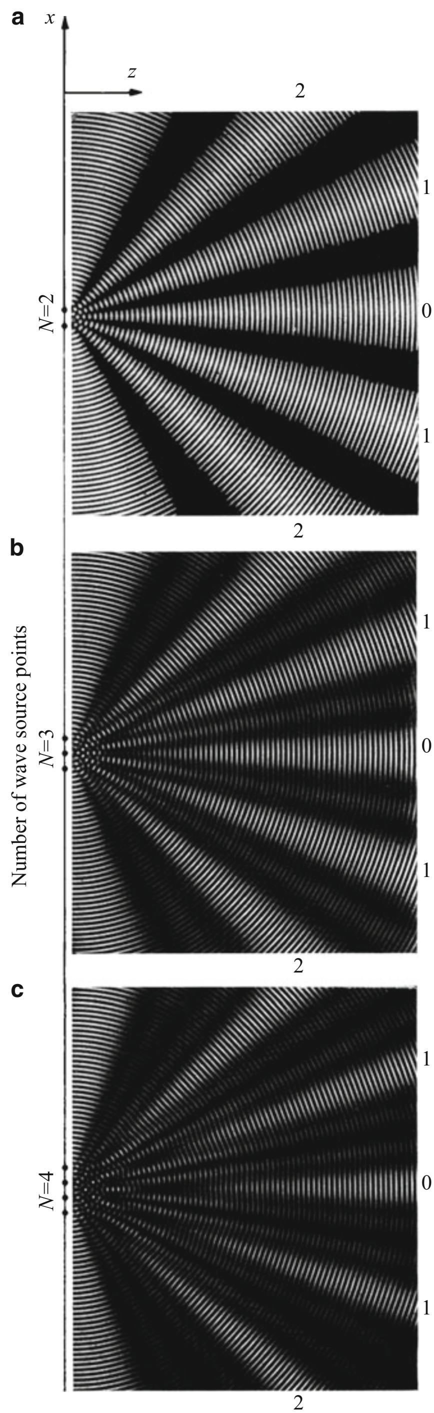
12.15 Narrowing of the Interference Fringes by a Lattice Arrangement of the Wave Sources

Figure 12.7 shows the experiment that Thomas Young used to demonstrate interference. In Fig. 12.39a, we repeat it as a model experiment by means of the superposition of two transparent wave images. This time, we do not place the two wave sources next to each other, but rather one above the other. A continuation of this model experiment makes use of a lattice of N wave source points, arranged equidistantly along a line. In Fig. 12.39b, there are three sources; in Fig. 12.39c, there are four, and so on. – This model experiment clearly demonstrates *two fundamental facts which are basic to all interference phenomena*[3]:

1. With an increasing number N of wave source points, the maxima seen already with two interfering wave trains remain, but they are

[3] Both of them can also be derived graphically using the same scheme as in Sect. 12.13.

Figure 12.39 A model experiment to show the interference of two, three and four wave trains from equidistant sources (marked with points at left). Two, three or four glass-plate images (cf. Fig. 12.29) are projected one above the other. The numbers refer to the order indices *m*.



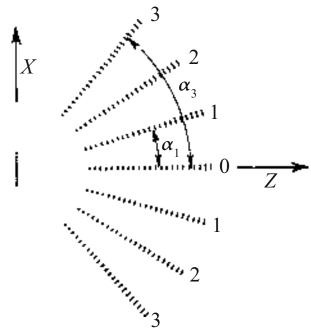


Figure 12.40 The interference maxima from a linear lattice; here as a schematic illustration. The numbers refer to the order indices *m*.

compressed into a more closely-spaced angular region: The interference fringes are *narrowed*.

2. Between each two neighboring maxima, $(N-2)$ *sub-maxima* appear, i.e. one in Fig. 12.39b, two in Fig. 12.39c, etc. At large values of N, the sub-maxima form a nearly continuous background. The resulting scheme is sketched in Fig. 12.40.

These facts, found here with the help of model experiments, play an important role in the precise measurement of wavelengths, especially in all the spectral regions of optics. We will therefore treat them in some detail experimentally in Sect. 12.20 (points 6–8).

We start by using narrow slits, spaced equidistantly, as wave source points, and we arrange for the waves to be incident in the z direction, as in Fig. 12.40. The waves which emerge from the slits are spread out over a large angular range due to diffraction (cf. Fig. 12.32, right), so that they superpose nearly as well as elementary waves and exhibit interference. This experimental trick, basically a minor point, has led to the terms *diffraction grating* or *optical lattice*.

For the angular dependence of the maxima of mth order, i.e. the interference maxima with a path difference of $\Delta = m\lambda$, when the waves are incident perpendicular to the lattice plane, we find

$$\sin\alpha_{\mathrm{m}} = \frac{m\lambda}{D} \qquad (12.6)$$

(m is the order index, and D is the distance between neighboring wave source points, called the lattice constant).

As a second possibility, we will use the mirror image of a first source as a second source, as described in the following: Imagine that in Fig. 12.39, the wave source is one point along the x axis and the second, third, fourth ... wave source points are its mirror images. Thus, in Fig. 12.41a, one wave source S is replaced by two mirror images S' and S'' produced by two levels of transmitting and reflecting planes. The waves reach the receiver along two paths which subtend the small angle 2δ. Their path difference Δ is shown in Fig. 12.41a.

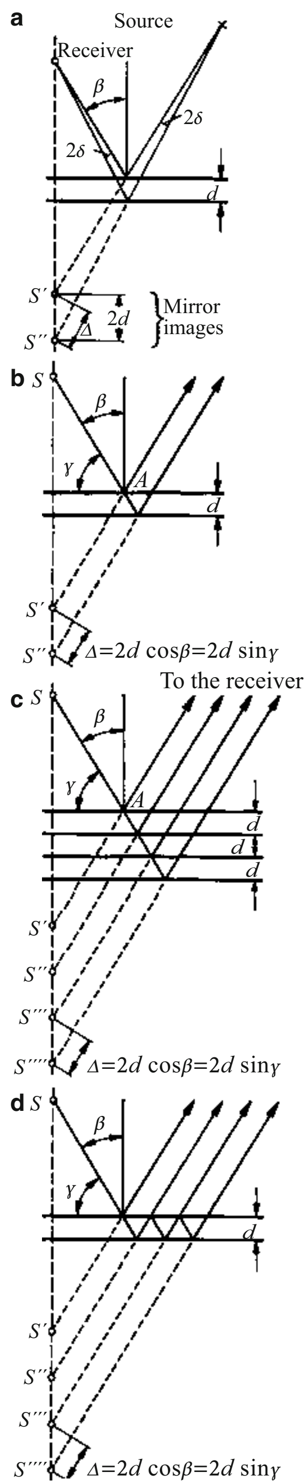


Figure 12.41 Interference arrangements in which mirror images S', S'', S''', ... of one source serve as wave source points. **a** POHL (divergent light beams); **b** HAIDINGER; **c** BRAGG; **d** PEROT and FABRY (**b**–**d** plane-parallel light beams). A demonstration follows in Sect. 12.20, Point 8.

In Fig. 12.41b, the receiver has been placed at a great distance and therefore the angle 2δ is practically zero. The waves reach the two reflecting surfaces along the same path. Maxima of the reflected waves, or minima of the transmitted waves, occur when the path difference obeys

$$\varDelta = 2d\cos\beta = 2d\sin\gamma = m\lambda \tag{12.23}$$

(m is an integer, γ is often called the Bragg angle).

In Fig. 12.41c, four transparent, reflecting surfaces are placed one behind the other at equal spacings d. In Fig. 12.41d, between two strongly reflecting but still somewhat transparent surfaces, multiple reflections occur. In both cases, the number N of wave source points is increased (S', S'', S''', ...) and thus the condition is fulfilled which leads to narrowing of the interference maxima. If we for example rotate the stacked transparent plates in Fig. 12.41c around an axis perpendicular to the plane of the page at the point A, then *sharp*, intense maxima appear in the directions of the arrows one after another, separated by *broad*, flat minima.

12.16 Interference of Wave Trains of Limited Length

Up to now, in treating the propagation of waves, we have tacitly made two assumptions: 1. The wave trains are excited with constant amplitudes and have unlimited lengths; and 2. the wave sources are point-like, i.e. the diameter of a wave source is presumed to be very small compared to the wavelength of the waves it emits. – If these two conditions are not fulfilled, then special features occur. They are particularly important for light waves. It is therefore expedient to discuss these effects in the section on *optics* (Vol. 2, Chap. 20).

12.17 The Production of Longitudinal Waves and Their Velocities

The knowledge that we have gained in this chapter will now be applied to the propagation of *3-dimensional* waves. For this purpose, *high-frequency longitudinal waves in air* are very well suited; these are short-wavelength sound waves.

First, some remarks about the formation of longitudinal waves. – A state can propagate as a wave if and only if it travels at a finite velocity. For the transverse surface waves on water, we have treated this fact for the time being as experimentally given. Its detailed discussion will follow in Sect. 12.21. – Longitudinal waves are formed

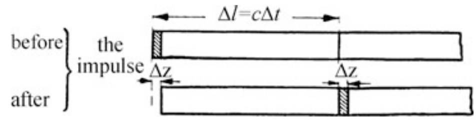


Figure 12.42 The calculation of the longitudinal sound velocity in a rod

when elastic disturbances propagate at finite velocities. In this case, we begin immediately with a quantitative treatment.

We let an impulse $F\Delta t$ act on the rod in Fig. 12.42 during the time Δt with the force F. It produces an elastic disturbance. This disturbance propagates to the right with the velocity c, so that within the time Δt, it affects a segment of the rod of length $\Delta l = c\Delta t$. This segment has a mass of

$$\Delta m = c\Delta t \cdot A\varrho \tag{12.24}$$

(ϱ is the density, A is the cross-sectional area of the rod).

The impulse $F\Delta t$ has two effects on the segment of the rod: First, it compresses the segment by the small amount Δz, upper image in Fig. 12.42. According to HOOKE's law, this is

$$\Delta z = \frac{1}{E}\Delta l \frac{F}{A} \tag{12.25}$$

(E is Young's modulus for the material of the rod; see Sect. 8.3).

Second, it gives the segment a momentum directed to the right:

$$\Delta m \frac{\Delta z}{\Delta t} = F\Delta t\,. \tag{12.26}$$

The segment of length Δl thus moves in the time Δt by Δz to the right; see the lower image. In the position shown there, the corresponding process repeats itself within the next time and length interval.

Combining Eqns. (12.24) to (12.26) yields

$$c\Delta t A\varrho \frac{\Delta z}{\Delta t} = \frac{\Delta z}{\Delta l} AE\Delta t\,, \tag{12.27}$$

and it follows from this that the velocity with which the longitudinal elastic disturbance moves, $c = \Delta l/\Delta t$ (usually called the sound velocity)[C12.10], is given by

$$c = \sqrt{\frac{E}{\varrho}}\,. \tag{12.28}$$

C12.10. Equation (12.28) holds only for *thin* rods. With other geometries, one cannot neglect the forces which occur due to the changes in the cross section of the rod (determined by the Poisson ratio μ, Sect. 8.3). Thus, the longitudinal sound velocity in an unbounded medium is
$$c_l = \sqrt{\frac{E}{\varrho}\frac{1-\mu}{(1+\mu)(1-2\mu)}}\,.$$
For transverse waves, this complication does not apply. In an unbounded medium, we have
$$c_t = \sqrt{\frac{G}{\varrho}}\,,$$
equal to the velocity of a torsional wave in a rod (corresponding to Eq. (11.6)), which one can represent as the superposition of two linearly-polarized transverse waves with a phase shift of 90°.
A *numerical example*: In an unbounded steel body ($\mu = 0.27$, Table 8.1), the longitudinal sound velocity is $c_l = 5.6\,\text{km/s}$.

Numerical example
For steel, $E = 2.0 \cdot 10^{11}\,\mathrm{N/m^2}$, $\varrho = 7850\,\mathrm{kg/m^3}$. Then the longitudinal velocity in a steel rod is

$$c = \sqrt{2.0 \cdot 10^{11}\,\frac{\mathrm{kg}}{\mathrm{m\,s^2}} / 7.85 \cdot 10^3\,\frac{\mathrm{kg}}{\mathrm{m^3}}} = 5.05\,\frac{\mathrm{km}}{\mathrm{s}}\,.$$

12.18 High-Frequency Longitudinal Waves in Air. The Acoustic Replica Method[C12.11]

C12.11. Sections 12.18 and 12.19 were the topic of POHL's farewell lecture, given on July 31st, 1952. An audio recording of the lecture is provided in the biographical film "*Einfachheit ist das Zeichen des Wahren*" ("Simplicity is the Mark of Truth") (Vol. 2).

As the source for high-frequency sound waves, we use the flue pipe which we encountered in Fig. 11.35. It emits longitudinal waves with spherical symmetry. Figure 12.43 can serve to illustrate this. It shows a section of a meridional plane as a stopped-motion image. We see a periodic distribution of air pressure and density. In the dark sketched wave crests, the pressure and density are greater, while in the light sketched wave troughs, they are less than in undisturbed air. The sinusoidal line added at an angle below represents the same phenomenon. The straight line indicates the normal air pressure p, while the sine curve shows its deviations Δp to higher and lower pressures. The absolute values of the amplitudes Δp_0 will be given later in Sect. 12.24. The whole distribution which is visualized as an instantaneous image in Fig. 12.43 moves with spherical symmetry outwards, at a velocity of around 340 m/s.

The changes in air density can be made directly visible by converting the travelling sound waves into standing waves. This is illustrated in Fig. 12.44 (following the scheme shown in Fig. 12.9, right). The

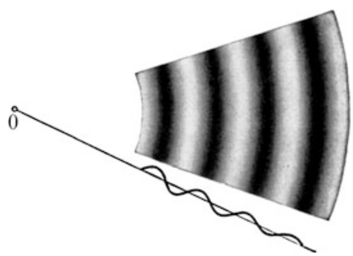


Figure 12.43 The propagation of spherical travelling sound waves in air (stopped-motion image). Medium gray indicates the normal density of air.

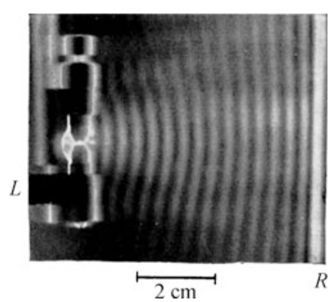


Figure 12.44 Standing sound waves in air in the interference field in front of a wall at R ($\lambda = 8$ mm, time exposure using the schlieren method (see Fig. 11.32 and Vol. 2, Sect. 21.11); L: supply hose for compressed air). In this image, up and down are reversed.

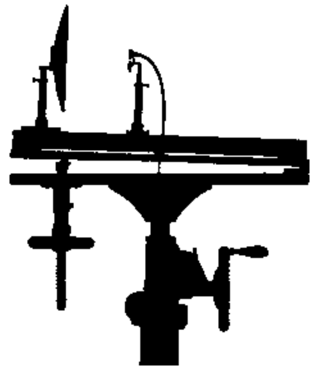


Figure 12.45 A sound source in the position in which it is used in Fig. 12.46; it is a flue pipe like that shown in Figs. 11.35 and 12.44

regions of constant density and those where the density changes periodically, i.e. the nodes and maxima, are observed as a schlieren pattern with a dark-field method.

Understanding a dark-field image requires an elementary knowledge of geometric optics; we have to understand the role of the pupils of the eyes. For the *acoustic replica method*, this is not necessary.

In the acoustic replica method, a collimated beam of waves (Fig. 12.45) passes at a grazing angle over the surface of a liquid (water or petroleum ether; Fig. 12.46). At the right end, the beam is reflected by a panel R, i.e. a "mirror" for sound waves. The reflected waves are superposed on the incident waves, producing standing waves in the air in front of the mirror. Directly beneath their pressure (and density) maxima, the liquid surface is slightly deformed[4]: It exhibits striations, as seen in Fig. 12.47. They can be observed best at a grazing angle. For demonstrations to a large audience, they can be shown as schlieren in a bright field, by using a basin with a transparent bottom and placing a small light source underneath it.

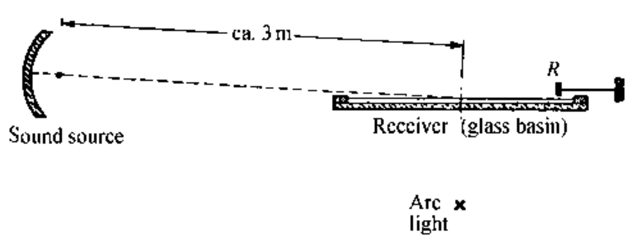


Figure 12.46 The acoustic replica method for demonstrating standing waves in a free sound field in air

[4] As a result of the pressure distribution in the standing wave, similar to that observed with the RUBENS flame tube (Fig. 11.30). This should be distinguished from transverse *surface waves* (for example as seen in Fig. 12.5 and the following figures).

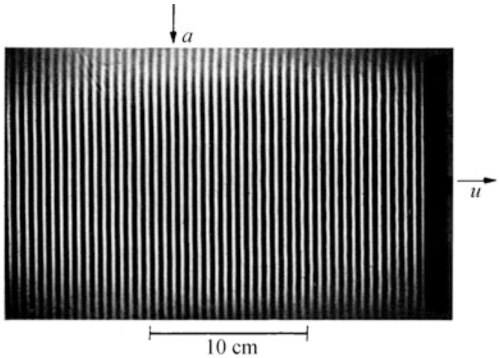


Figure 12.47 The interference field of acoustic plane waves in front of a flat mirror (reflector); standing waves (time exposure using the acoustic replica method; $\lambda = 1.15\,\text{cm}$, $\nu = 3 \cdot 10^4\,\text{Hz}$). A moving mirror ($u \neq 0$) would make the interference fringes drift, due to the DOPPLER effect.

If we move the reflector in the direction of the incident waves, or opposite to them, at a velocity u ($\ll c$), then *the interference field also moves*. The nodes of its standing waves pass by a given observation point (e.g. the point marked with the arrow a in Fig. 12.47) at the "beat frequency" $\nu_B = 2u/\lambda$. This is a result of the DOPPLER effect which acts twice (on the incident and on the reflected waves)[C12.12].

C12.12. This is a particularly impressive and perhaps surprising demonstration of the DOPPLER effect (Sect. 12.2)!

If a wave of frequency ν strikes the moving reflector, the latter *receives* the wave with the frequency $\nu'' = \nu(1 \pm u/c)$. The reflected wave, which is "emitted" by the reflector, has the frequency $\nu' = \nu''(1 \pm u/c) = \nu(1 \pm u/c)^2$, and for $u \ll c$, we have $\nu' = \nu(1 \pm 2u/c)$. Thus, two oppositely-travelling waves with a frequency difference of $\nu' - \nu = \Delta\nu = 2u/\lambda$ interfere with each other. This frequency difference causes the interference fringes to *wander* (Sect. 12.4); here the *nodes* of the standing wave. They pass the observation point, e.g. a, at the beat frequency $\nu_B = \Delta\nu = 2u/\lambda$. It follows from this that $\nu_B\lambda/2 = u$; or, in words: Using the readily measurable beat frequency ν_B, (for example with electromagnetic waves of short wavelength), the velocity u of the reflector (e.g. a moving automobile!) can be determined (**Exercise 12.7**).

In the acoustic replica method, the form of the reflector R can be varied in many ways. Figure 12.48 shows some examples (time exposures!).

Figure 12.49 is also very instructive. There, the reflector for the acoustic replica method is a hand. In the course of the demonstration, its shape is changed. In the language of optics, we say that the hand is a "non-luminous source"; we see the "secondary radiation" which it "emits" by reflection. Acoustically, this means that we "see" the high-frequency acoustic radiation with which bats can recognize obstacles and find their prey in complete darkness or without using their eyes. This is the ancient acoustic archetype of radar technology for localizing aircraft or other objects.

"This is the ancient acoustic archetype of radar technology for localizing aircraft."

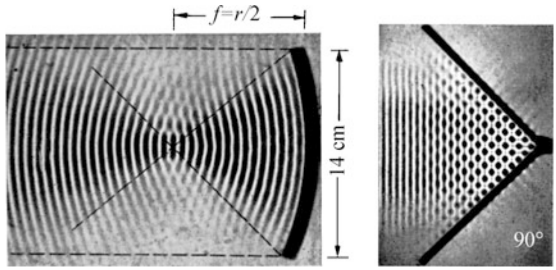


Figure 12.48 The interference fields of short-wavelength acoustic waves in front of a concave cylindrical mirror (*left*), and a 90° angular mirror (*right*). – At the left, the focal point is clearly visible (compare Fig. 12.21). – At the right, the acoustic replica method shows an especially pronounced interference pattern. It is suitable for convenient and low-cost detection of high-frequency sound waves.

Figure 12.49 The acoustic replica method makes visible the interference field of high-frequency sound waves in front of a hand which is reflecting them (a 'snapshot', i.e. the hand did not move during the exposure time)

12.19 The Radiation Pressure of Sound. Sound Radiometers

For the quantitative investigation of sound fields, the *acoustic* (or sound) *radiometer* is particularly useful. This instrument is based on a little-known but significant fact: Every surface which is struck by sound waves experiences a one-sided pressure in the direction of the sound waves. It is called the *radiation pressure* of the sound waves, analogous to the radiation pressure of light. This constant, one-sided pressure should not be confused with the sinusoidally varying pressure of the sound waves themselves (see Sects.12.18 and 12.24). (A thin membrane struck by sound waves not only oscillates at their frequency, but also bulges to one side in the direction of propagation of the sound waves!)

For a qualitative demonstration of the radiation pressure, the small "pinwheel" sketched in Fig. 12.50 can be used. We place it in front of a concave mirror so that the focus point of the mirror is on one of its blades. The higher the acoustic radiation power of the sound

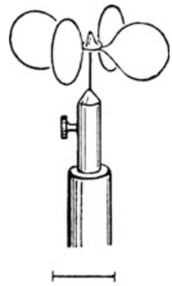


Figure 12.50 A "pinwheel" as an indicator for short-wavelength sound waves

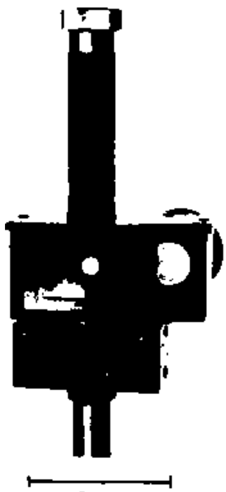


Figure 12.51 An acoustic radiometer. At the right, we see the circular rotor behind an inclined glass window, and behind it on the housing, the edge of the entrance aperture. For receiving a collimated beam of sound waves, one puts this aperture at the focus point of a concave mirror which concentrates the waves **(Video 12.3)**.

Video 12.3:
"The acoustic radiometer"
http://tiny.cc/gggvjy

waves, the faster the rotation frequency of the wheel. It represents a practical receiver for high-frequency sound waves.

If we use only *one* blade, and replace the bearing made of a glass thimble and a sharp pin by a fine metal band which is under tension, then we have a *sound radiometer*, a quantitative measuring instrument. Today, one can purchase small, easy-to-use models with magnetic damping and a short, aperiodic reaction time (ca. 2 s). Figure 12.51 shows an instrument of this type. Its scale deflection can be read off using a mirror and a light-beam pointer.

We will explain the origin of the radiation pressure of sound waves and its magnitude by referring to Fig. 12.52. The two straight lines indicate the boundaries of a collimated beam of travelling sound waves. Within it, the air particles flow sinusoidally back and forth in the direction of the double arrows with a maximum velocity u_0. This decreases the pressure p within the beam according to BERNOULLI's equation by an amount $\frac{1}{2}\varrho u_0^2$ (ϱ is the density of the air). As a result, air flows inwards from outside the beam. If the beam now strikes an absorbing wall at the right with perpendicular incidence, the velocity of the air particles at the wall is zero; thus the pressure there increases by an amount equal to the stagnation pressure $p_S = \frac{1}{2}\varrho u_0^2$. This is the radiation pressure. – The quantity $\frac{1}{2}\varrho u_0^2$ is however

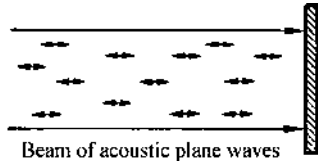


Figure 12.52 The origin of acoustic radiation pressure

at the same time given by the quotient

$$\frac{\text{Kinetic energy within the volume } V \text{ of the sound field}}{\text{Volume } V \text{ of the sound field}}\;;$$

that is, the spatial (volume) density δ of the acoustic energy (the derivation is given in Sect. 12.24). Then for the radiation pressure of a sound wave, we have[C12.13]

$$p_{\mathrm{S}} = \delta = \frac{\text{Radiation strength } b}{\text{Sound velocity } c}\,. \qquad (12.29)$$

It is doubled when the irradiated surface is completely *reflecting*.

C12.13. The relation between radiation pressure and radiation energy density (Eq. (12.29)) is complex. A detailed treatment is given by Boa-Teh Chu and Robert E. Apfel, *Journal of the Acoustical Society of America* **72**, 1673 (1982).

12.20 Reflection, Refraction, Diffraction and Interference of 3-Dimensional Waves

Three-dimensional waves, whose wavelengths are of the order of 1 cm, are especially suitable for experimentally demonstrating the most important fundamentals of wave theory. We make use of *high-frequency sound waves in air*, usually in the form of parallel-collimated beams (Fig. 12.45). As receiver, we use a sound radiometer (Fig. 12.51).

Most of these experiments could also be demonstrated using electromagnetic waves. Often, one can even use the same accessories for both types of waves. We need only replace the flue pipe as wave source by one of the small, commercially-available transmitters for electric or electromagnetic waves (e.g. microwaves), and the sound radiometer by a small receiver antenna and detector.

Out of the large number of impressive demonstration experiments, we can present only a small selection in the following sections.

1. *Shadowing*: We direct the sound source (Fig. 12.45) towards the receiver and place an obstacle, for example a human body, in the sound-wave beam.

Shadowing of sound waves can, by the way, be demonstrated quite convincingly without any instrumentation at all. Rub the thumb and forefinger of your right hand together at about 20 cm from your right ear. You will hear a high tone, not dissimilar from that of our flue pipe. Then use your left hand to block your right ear. You will no longer hear anything at all, since the left ear lies completely within the sound shadow of your head.

2. *Reflection, lattice planes as a mirror*: The reflection of sound waves has already been demonstrated in some of our earlier experiments (e.g. Figs. 12.44–12.49). Here, a few supplementary remarks will suffice: The reflecting surfaces need not be smooth; they can in fact consist of *lattice planes*. Two examples are illustrated by Fig. 12.53. There, the distance between the balls or the apertures from their neighbors must be of the order of the wavelength of the sound waves. Then, one can readily demonstrate their specular reflection from these lattice planes using the setup sketched in Fig. 12.54.

In this setup, the angle of incidence β_2 can be varied by moving the sound source which is attached to a swivel-mounted arm. A small auxiliary apparatus (a pantograph) moves the reflecting surface R at the same time by an angle of β_2. Then the reflected beam maintains a fixed direction.

Therefore, we can use a fixed receiver. – We find a strong reflectivity from the lattice planes (Fig. 12.53) at every arbitrary angle of incidence.

The reflection of sound waves from the interface between warm and cool air is also quite impressive. Figure 12.56 shows a suitable setup. Using a comb-shaped gas burner (Fig. 12.55), we prepare a nearly planar vertical wall of warm air, with a low density. It clearly reflects

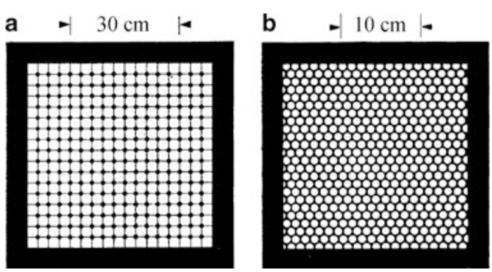


Figure 12.53 Lattice planes which act as mirrors for sound waves

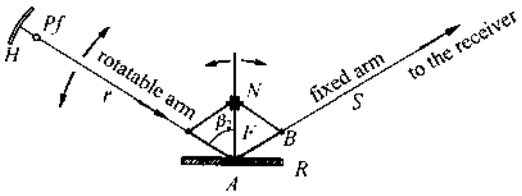


Figure 12.54 Demonstrating the law of reflection for a fixed direction of the mirrored beam; the sound source is the same as in Fig. 12.45. The receiver is a concave mirror at whose focus the input aperture of the sound radiometer (or a microphone) is located.

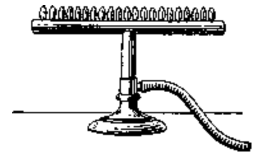


Figure 12.55 The gas burner used to prepare a vertical layer of warm air in Fig. 12.56

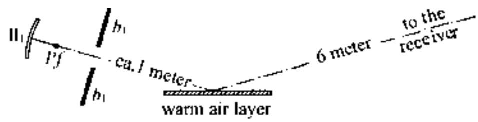


Figure 12.56 Specular reflection of a beam of acoustic plane waves by a hot air layer[C12.14]

C12.14: Air at the surface of a lake is often cooler than the air higher up. Then sounds, for example the voices of people at water level, are reflected at the cool/warm air interface and conducted along the surface as in a two-dimensional speaking tube, so that the sound "carries further".

the collimated beam from the source, although not as precisely as a wooden or metal mirror.

3. *Refraction*: For the demonstration of refraction of sound waves, we use a prism which is filled with carbon dioxide gas (Fig. 12.57). Its transparent walls are best made of silk (paper, cellophane, or plastic are nearly opaque to the sound waves). After the prism is filled with carbon dioxide, we observe an angle of refraction amounting to $\delta = 9.8°$.

The angle of incidence α and the angle of refraction β are sketched for the second interface. The angle α is 30° for the prism. From the sketch, we can read off $\beta = \alpha + \delta$, and thus $\beta = 39.8°$. From this, we find a relative index of refraction of

$$n_{CO_2 \to \mathrm{air}} = \frac{\sin 30°}{\sin 39.8°} = \frac{0.5}{0.64} = 0.78\,.$$

4. *Scattering*: We point the beam of sound waves from the source directly towards the receiver and then interpose the hot flame gases from a gas burner which is swung back and forth through the beam; or we allow the beam to pass through a stream of carbon dioxide gas poured from a watering can. In both cases, the waves are scattered randomly in all directions by reflection and refraction ("diffuse reflection"), so that they no longer reach the receiver. Nothing more can be discerned of the originally sharply collimated beam of sound

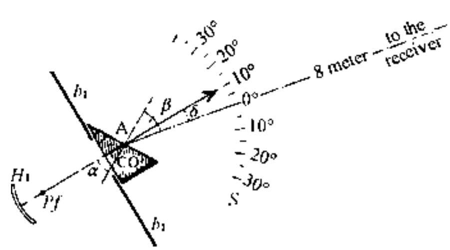


Figure 12.57 The refraction of a beam of acoustic plane waves by a prism filled with CO_2 gas. α is the acute angle of the prism, and at the same time the angle of incidence at the second face of the prism. The scale and the receiver are fixed, while the prism and the flue pipe *Pf* can be rotated (see also Fig. 12.60).

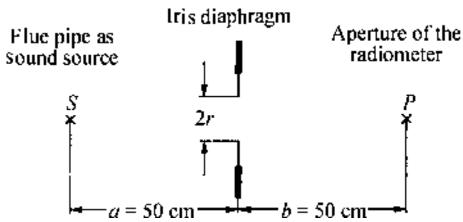


Figure 12.58 Masking out FRESNEL zones for short-wavelength sound waves (**Video 12.4**)

Video 12.4:
"Fresnel's zones"
http://tiny.cc/fggvjy
A setup similar to that in Fig. 12.58 with a variable opening $2r$ (iris diaphragm) and a microphone as sound detector at P is used to show the variations of the amplitude (*not* the radiation power) on masking different numbers of zones.

waves; it has been completely destroyed by *schlieren* in the air or the *turbid medium*.

5. *Fresnel zones.* We refer to Sect. 12.14. – In Fig. 12.58, let S be a small flue pipe as wave source and the observation point P the entrance aperture of the receiver (acoustic radiometer). In the center between the two is a large iris diaphragm, framed by an opaque plate. The distances a and b are adjusted to 50 cm, and the wavelength of the sound emitted by the source is chosen to be $\lambda \approx 1$ cm. Then from Eq. (12.21), we obtain

for the	first	second	third zone etc.
the diameter $2r_m =$	10 cm	14.1 cm	17.3 cm etc.

The indications of the radiometer are proportional to the radiation power which passes from the iris to the observation point P. If for example we choose $2r_1 = 10$ cm, then we observe a signal of $\alpha = 16$ scale divisions. When we *open* the iris from 10 cm to 14 cm, the power arriving at P is *reduced*; we find only $\alpha_2 = 3$ divisions. A further opening to 17 cm again increases the power to ca. 15 divisions, and so forth. Figure 12.59 shows a complete series of measurements.

6. *Diffraction by a slit*: The phenomena that we encountered in Figs. 12.14 and 12.32 are demonstrated. The setup is sketched in Fig. 12.60, and Fig. 12.61 shows an image of the slit, while Fig. 12.62 gives the results of the measurements.

Compare with Fig. 12.35: The values indicated by the sound radiometer are proportional to the *squares* of the amplitudes, so that the submaxima alongside the principal maximum as seen in Fig. 12.62 are relatively much smaller than in Fig. 12.35.

Figure 12.63 shows the same measurements as Fig. 12.62, but now represented in polar coordinates. Here, the radius r represents the signal indicated by the radiometer, or the radiation power, which is proportional to it. – Polar coordinates are preferred for technical purposes ("directional characteristic").

7. *Narrowing of interference fringes by a lattice array of slits serving as wave sources*: Interference fringes become more sharp and narrow when the wave source points are arranged in a regular lattice. This was derived using model experiments in Sect. 12.15. Experi-

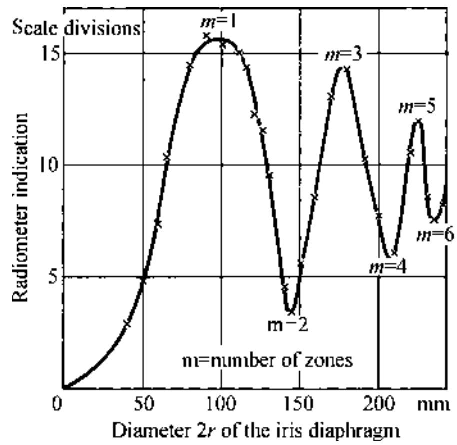


Figure 12.59 The radiation power arriving at the observation point P in Fig. 12.58 for various settings of the diameter $2r$ of the iris diaphragm

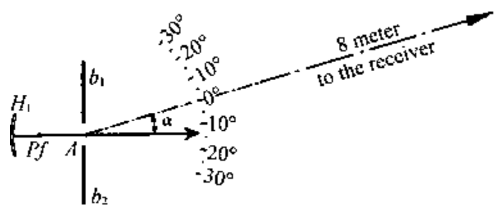
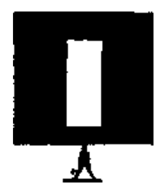


Figure 12.60 Passage of acoustic plane waves through a slit

Figure 12.61 The diffraction slit used in Fig. 12.60; its width is $B = 11.5\,\mathrm{cm}$



mentally, lattice arrangements of wave source points can be realized either in the form of apertures (usually slits) or as mirror images. We start with the first method; the second will be discussed under item No. 8. We first demonstrate the transition from YOUNG's interference experiment to a lattice of wave sources. To this end, the wide slit in Fig. 12.60 (see Fig. 12.61) is replaced first by two and then by five equidistant *narrow* slits (Fig. 12.64). The principal maxima of the interference pattern keep their positions during this process (Fig. 12.65), but with five interfering wave trains, they are considerably stronger and narrower than with two. The small submaxima that lie between them form a nearly continuous background. Five slits already demonstrate the characteristics of a typical *lattice* or *grating*[5],

[5] Also called *diffraction grating*. Concerning this name, we refer to Sect. 12.15.

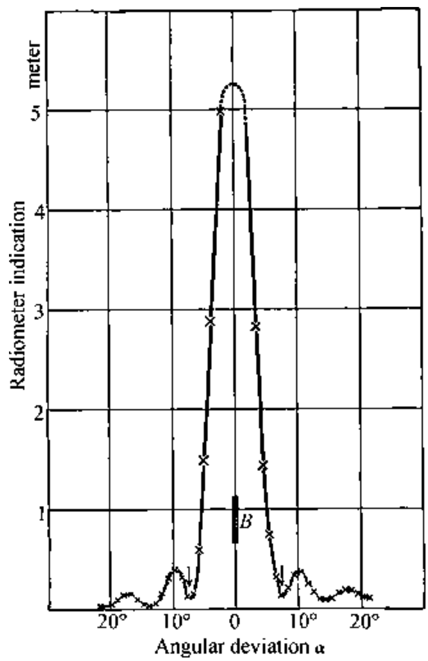
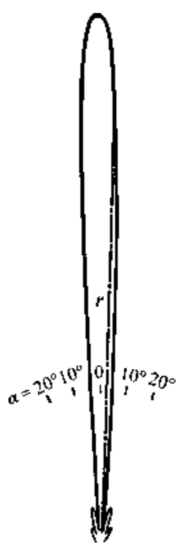


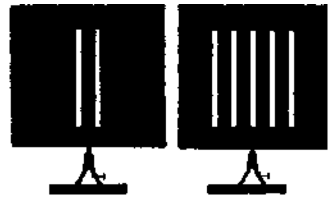
Figure 12.62 The diffraction pattern of the slit shown in Fig. 12.61 at a wavelength of 1.45 cm. The shaded area B marks the geometric limits of the beam (geometric shadow).

Figure 12.63 The diffraction pattern from Fig. 12.62, shown in polar coordinates



as used in optics for spectral investigations. Under the topic 'optics' (Vol. 2, Chaps. 21 and 22), we will treat the properties of optical gratings in more detail.

Figure 12.64 The interference of two and five wave trains from equidistant slits; the spacing of neighboring apertures (center to center) is called the lattice constant D



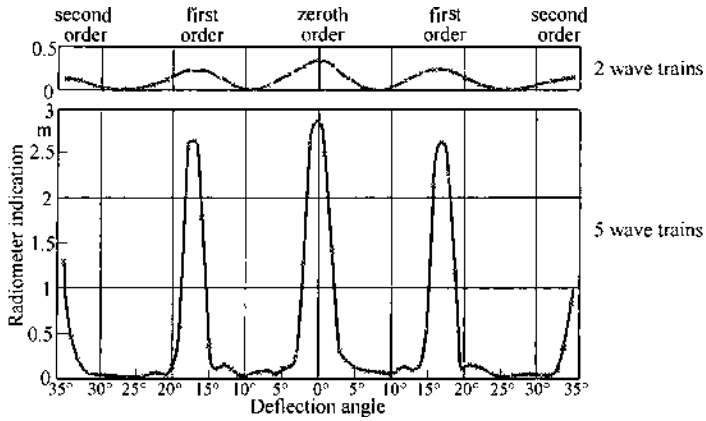


Figure 12.65 The transition from YOUNG's double-slit experiment to a lattice or grating (Fig. 12.64), i.e. a lattice-like arrangement of slits which serve as wave source points. This results in a narrowing of the interference fringes by superposition of more than two wave trains.

8. *Narrowing of interference fringes by a lattice array of mirror images serving as wave sources*: We refer to Fig. 12.41c and make use of the experimental setup shown in Fig. 12.54. The mirrors A consist of four lattice planes with dimensions as given in Fig. 12.66, left. We thus use only four mirror images as wave sources. Nevertheless, the demonstration experiment shows two rather sharp interference maxima ("spectral lines" with the indices $m = 3$ and $m = 4$) (Fig. 12.66, right). In the model experiment (Fig. 12.39c), we could clearly recognize $(N-2) = 2$ submaxima between the principal maxima. They are not visible in this demonstration experiment.

9. The *interferometer*: The archetype of all interferometers is the interference setup which was described by THOMAS YOUNG in 1807 for light waves (two wave trains and two slits as source points (Figs. 12.39a and 12.64)). It is still in use today for many measurements in laboratories and in technology.

The radiation power transmitted by the two narrow slits is rather small. Therefore, later on other designs were found which make it possible to observe interference from collimated beams of large cross-section. All of them use reflection or refraction to separate one beam of waves into two beams. Out of the many designs, we will show here only one which is particularly important in physics,

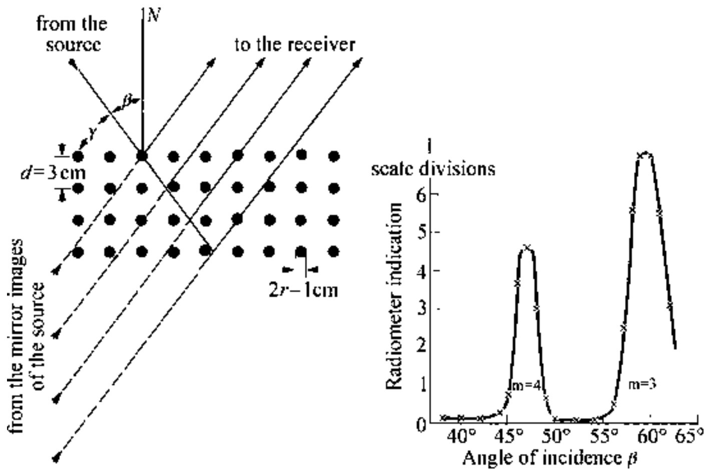


Figure 12.66 The reflection of sound waves ($\lambda = 1.03$ cm) from four equidistant lattice planes. According to Fig. 12.41c, these produce four mirror images as wave sources, arranged in a lattice. For the angles β at which the reflected waves exhibit maxima, while the transmitted waves are minimal, cf. Eq. (12.23). γ is often called the Bragg angle.

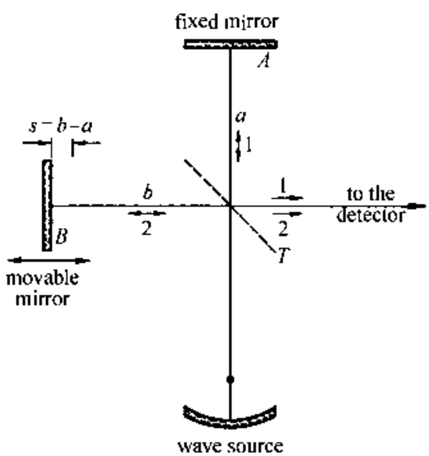


Figure 12.67 The interferometer setup of A.A. MICHELSON

the MICHELSON interferometer (Fig. 12.67). In this design, the two reflecting surfaces are placed perpendicular to one another; T is a "beam splitter", i.e. a surface which reflects about half of the incident waves and transmits the other half. The phase difference of the resulting two wave trains is determined by their path difference s.

12.21 The Origin of Waves on Liquid Surfaces

We have discussed the most important facts concerning the propagation of waves or radiation using longitudinal waves in air and transverse surface waves on water. We showed how longitudinal waves are produced in Sect. 12.17. In this section, we will treat the formation of surface waves. The results will lead us to some insights which are significant for waves of all types.

The waves on liquid surfaces can be represented by a simple picture of sinusoidal oscillations only in the limiting case of very small amplitudes. In general, the wave troughs are broad and flat, while the crests are narrow and high. Figure 12.68 shows an instantaneous image of a water wave which is travelling to the right. – The formation of such a wave can be observed with a wave tank. This is a long, narrow channel made of sheet metal with glass windows in its sides (around 150 × 30 × 5 cm). It is half-filled with water. As we have seen before, aluminum flakes are mixed into the water and can be observed as suspended particles (Chap. 10). A movable bar serves to initiate the wave motion; it can be displaced up and down by a motor. When the wave travels along the channel, we can observe a streamline pattern as shown in Fig. 12.69. This is a time exposure of 0.04 s duration. The streamline pattern shown in the figure would be seen by an observer who is at rest in the lecture room. It shows the distribution of directions of the velocities of the water particles.

In a wave, the motion of the liquid is not stationary. As a result, the paths followed in the course of time by individual water particles are by no means identical to the streamlines (cf. Sect. 10.5). These

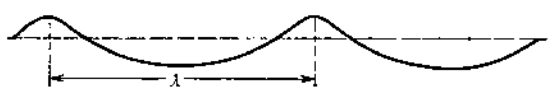


Figure 12.68 The profile of a water wave of small amplitude. The wave crests do not "tip over" and no foamy breakers are formed.

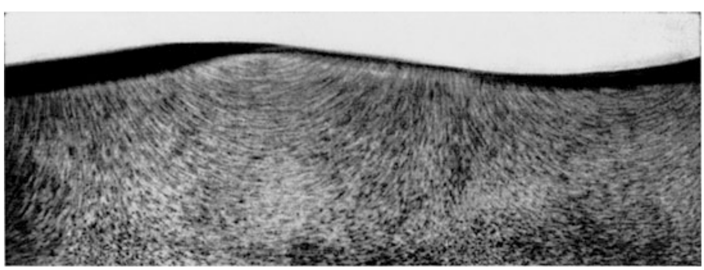


Figure 12.69 Streamlines in a travelling water surface wave (photographic positive exposure with bright-field illumination)

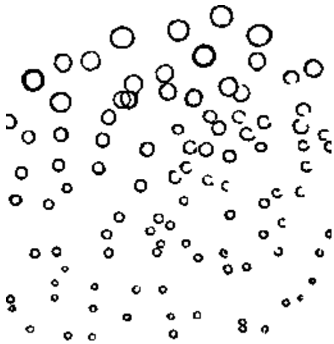


Figure 12.70 The circular motion of individual liquid 'volume elements' (the orbital motion) in a travelling water wave (photographic negative with dark-field illumination). The upper edge of the picture does not show the outline of a wave, but rather is due to the random distribution of the aluminum flakes.

paths look quite different. For moderate wave amplitudes, they are to a good approximation circular. We can observe these circular orbits both at the surface and also at considerable depths in the water. However, the diameter of the circular orbits is greatest for the water particles in the uppermost water layers, near the surface.

To demonstrate these circular orbits of the individual water particles (the "orbital motion"), we add a small amount of suspended aluminum flakes to the water. Furthermore, we take the period of the wave motion to be the exposure time for our photographs. In this way, we can obtain images such as the one shown in Fig. 12.70.

Based on our experimental observations, we arrive at the scheme which is sketched in Fig. 12.71. It shows the circular orbits of several liquid particles which are near the surface. Their diameter $2r$ is equal to the difference in height between the tips of the wave crests and the bottoms of the wave troughs.

We will call the orbital velocity on the circular paths w, i.e.

$$w = \frac{2\pi r}{T}\,.$$

The time T for a complete rotation corresponds to the advance of the wave by a full wavelength λ.

Figure 12.71 The correlation between the streamlines and the circular motions within travelling water waves. The horizontal row of points indicates volume elements at the water surface which are in their rest positions; the clockwise circular arcs connected to them show their paths. By connecting the small *arrow points*, we obtain the profile of the wave (which is travelling to the right) at the end of the next time interval. The circular motion is shown only for each second velocity vector arrow.

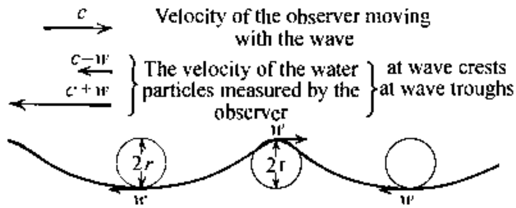


Figure 12.72 The orbital motions (*orbital velocity* w) of the water particles as seen by an observer who is *moving with* the travelling wave

To simplify the computation, we assume a surface between water and air. Initially, we will ignore the density and motions of the air relative to those of the water. Furthermore, from now on we assume that the observer is moving to the right at the phase velocity c. For this "observer in the moving frame", the wave as a whole is at rest; its outline appears to be frozen in its motion.[6] But as a result, the individual liquid particles are zipping past the observer to the left at high velocities (Fig. 12.72).

Thus, a volume element of the water in a wave trough has a velocity of $u_1 = c + w$, and at the crest of a wave, a velocity of $u_2 = c - w$. These velocities produce *stagnation pressures*; in a trough, $p_1 = \frac{1}{2}\varrho(c+w)^2$, and at a crest, $p_2 = \frac{1}{2}\varrho(c-w)^2$. In a wave trough, the stagnation pressure p_1 acts to produce a deepening; at a crest, the pressure p_2 tends to produce a flattening. In a trough, the surface of the liquid is subject to two pressures with forces that are in mutual equilibrium: The difference of the two stagnation pressures, i.e. $p_1 - p_2 = 2\varrho wc$, pulls downwards. A static pressure $p = \varrho gh = \varrho g2r$ pushes upwards. This static pressure results from the vertical distance h between the crests of the wave and its troughs. It follows from $p_1 - p_2 = p$ that

$$wc = gr\,. \qquad (12.30)$$

For small orbit amplitudes r, we find

$$w = \frac{2\pi r}{T} \quad \text{and} \quad T = \frac{\lambda}{c}\,, \quad \text{thus} \quad r = \frac{w}{c} \cdot \frac{\lambda}{2\pi} \qquad (12.31)$$

and so

$$c^2 = \frac{g\lambda}{2\pi}\,. \qquad (12.32)$$

(For example ground swells, with $\lambda = 1\,\text{km}$, $T = 25.4\,\text{s}$ and $c = 142\,\text{km/h}$; or $\lambda = 50\,\text{m}$, $T = 5.7\,\text{s}$ and $c = 32\,\text{km/h}$).

[6] The "observer in the moving frame" can thus apply the knowledge gained from Figs. 10.11 and 10.12.

In deriving this equation for the calculation of the static pressure, the *surface tension* ζ was neglected in comparison to the weight of the liquid. This is permissible down to wavelengths of the order of 5 cm. For still smaller waves, the term $2\pi\zeta/\lambda\varrho$ should be added into Eq. (12.32), and then we obtain

$$c^2 = \frac{g\lambda}{2\pi} + \frac{2\pi\zeta}{\lambda\varrho}\,. \tag{12.33}$$

The *phase velocity* c for surface waves as derived from this equation is shown later as a graph in Fig. 12.74. – When the first term is dominant, we speak of *gravity waves*. When the second dominates, we have *capillary waves*. For capillary waves alone, we find

$$c^2 = \frac{2\pi\zeta}{\lambda\varrho}\,. \tag{12.34}$$

Derivation

The static pressure $p = 2r\varrho g$ which results from the height difference h between the wave crests and the troughs of the waves can be neglected; instead, we consider the static pressure which is due to the *curvature* of the wave crests and troughs, $p = \zeta/r + \zeta/r = 2\zeta/r$ (from Eq. (9.9; see Fig. 12.73). It provides the counter-force to the difference of the two stagnation pressures, i.e. the pressure $p_1 - p_2 = 2\varrho wc$. We find

$$2\varrho wc = \frac{2\zeta}{r}\,. \tag{12.35}$$

For the circular orbits of the water particles, we found Eq. (12.32). For $w = c$, a circular orbit of diameter $2r = \lambda/\pi$ provides a good approximation to the curvature of a sinusoidal wave. Inserting $w = c$ and $r = \lambda/(2\pi)$ into Eq. (12.35), we obtain Eq. (12.34).

Equation (12.32) remains applicable down to a water depth of $h \approx 0.5\lambda$. In the opposite limit of a vanishingly small water depth h, the propagation velocity of the shallow-water waves becomes independent of λ; the same value holds for all wavelengths[C12.15]:

$$c^2 = gh\,. \tag{12.36}$$

Consequences: Long ground swells of small amplitude can rise up to disastrous heights when they run onto a beach with a gentle slope

C12.15. Equation (12.36) is important for the demonstration of how surface waves break, Figs. 12.19, 12.22 and 12.23. It is derived here for amplitudes comparable to the water depth, but holds also for smaller amplitudes. The tsunami phenomenon can also be understood from energy conservation: The wave loses kinetic energy near the shore, due to the reduction of c, and this is transferred to its potential energy, so that its height (amplitude) increases. See also **Video 12.2** (http://tiny.cc/tfgvjy).

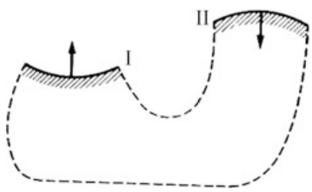


Figure 12.73 The addition of two static pressures produced by two curved surfaces. The liquid, bordered here by dashed lines, has an arbitrary shape. Its surface I presses upwards and is pressed upon (owing to the isotropy of pressure) by the surface II.

(tsunami). The reduced velocity near the shore causes a "pile-up" effect so that the water from the rear of the swell overtakes the water at its front; a very large volume of water thus reaches the shore in a short time, and can inundate the nearby (normally dry-land) regions up to depths of more than 10 m.

Derivation

In deep water, the gravity waves propagate to the right as shown in Fig. 12.71. There, with the decay of the wave at the left side and its buildup on the right side of each wave crest, the velocity vectors trace out circles. – When the depth h of the water is reduced, we arrive at the limiting case of *ground waves*. The amplitude of the gravity waves has become practically equal to the depth h of the water at rest. Below a wave trough, the water is flowing along a horizontal line. It flows to the left at a velocity $w = c$ in a wave trough which is moving to the right at a velocity c; for an observer who is "moving with the wave", this velocity is thus $c + c = 2c$. For such an observer, a stagnation pressure of $\frac{1}{2}\varrho(2c)^2 = 2\varrho c^2$ is in equilibrium with a static pressure. The latter arises as with all gravity waves through the height difference between the crests of the waves and their troughs; thus

$$2\varrho c^2 = 2h\varrho g \quad \text{or} \quad c^2 = gh\,.$$

Thus far, we have neglected the fact that a second medium is located above the liquid surface. This medium is air, and we have specifically neglected its effects. We now relax this constraint. Above the liquid of density ϱ, a second fluid of density ϱ' is presumed to be present. Then, instead of Eq. (12.33), we have

$$c^2 = \frac{\varrho - \varrho'}{\varrho + \varrho'}\frac{g\lambda}{2\pi} + \frac{2\pi}{\lambda}\frac{\zeta}{\varrho + \varrho'}\,. \tag{12.37}$$

Examples

We offer three examples:

1. In two layers of the atmosphere (one above the other), as a result of temperature differences, the mass densities ϱ and ϱ' may be different. Then at the interface between the layers, there can be waves. They can be seen through the periodic condensation of water in the form of 'ribbed' cirrus clouds.

2. *Dead water*: Not far from the mouths of rivers, especially in Scandinavian fjords, one can frequently observe the phenomenon of "dead water". Ships which are moving slowly, i.e. at 4 to 5 knots ($\approx$ 8–10 km/h), are suddenly brought to a stop by an apparently invisible force, while sailing vessels no longer obey the helm.

Explanation: Fresh water with a lower density forms a layer above salt water with a higher density. The vessel reaches down into the interface region between the two layers. Its motion generates high-amplitude waves along this interface, which however remain invisible to the eye above the water surface. The visible water surface (interface between water and air) remains practically at rest. The vessel must provide all the energy for this wave motion; this causes its noticeable slowing down. This case is thus similar to the formation of a flow resistance for objects in a fluid flow due to the formation of vortices behind it.

3. For demonstration experiments, we sometimes need waves with a very slow propagation velocity. Then we can pour a layer of petroleum onto water, mark the interface with aluminum dust, and insert a flat excitation

bar into the interface region. It can produce waves of high amplitude there, while the surface of the petroleum, its interface with the air, remains practically at rest.

12.22 Dispersion and the Group Velocity

The content of Eq. (12.33) is graphically represented in Fig. 12.74: For surface waves, the phase velocity c depends on their wavelength (top image), and the waves exhibit *dispersion*, defined as the quotient $\mathrm{d}c/\mathrm{d}\lambda$, which is a function of their wavelength (bottom image). – If dispersion is present, then the phase velocity $c = \nu \cdot \lambda$ can always be determined experimentally when the waves are excited at a frequency ν and their wavelength λ can be measured[7]. Two examples can be seen in Fig. 12.75.

When the excitation is limited to a single frequency, this *instantaneous image* of the travelling *wave* (Fig. 12.1) corresponds to the curve of a sinusoidal *oscillation* (Fig. 4.12). If waves are excited at

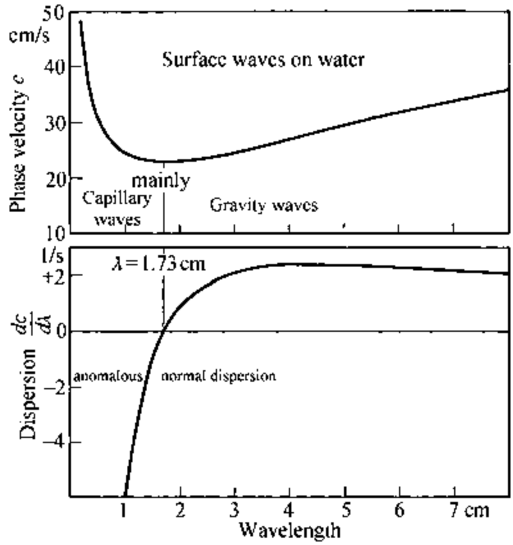


Figure 12.74 The phase velocity (*top*) and dispersion (*bottom*) of flat, nearly sinusoidal surface waves on water with various wavelengths as used for demonstration experiments[C12.16]

C12.16. An example of the dispersion of these waves is shown in **Video 12.2**, where a short rectangular pulse, produced by dipping an excitation bar into the liquid just once, is stretched out into a long pulse (see also the "aged capillary waves" in Fig. 12.81). Short waves propagate faster than longer waves.

Video 12.2: "Experiments with water waves" http://tiny.cc/tfgvjy.

[7] This procedure is often considerably more precise than a direct measurement of the phase velocity from the distance travelled and the time required. The method using (KUNDT's) dust figures has long been applied in acoustics, and today, an analogous method is preferred for electric waves of short wavelength.

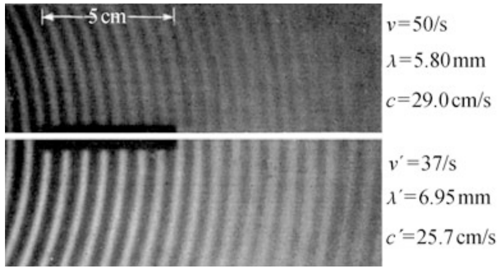


Figure 12.75 Measurement of the phase velocity of sinusoidal capillary waves on a water surface. We can see the shadow of a 5 cm long ruler.

two or more frequencies, then the *instantaneous images* of the travelling waves correspond to the graphs which result from superposition of sinusoidal oscillations of different frequencies and amplitudes. If dispersion is also present, then during the propagation of the resulting waves, the shape of the instantaneous images changes. With two excitation frequencies in the ratio 1 : 2, for example, the two curves S_r in Fig. 11.11 are transformed periodically into one another along the path of propagation.

These changes of shape are important in limiting cases. If the frequencies involved lie e.g. within a narrow region between ν and $\nu \pm \mathrm{d}\nu$, they produce wavelengths between λ and $\lambda \mp \mathrm{d}\lambda$. Then as a result of dispersion $\mathrm{d}c/\mathrm{d}\lambda$, a prominent point on the resulting waveform, e.g. the highest wave crest, does not move at the phase velocity c which belongs to λ, but rather at the *group velocity*[8]

$$c^* = c - \lambda \frac{\mathrm{d}c}{\mathrm{d}\lambda} \,. \tag{12.38}$$

To derive Eq. (12.38), imagine that in Fig. 12.76 there is a wave travelling to the right, which has resulted from the superposition of two sinusoidal waves. Its 'snapshot' image A is a "beat curve" (something like that in Fig. 11.10). The longer of the two sinusoidal waves (B) is assumed to have the wavelength λ and the phase velocity c; the shorter wave (C) has $\lambda' = \lambda - \mathrm{d}\lambda$ and the phase velocity $c' = (c - \mathrm{d}c)$.[9] Within the resulting wave (A), the two sinusoidal waves cannot be distinguished in any form. In order to find the velocity with which the resulting wave (curve A) is moving to the right, we need to define a *marker* on the waveform. It is convenient to use a maximum (crest) of the wave A. In the instantaneous image A, it is marked with the double arrow 1. This maximum lies over the wave crests γ and d (curves B and C). After a propagation time of Δt, both maxima have moved to the right. The maximum γ has travelled along a distance $s = c\Delta t$, while the maximum d has travelled a somewhat

[8] This term can lead to misunderstandings; note the last paragraph of this section.
[9] Corresponding to normal dispersion in optics (Vol. 2, Sect. 27.2). See also Fig. 12.74.

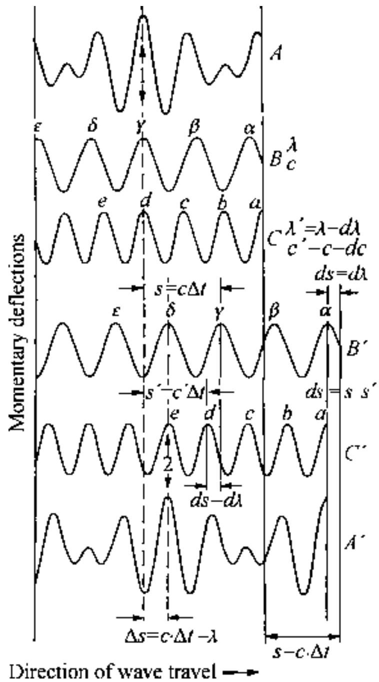


Figure 12.76 Wave graphs ("snapshot images") to illustrate the definition of the group velocity. We have chosen $\mathrm{d}\lambda = \Delta t \mathrm{d}c$ here.

shorter distance $s' = (c - \mathrm{d}c)\Delta t$. The lead $(s - s') = \mathrm{d}s = \mathrm{d}c\Delta t$ gradually approaches the value $\mathrm{d}\lambda$. This case is sketched in the three lower curves: The point of equal phases is now at the wave crests δ and e. That is, the maximum, the marker of the wave group, has not moved along a distance $c\Delta t$, but only along the smaller distance $\Delta s = (c\Delta t - \lambda)$. Therefore, the velocity of the marker, i.e. the group velocity, is

$$c^* = \frac{c\Delta t - \lambda}{\Delta t},$$

and this leads to Eq. (12.38), since we chose $\mathrm{d}s = \mathrm{d}c\Delta t = \mathrm{d}\lambda$.

The consequences of Eq. (12.38) can be – quantitatively! – explained by a demonstration experiment. The necessary apparatus is described in Fig. 12.77:

Two waves of differing wavelengths are represented by the shadow images of two gears, B and C. Black gear teeth mean wave crests, and the white gaps between them are wave troughs. These "waves" do not travel as in Fig. 12.76 along straight-line paths, but rather on circular paths. – The two gears are mounted on the same shaft, one behind the other, and can be rotated independently. Then in the shadow image A, the beat curve resulting from their superposition can be seen; it exhibits four wave groups. The gears are driven by a slow-running synchronous motor M. The velocities c and $(c + \mathrm{d}c)$ of the two gears can be conveniently adjusted by using pulley wheels of different diameters driven by elastic belts. A pointer Ph permits us to measure the phase velocities of the individual waves with a stopwatch.

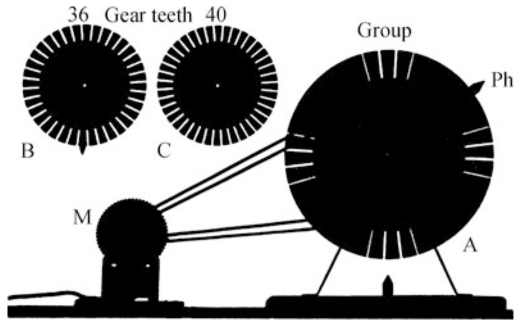


Figure 12.77 The quantitative demonstration of the group velocity (details are given in the text). The "rectangular" profile of the waves poses no problem here, just as in other geometric model experiments on wave theory, for example in Sects. 12.12 and 12.15. M is an electric motor[C12.17].

C12.17. E. Mollwo, *Physikalische Zeitschrift* **43**, 257 (1942).

We can choose at will whether the longer wave λ (B) or the shorter ($\lambda - \mathrm{d}\lambda$) ($C$) has the greater phase velocity. In the first case, the group velocity c^* is smaller than the phase velocity c, and the phase marker *Ph* overtakes the groups. In the second case, the groups overtake the phase marker. In the limiting case

$$\frac{\mathrm{d}c}{\mathrm{d}\lambda} = 0\,,$$

the groups have exactly the same velocity as the phase. In the limit $c\mathrm{d}\lambda = \lambda\mathrm{d}c$, the group velocity becomes $c^* = 0$. The groups remain fixed at one spot.

A strictly sinusoidal wave train in the mathematical sense has no beginning and no end, either in time or in space. Its spectrum will contain a single *spectral line*, as in Fig. 12.78a. Every sinusoidal wave that occurs in nature however has a beginning and an end. Therefore, it appears in its spectrum as a *very narrow band* (Fig. 12.78b). On both sides of the sine wave frequency, it includes a region $\mathrm{d}\lambda$, which is very narrow compared to its mean wavelength λ. This narrow region is filled with a dense series of spectral lines (a FOURIER integral). For the purposes of computations, we can as an approximation replace this dense series with two spectral lines that fall within the narrow region (Fig. 12.78c). We made use of this fact above, in order to derive Eq. (12.38) for the group velocity c^* and to demonstrate c^* experimentally. As a result of this approximation, instead of a single group, we obtained a sequence of identical groups (Fig. 12.77). *According to Eq. 12.38, a group velocity is defined only for wave groups whose spectrum includes a narrow range of wavelengths between λ and $\lambda \pm \mathrm{d}\lambda$.* One can thus not consider the group velocity to be simply the velocity of any arbitrary wave group.

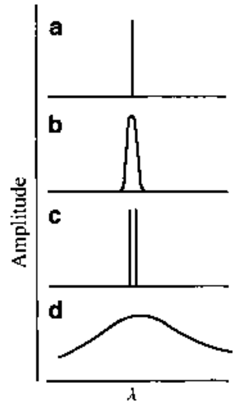


Figure 12.78 The spectra of unbounded waves (**a** and **c**), and of a short (**d**) and a long (**b**) wave group (short refers here to the *spatial* length of the group, not to its *wave*length)

12.23 The Excitation of Waves by Aperiodic Processes

The greatest contrast to mathematical sine waves is the propagation of *aperiodic* processes, e.g. the practically aperiodic wave group shown at the top of Fig. 12.79. Such wave groups have a broad, continuous spectrum; strictly speaking, they contain an unlimited range of wavelengths on each side of the mean wavelength λ (Fig. 12.78d).

If such an aperiodic wave group passes through a medium *without dispersion*, that is a medium in which the phase velocity is the same for all wavelengths, then the shape of the group along its path through

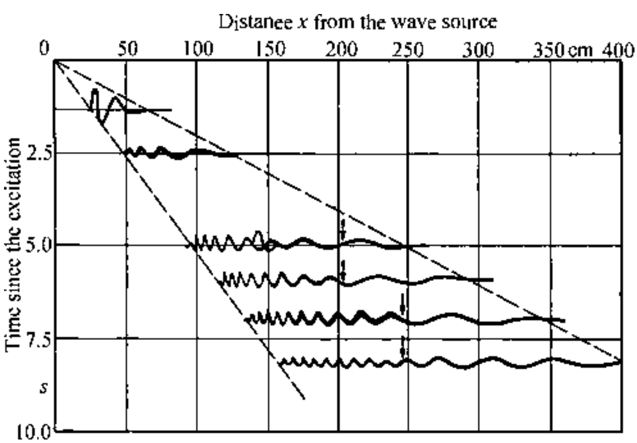


Figure 12.79 The time dependence of the production of gravity waves by an aperiodic disturbance of the water surface at the time zero ('snapshots', extended by computed values to the right of the arrows, because the wave tank was too short); upper image, 1.3 s after a single (i.e. aperiodic) pulse of the excitation bar (in a wave tank as described in Sect. 12.21)

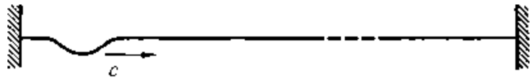


Figure 12.80 Along an elastic cord, an aperiodic wave group propagates without any change in shape (here, a 10 m long helical spring); its velocity follows from Eq. (11.4)

the medium remains unchanged. In this case, one can readily determine the common phase velocity c for all the waves experimentally. Transverse waves along an elastic cord offer a good example; cf. Fig. 12.80.

If, in contrast, an aperiodic wave group passes through a medium *with dispersion*, then the shape of the group will continually change along its path. This can be seen for example in gravity waves on a water surface, Fig. 12.79. (Their dispersion is *normal*, i.e. their phase velocities c increase with increasing wavelength.) – At the beginning of an experiment (upper left), an excitation bar is thrust into the water surface. After 1.3 s, we can observe a nearly aperiodic wave group. In the course of time, the group becomes longer and longer; at its front, longer waves form, while shorter ones collect behind the group.

Here, we have used a wave tank ca. 3 m long and 60 cm deep, with a straight bar for wave excitation. This bar is thrust once, aperiodically, into the surface of the water. The formation of perturbing capillary waves can be suppressed by putting a very thin layer of fatty-acid molecules onto the liquid surface; usually, it is sufficient to dip one's hand into the water[C12.18]. The shape of the waves can be observed through windows on the sides of the tank. If no windows are available, one could observe the mirror images of a long fluorescent tube reflected in the water surface. These were photographed to obtain Fig. 12.79.
We can consider a similar experiment on a larger scale: In Fig. 12.79, imagine that the zero point at the upper left is a storm zone near Cape Horn. Then when the Atlantic is free of storms, a ground swell of a few centimeters amplitude propagates as far as the south coast of England. At first, after several days, waves of lengths around several 100 m and periods of about 20 s appear. The waves which follow later have gradually decreasing wavelengths and periods. For more details, see e.g. https://en.m.wikipedia.org/wiki/Swell_(ocean)#.

C12.18. A very well-formed (circular) wave train has been observed with gravity waves of wavelengths 5–20 m and used to verify Eq. (12.32). (E. Mollwo, *Optik* **4**, 450 (1948/49)).

These observations are very instructive; they show that dispersion alone suffices to *convert* an aperiodic process (in optics, for example "incandescent light") into periodic waves (in optics "monochromatic light"). It is not at all necessary to give the dispersive medium (e.g. glass) a particular geometric form (e.g. as a prism; cf. Vol. 2, Sect. 16.10). Dispersion can however never produce strictly sinusoidal ("monochromatic") waves. Even short wave trains from long wave groups produced by dispersion consist of waves from a range between λ and $(\lambda + d\lambda)$. In their spectra, they are represented by a narrow band, as shown in Fig. 12.78b. Therefore, spectral appa-

Figure 12.81 "Aged" capillary waves travelling to the right on a water surface; at the left side of the group, the wavelength is approximately 1.7 cm

ratus that are based on dispersion, for example prism spectrographs, have a limited resolving power, $\lambda/\mathrm{d}\lambda$.

For each short wave train from a long wave group produced by dispersion, we can specify a group velocity c^*. For the case of gravity waves, c^* is always less than their phase velocity c at all wavelengths, but c^* increases along the direction of propagation of the group. For this reason, the wave group in Fig. 12.79 becomes longer and longer, the further it moves away from its source.

The dispersion of the capillary waves can be treated in a similar manner, but now, $\mathrm{d}c/\mathrm{d}\lambda$ is negative, so that the dispersion is "anomalous": the short waves travel faster than the long waves (Fig. 12.17). At the front of a wave group, new crests are formed. At first, we can discern around 15 wave crests; but the shorter waves are damped much more strongly by frictional effects in the surface layers as they propagate through a wavelength than the longer waves (cf. Sect. 9.5, point 5). Therefore, the short waves at the front of the group decay more quickly. After a propagation time of around 2.5 s, an *aged* group exhibits the form shown as a photographic image in Fig. 12.81. At the left, waves of about 1.7 cm wavelength remain. Explanation: At a wavelength of $\lambda = 1.73$ cm, the curve of the phase velocity c has its minimum (Fig. 12.74, top). In the region of this minimum, $\mathrm{d}c/\mathrm{d}\lambda \approx 0$; thus, the group velocity c^* and the phase velocity c are practically identical; the aged group in Fig. 12.81 can then continue on its way without any further noticeable change in its shape[C12.19].

All of the observations described in this section can be conveniently experienced on the smooth water surface of a pond. There, one frequently sees a very noticeable phenomenon: If a small object (a stick or a fishing line) is moved relative to the water surface within its plane, we can see *standing* waves in front of it, i.e. waves moving at the same speed as the object. They occur just when the relative velocity $u > 23$ cm/s, i.e. when it exceeds the minimum value of the phase velocity in Fig. 12.74. Then, the slowest waves can no longer 'run ahead' of the moving object.

C12.19. Such an "aged" wave group of capillary waves can be seen in **Video 12.2** (at 2:30 min.); see also Fig. 12.17.

Video 12.2:
"Experiments with water waves."
http://tiny.cc/tfgvjy

"All of the observations described in this section can be conveniently experienced on the smooth water surface of a pond."

12.24 The Energy of a Sound Field. The Wave Resistance for Sound Waves

The energy density δ of a sound field is defined by the quotient

$$\delta = \frac{\text{Energy of oscillation within the volume } V}{\text{Volume } V}. \qquad (12.39)$$

The waves are presumed to be weakly divergent, i.e. they form a beam with a small divergence (opening angle[10]), but are practically still plane waves, which strike a surface A perpendicular to it. Then within a time t, they convey to this surface an energy

$$E = \delta A c t \,, \tag{12.40}$$

i.e. the entire energy which was previously contained in the volume Act. The area A is "irradiated". Its *irradiance* is defined by the ratio

$$E_\mathrm{e} = \frac{\text{Incident radiation power}}{\text{Irradiated surface area}} \,, \tag{12.41}$$

thus

$$E_\mathrm{e} = \frac{E}{tA} = \frac{\delta A c t}{tA} = \delta c \,.$$

From this, we obtain for the irradiance the important relation:

$$E_\mathrm{e} = \delta c \,. \tag{12.42}$$

Its units are for example $\mathrm{W/m^2}$.

The oscillation energy in the sound field is the sum of the oscillation energies of all the volume elements of the air that are vibrating along the direction of propagation of the sound waves. The energy of each sinusoidal oscillation can be computed either as the maximum value of its potential energy or the maximum of its kinetic energy. Think of a simple pendulum: At its maximum deflection, all of its oscillation energy is in the form of potential energy; when it is passing through its rest position, its energy is all kinetic. At all intermediate positions, the total energy is composed of both potential *and* kinetic energy. The same is true for sinusoidal sound waves.

The maximum velocity of the air particles, i.e. the *velocity amplitude* (technically: their "rapidity"), will be denoted by u_0. The greatest deviation of the air pressure from its value in air at rest, i.e. the *pressure amplitude* of the sound waves, will be called Δp_0 (known technically as the *sound pressure amplitude* or *sound pressure level*, SPL). Then a quantity of air of volume V and density ϱ contains the kinetic energy

$$E_\mathrm{kin} = \frac{1}{2}\varrho V u_0^2$$

and its acoustic energy density[C12.20] is

C12.20. Note: δ is equal to the sound pressure (acoustic radiation pressure) p_S; cf. Eq. (12.29).

$$\delta = \frac{1}{2}\varrho u_0^2 \,. \tag{12.43}$$

Starting from the potential energy, after a brief computation we obtain an expression for the acoustic energy density:

$$\delta = \frac{1}{2}\frac{(\Delta p_0)^2}{c^2 \varrho} \,. \tag{12.44}$$

[10] The details of divergence, radiation power and related concepts can be found in Vol. 2, Chap. 19.

Derivation
Beginning with Eq. (5.8), we find $E_{\text{pot}} = \frac{1}{2}\Delta V \Delta p_0$, and the energy density

$$\delta = \frac{1}{2}\frac{\Delta V}{V}\Delta p_0\,. \tag{12.45}$$

We use the modulus of compression of a gas,

$$K = V\frac{\Delta p_0}{\Delta V}$$

and the velocity of sound[C12.21]

$$c = \sqrt{\frac{K}{\varrho}}. \tag{12.46}$$

Combining these equations yields Eq. (12.44).

C12.21. Derived in an analogous way as for a solid rod (Sect. 12.17); however, the modulus of elasticity E of the solid material must be replaced by the modulus of compressibility of the gas.

During each sinusoidal oscillation, the velocity amplitude u_0 and the maximum deflection x_0 are related to each other by the equation

$$u_0 = \omega x_0 \tag{4.12}$$

$$(\omega = 2\pi\nu = \text{circular frequency})\,.$$

We thus obtain a third expression for the acoustic energy density, this time containing the *frequency* of the sound waves:

$$\delta = \frac{1}{2}\varrho\,\omega^2 x_0^2\,. \tag{12.47}$$

The above equations are by no means applicable only to air; they apply to every medium that carries sound waves. The energy density increases as the square of the frequency. As a result, high-frequency sound waves give rise to striking phenomena, for example cavitation in liquids. In Fig. 12.82a, standing acoustic waves ($\lambda \approx 1$ cm) are produced by reflection from a ground-glass plate serving as a mirror. We place a soap film in the region of the sound waves. It cuts the wave crests and troughs into strips perpendicular to the plane of the page. In the crests, the film becomes cloudy due to cavitation, i.e. the segregation of gas bubbles. This can be projected as a shadow image onto the glass plate (Fig. 12.82b).

All three characteristic quantities of air vibrations, namely the amplitudes u_0 of the *velocity*, Δp_0 of the *pressure changes* and x_0 of the *deflection* can be measured directly.

1. *The measurement of the velocity amplitude* u_0 is carried out by making use of hydrodynamic forces.

Example
The RAYLEIGH disk. In the sound field, a thin disk the size of a coin is hung suspended, free to rotate. It carries a mirror for a light pointer and is protected from air currents by a small gauze cage. Let the surface normal of the disk be inclined at an angle ϑ of ca. 45° relative to the direction of

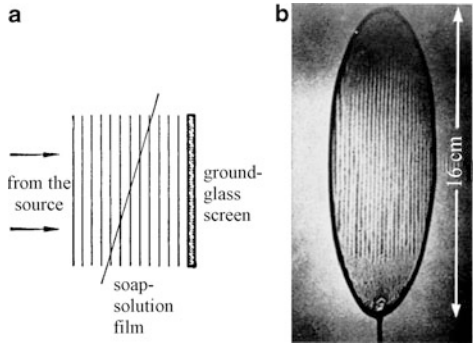


Figure 12.82 Sound waves of high energy density cause cavitation in a soap film, which is cut at an angle by a field of high-frequency standing sound waves ($\lambda = 1$ cm; the source is the same as in Fig. 12.45)

propagation of the sound waves. The airflow of the waves passes around the disk with the well-known streamline pattern as shown in Fig. 10.16b. The disk experiences a torque[C12.22] of

C12.22. This torque is maximal when the disk is inclined at an angle of 45° relative to the undisturbed airflow.

$$M = \frac{4}{3} \delta\, r^3 \sin 2\vartheta \tag{12.48}$$

(r = Radius of the disk, δ = energy density of the sound waves).

2. To measure the *pressure amplitude* Δp_0, one generally uses a capacitive microphone.

3. To measure the *maximum deflection* x_0, one places tiny dust particles into the sound field and observes their pendulum orbits with a microscope. The small spheres are pulled along by the internal friction (viscosity) of the gas (Sect. 10.3). They exhibit nearly the same amplitudes (maximum deflections) as the surrounding volume elements of the air. However, this method is applicable only at high energy densities δ.

We combine Eq. (12.46) with Eqns. (12.43) and (12.44) to obtain

$$\frac{\Delta p_0}{u_0} = c\varrho = \sqrt{K\varrho}\,. \tag{12.49}$$

This ratio of the pressure amplitude to the velocity amplitude is called the *acoustic wave resistance Z* (compare Vol. 2, Sect. 12.7.)

This wave resistance determines the reflection at the interface between two media. If a plane wave is incident perpendicular to the interface with a medium having a different wave resistance, then the ratio R is[C12.23]

C12.23. The acoustic wave resistance $Z = c\varrho$ corresponds to the wave resistance $Z = \sqrt{\mu_0/(\epsilon\epsilon_0)}$ for electromagnetic waves (Vol. 2, Sect. 12.7). The reflection coefficient derived from FRESNEL's formulas (Vol. 2, Sect. 25.8) is $R = ((Z_1 - Z_2)/(Z_1 + Z_2))^2$. This yields Eq. (12.50).

$$R = \frac{\text{Reflected radiation power}}{\text{Incident radiation power}} = \left(\frac{c_1\varrho_1 - c_2\varrho_2}{c_1\varrho_1 + c_2\varrho_2}\right)^2 . \tag{12.50}$$

It is called the *reflection coefficient* or *reflectivity*. The reflected and the incident waves superpose to give a *resultant wave*.

In the technical and acoustics literature, for a sinusoidal sound wave of radiation power $\dot{W}_1$ or pressure amplitude Δp_1, the absolute values are often not quoted, i.e. for example $\dot{W}_1$ in watt or Δp_1 in newton/meter^2, but instead only the relative values. These are *compared* to a reference power $\dot{W}_2$ or a reference amplitude Δp_2, which must be *precisely specified in every case*. Then, either the quantity

$$x = 10 \cdot \log \frac{\dot{W}_1}{\dot{W}_2} = 20 \cdot \log \frac{\Delta p_1}{\Delta p_2} \tag{12.51}$$

is computed, or else the quantity

$$y = \frac{1}{2} \ln \frac{\dot{W}_1}{\dot{W}_2} = \ln \frac{\Delta p_1}{\Delta p_2} \,. \tag{12.52}$$

Both of these quantities are pure numbers. They are combined with the number 1 as multiplier and we give the number 1 two new names, namely in the first case *decibel* (dB), and in the second case *Neper* (Np). – If for example we use a reference pressure of $\Delta p_2 = 1\,\mathrm{N/m^2}$ ($= 1\,\mathrm{Pa}$), then the specification "–60 dB" implies a pressure amplitude of $\Delta p_1 = 10^{-3}\,\mathrm{N/m^2}$. These special names for the number 1 have an advantage: They remind us which equation for the power or pressure comparison has been employed. Their great disadvantage, however, is that one can interpret decibel or Neper (and, later, the *phon*) incorrectly to be *units*, similar to the ampere, kilogram, candela etc.
In technology, the (effective) reference pressure is often taken to be $\Delta p_2 = 2 \cdot 10^{-5}\,\mathrm{N/m^2}$. Then the quantities x and y, i.e. the logarithms of the *relative* measured pressures, are quoted as *absolute* sound-pressure levels! Often, decibels are also used to denote *orders of magnitude*, in contrast to the usual form. Then, for example, $80\,\mathrm{dB} = 10^8$, $20\,\mathrm{dB} = 10^2$, $-30\,\mathrm{dB} = 10^{-3}$ etc.

12.25 Sound Sources

In the ripple tank, we could get a good intuitive picture of the mechanism of wave emission. We could see how the excitation bar dips rhythmically into the water surface and displaces some water at the frequency of its vertical vibrations. This experiment can be applied in analogous fashion to the spatial emission of longitudinal elastic waves in air, water, etc. If a sphere changes its volume in the rhythm of sinusoidal oscillations, one is dealing with an *ideal sound emitter*, the *breathing sphere*. All the points on its surface vibrate with the same phase, and the result is a completely symmetric emission of spherical waves. This ideal sound emitter has yet to be technically implemented. But some solutions to this problem approach the ideal rather closely. Thick-walled containers with a vibrating membrane are good examples in this connection. The membrane is best driven *electromagnetically* from within the container. Applying this principle to acoustic signals in water, using membranes of around 50 cm diameter, power levels of the emitted sound waves in water of up to 0.5 kW have been achieved. The "membrane" or diaphragm is in this example a steel plate of ca. 2 cm thickness!

Figure 12.83 The sound emission of a vibrating string

In the *simplest* vibrational mode, the membrane of a sound source oscillates with the same phase all over its surface, and exhibits no nodal lines except at its rim. Furthermore, as a rough approximation, we will consider its amplitude to be constant over the whole surface area. Then we have physically quite similar conditions to the constant-phase emission of waves from the aperture of a slit in Fig. 12.12. We can thus consider the emission of the waves into a spatial cone under the right circumstances, similar to what is shown in Fig. 12.14. In this case, the diameter of the membrane must be a multiple of the emitted wavelength.

Passable sound sources are also the open ends of short, thick oscillating air columns. *Very poor emitters* are, in contrast, the *strings* which are often employed in musical instruments.

In Fig. 12.83, the black disk represents the cross-section of a vibrating string which is perpendicular to the plane of the page. The string is just beginning a vibration in the direction of the arrow, downwards. It then “displaces”, roughly speaking, the air on its lower side, and there, a wave train begins with a crest. At the same time, the string leaves, again roughly speaking, an empty space on its upper side, and there, a wave train begins with a trough. The two waves have a phase difference of practically 180° in every direction and cancel each other almost completely by interference. Therefore, the vibrating string is a very poor emitter of sound waves. Similar considerations apply to *tuning forks*.

For practical applications, the vibrations of strings and of tuning forks must therefore first be transferred to good emitters. To this end, one arranges a suitable mechanical connection between the strings or tuning fork and a good emitter, which is thus excited to forced vibrations. Under some circumstances, one can make use of the special case of resonance to obtain large amplitudes. The emitter is then set up to have weak damping and its resonance frequency is matched to that of the tuning fork or string. To illustrate this, we offer two examples:

1. In Fig. 12.84a, a piece of twine is held at the right by a hand. Two fingers of the other hand rub its left end. The twine begins to vibrate like a violin string, but it emits practically no sound. Then we attach its right end to a good emitter, e.g. a short metal or cardboard cannister with a membrane (Fig. 12.84b). Now the vibrations are emitted as sound waves which can be clearly heard at some distance.

“One often hears that ‘The vibrations were amplified by resonance’. This is a rather skewed way of expressing the situation.”

2. A tuning fork is mounted on a short wooden box which is open at one end, usually called a resonance chamber or sound box. One often hears that “The vibrations were amplified by resonance”. This is a rather skewed way of expressing the situation. The essential point is only the relatively good emission properties of the box; the

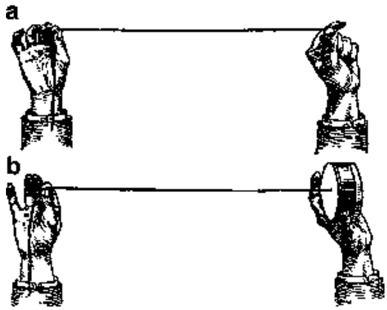


Figure 12.84 Coupling of a poorly-emitting string to a membrane which is a good emitter

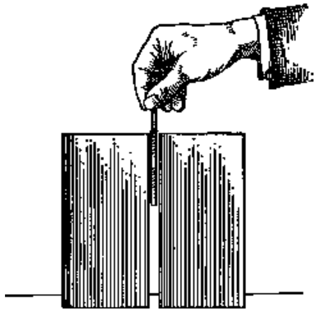


Figure 12.85 Improvement of the sound emission from a tuning fork by a wall on each side (W. BURSTYN)[C12.24]

C12.24. W. BURSTYN constructed among other things electric musical instruments at the beginning of the 20th century.

resonance is simply an aid to transferring the vibrations. We can demonstrate this with a surprising experiment. We bring one tine of a tuning fork as shown in Fig. 12.85 into the gap between two walls which are large in comparison to the wavelength of its tone. The tuning fork can now be heard at some distance away. This is because the interference of the waves from the inner and outer sides of the tine is now considerably reduced and the tuning fork has thus been made into a tolerably good emitter.

In the case of musical instruments, e.g. those of the violin family, the situation is rather complex. The strings and the sound box form a confusingly coupled system (Sect. 11.15). The box itself has a whole series of resonance frequencies. In producing its forced vibrations, certain frequencies of the vibrating strings are thus strongly favored. Figure 12.86 shows a vibration curve and the associated vibration spectrum of a tone from a violin. The sound box of a violin is stiffened from within by the sound post. The top plates, considered as membranes, are by no means small compared to all the wavelengths used in music. Therefore, the emission can exhibit strongly-preferred directions.

With the topic of violin characteristics, we come to a discrimination which is of some importance technically, between *primary* and *secondary sound sources*. Primary sound sources (emitters) have to produce vibrations with a certain spectral composition. We condone the fact that every individual primary sound source, for example every musical instrument, has a right to its characteristic spectrum; or, stated physiologically, to its own particular timbre. The situation is quite different for secondary emitters. A typical modern represen-

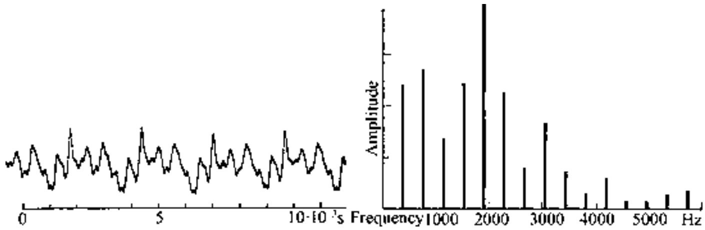


Figure 12.86 The vibration curve (at *left*) and the frequency spectrum (*right*) of a violin tone (H. BACKHAUS)[C12.25]

C12.25. The frequency spectrum (FOURIER analysis, see Sect. 11.3) of the vibration curve shows the amplitudes of the fundamental tone and the harmonics, referred to as partial tones, of the g_1 note of the d-string of an Antonius Stradivarius violin dating from 1707. The frequency of the fundamental is 392 Hz. The harmonics determine the timbre, which can vary considerably from instrument to instrument. (After Fig. 11 from H. Backhaus, *Die Naturwissenschaften* **17**, 811 (1929).)

tative of this latter group is the loudspeaker. Loudspeakers have no free choice in their frequency spectra. They are required to emit – in the form of sound waves – the electrical oscillations which are sent to them, without giving preference to any partial tone components.

12.26 Aperiodic Sound Sources and Supersonic Velocities

In Sect. 12.10, we encountered shockwaves and conical waves. They are generated when an object moves aperiodically and its velocity exceeds that of the sound waves (or other waves). This case often occurs in the air. As examples, we could mention the end of a whiplash, a bullet, or an aircraft flying at supersonic speed. These rapidly moving bodies generate conical waves[11]. Figure 12.87 shows such a conical wave from a bullet, just at the moment when it has overtaken the muzzle report ('bang') of the gun. The opening angle of the sound wave cone is called MACH's angle (Sect. 12.11). A single conical wave is heard by the ear as a *'boom'* or *'bang'*; a periodic series of these waves sounds like the tone from a trombone.

In the acoustic investigation of bullets, it must be taken into account that sound waves of large amplitudes can propagate at higher velocities than

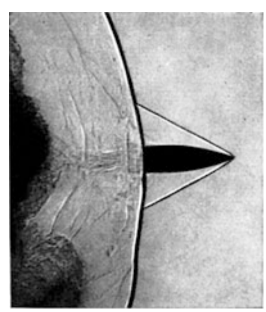


Figure 12.87 The muzzle report of a gun and the conical wave from a bullet (the black cloud at left is gas from the powder explosion). (This and the following picture are photos taken by C. CRANZ using the schlieren method.)

[11] Journalists will then report that "the sound barrier has been broken".

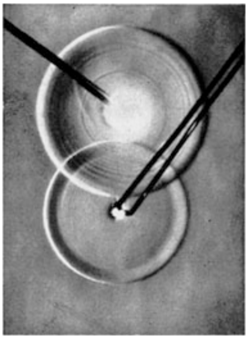


Figure 12.88 Shock waves produced by two *simultaneous* sparks with different electric currents (larger at the *top*)

the normal speed of sound. Near electric sparks, one can readily observe sound waves whose velocities are around 500 m/s (Fig. 12.88).

12.27 Sound Receivers

Sound receivers can be divided into two groups in the sense of *limiting cases*: Pressure receivers and velocity receivers.

1. *Pressure receivers*: The majority of pressure receivers consist of membranes fixed at their edges. They can be attached within capsules, walls, funnels etc. Examples: Microphones of all types, as well as the eardrum.

All pressure receivers carry out *forced vibrations* in the sound field. Their amplitudes are independent of their orientation within the sound field; the air pressure is independent of direction (pressure isotropy; cf. Sect. 9.3). Every living-room barometer demonstrates this. A barometer is in the end just a pressure receiver for longitudinal waves in the air. But in the case of atmospheric pressure fluctuations, we are usually dealing with oscillation processes of extremely low frequency.

Technically, microphones are much more important today than all other types of pressure receivers. Radio broadcasting and sound recording have raised the standards here enormously. Over a broad frequency range from around 100 up to 10 000 Hz, they are required to maintain the original amplitude ratios. As with all forced oscillations, this requirement can be bought only at the price of a greatly reduced sensitivity.

2. *Velocity receivers*. In the case of velocity receivers, the velocity amplitude of the alternating air flow of a sound wave is used to produce forced oscillations. This can be most readily made clear by giving an experimental example.

In Fig. 12.89, we see a thin glass fiber about 8 mm long which is set up as a small leaf spring perpendicular to the direction of the

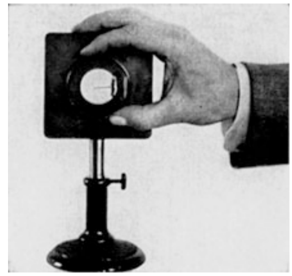


Figure 12.89 A fine glass fiber as a detector of motion (velocity receiver) (in reality, it is only 0.028 mm thick)

propagating sound waves (microprojection!). Periodic changes in the air pressure have no influence at all on this fiber. On the contrary, the alternating air flow in the direction of the oscillating air particles pulls it along by internal friction and thereby excites it to forced vibrations (here using a reed pipe as sound source, close to the fiber). This fiber is a typical velocity receiver. It at the same time exhibits an important and – for *velocity receivers* – characteristic property: We find that *its amplitude depends on its orientation in the sound field*. If it is placed with its long axis parallel to the direction of sound propagation, the fiber remains at rest.

Velocity receivers can be used as *directional receivers*. Imagine two such fibers oriented symmetrically to the long axis on each side of a moving body. If there is a straight-line course to the sound source, both receivers will answer with the same amplitude. Sideways deviations from the straight course will give rise to inequalities in the forced amplitudes of the two fibers.

Pressure and velocity receivers are, as mentioned, limiting cases. Every momentum transfer from pressure requires a wall which is not noticeably deformed by the pressure. The forced amplitudes of the wall must remain small compared to the deflections x_0 of the oscillating air or water particles. Air has a low density ϱ and therefore larger deflections x_0 (compare Eq. (12.47)). Therefore, pressure receivers can be constructed successfully for air, but only with difficulty for sound waves in water. Pressure receivers in air can also become velocity receivers in water.

12.28 The Sense of Hearing

Hearing and our auditory organs are for the most part the objects of physiological and psychological research. Nevertheless, for physical purposes, we want to summarize the most important facts concerning them. Likewise, in optics one should know the properties of the eye, at least in general terms.

1. Our ears react to mechanical vibrations in the very broad frequency range from about 20 Hz up to around 20 000 Hz. The human ear thus encompasses a spectral range of at least 10 octaves ($2^{10} = 1024$).

The upper limit of this range decreases with increasing age of the hearer.

2. Sinusoidal vibrations produce a sensation of a pure *note* in the ear. Every note has its particular pitch. The pitch of a note is a sensory perception and as such is not accessible to physical measurements. The pitch of a note depends in the main on the frequency of the waves, but to some extent also on the strength of irradiation at the ear. Unfortunately, we generally speak of the frequency of a note. This is of course convenient, but it is a rather lax form of expression; what is meant is always the *pitch* of the note as experienced by the ear when it is excited by a sinusoidal wave of the given frequency at a medium irradiation strength.

3. In the most favorable frequency range, our ears can distinguish two frequencies that differ by only 0.3 %. The ear thus has a "spectral resolving power" $\nu/\mathrm{d}\nu$ of around 300. This corresponds in optics to what one can achieve with a glass prism of a basal thickness of about 1 cm.

4. The ear responds to non-sinusoidal vibrations by perceiving a *tone*. *A tone is independent of the phase differences between the individual sinusoidal partial tones.* This is a fundamental insight based on the observations of GEORG SIMON OHM. It corresponds to HELMHOLTZ's interpretation that the inner ear functions as a spectral apparatus for sound waves (cf. Sect. 12.30).

A musical chord corresponds to a frequency spectrum with a particular structure, characterized by the ratio of the frequencies and amplitudes of its spectral lines. The absolute value of the fundamental frequency is unimportant. Two sinusoidal oscillations of nearly the same energy density with a frequency ratio of 1:2 always yield the chord called an "octave", etc.

5. The most important tones are the *syllables* of language. – In normal reading, the eye is stimulated by a temporal sequence of two-dimensional symbols. They are distinguished by a spatial sequence of individual elements, namely letters or syllables. Often, we can 'hear' the writer speaking when we read a text written by someone we know. In normal hearing, however, the ear is stimulated by a temporal series of air-pressure variations. These have spectra of differing forms. Voiceless consonants are characterized by broad, continuous spectra with a variety of shapes. Voiced consonants and vowels can be represented as frequency spectra. One finds the same spectral lines independently of the pitch of the voice (bass, tenor etc.) in regions which are characteristic of the particular vowel. These are called the *formant ranges* (Fig. 12.90). They are finally just the damped normal modes of the oral cavity etc., which are periodically excited through impulses of air pressure from the larynx along the lines of Fig. 11.16. In general, the frequency and intensity of this excitation is constantly changing. These changes determine the *intonation* of the speaker. If in contrast the excitation is continuous, with a constant frequency and intensity, we experience the solemn intonation of a priest at prayer.

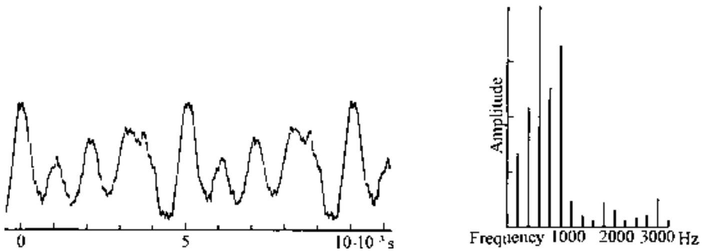


Figure 12.90 *Left*: Vibration curve of the vowel ‘a’ as sung by a male voice (recording by FERDINAND TRENDELENBURG). The impulse frequency of the larynx was 200 Hz. *Right*: Representation of this vibration curve in the form of a frequency spectrum. The principal formant range is clearly visible.

The time sequence of the language elements, i.e. their spectra, can be continually recorded on tape (Fig. 12.91). This yields a script whose spatial sequence of elements consists of spectra. This script can be read, much like e.g. Morse code, but only after considerable practice.

Spoken language can be reproduced by mechanical means. Girls play with dolls which have a bellows in their torsos and can say ‘mama’ and ‘papa’. – A speaking machine was described in 1791 by WOLFGANG V. KEMPELEN[12]. It contains apparatus to produce tones and hissing sounds which could be activated by openings and keys operated by the hands and fingers. Newer versions of such speaking machines make use of electronic devices.

6. At higher energy densities, the ear can hear *difference frequencies*. It hears not only two notes of frequencies ν_2 and ν_1, but also a third note of frequency $\nu_2 - \nu_1$. Difference frequencies can be readily demonstrated using organ pipes. Occasionally, still other “combination frequencies” can be heard, e.g. the “sum frequency” $(\nu_1 + \nu_2)$ or the frequency $(2\nu_1 - \nu_2)$ (cf. Sect. 11.12, point 2).

7. The rise and decay times of the ear are not well known; they appear to be of the order of some 10^{-2} s.

8. With two ears, we can recognize the *direction* from which sound waves are arriving. This works best for tones and noises with a sharp onset or with repetitions of characteristic details. The decisive factor is the time difference between the excitation of the left ear and of the right ear by the same portion of the sound wave curve (compare Fig. 12.92). At frequencies of several 1000 Hz, differences in the irradiation power due to the shadowing effect of the head also play a role (cf. Sect. 12.19).

[12] His book on “The Mechanism of Human Speech along with a Description of a Speaking Machine” was published in Vienna in 1791 by J.V. Degen Publishers.

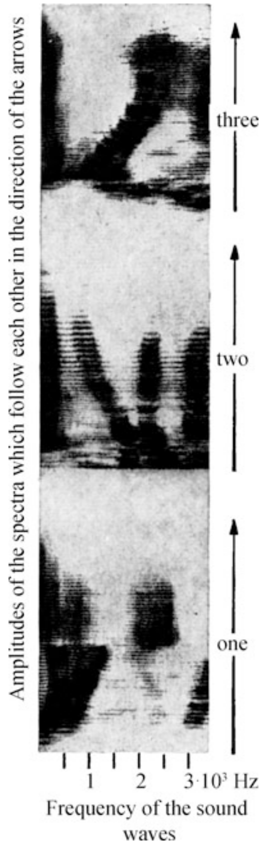


Figure 12.91 A time sequence of a number of spectra (around 200) corresponding to counting from 1 to 3 in English. The spectra, which follow each other from bottom to top in a dense series, each have a height of only a few tenths of a millimeter in the image. The amplitudes are not represented as in Fig. 12.90 (right side) in terms of ordinate values of different heights, but rather as the darkness (density) of a photographic plate.

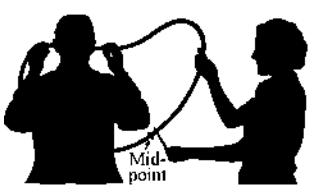


Figure 12.92 *Directional hearing* Hold the two ends of a piece of hose around 2 m long in your ears and let an assistant tap on the hose near its midpoint. The apparent direction of the sound deviates from the medial plane of the head when the position of the tapping is more than 0.5 cm from the midpoint of the hose. Our sense of hearing thus reacts to difference in the arrival times of sounds of $\Delta t = 3 \cdot 10^{-5}$ s. At $\Delta t = 60 \cdot 10^{-5}$ s (corresponding to 20 cm path difference, roughly the diameter of the head!), we localize the apparent sound source at right angles to the medial plane.[C12.26].

C12.26. At the left: E. MOLLWO, Dr. rer. nat. (Göttingen 1933); at the right: H. KELTING, Dr. rer. nat. (Göttingen 1944).

12.29 Phonometry

Phonometry in acoustics corresponds to photometry in optics. Both evaluate a type of physical radiation not according to its power (energy/time = energy current, measured in watt), but rather according to its effect on our *sensory organs*, i.e. on our ears and eyes. *Just as photometry holds only for observers with normal sight* (compare Vol. 2, Chap. 29), *phonometry applies only to observers with normal hearing*.

All sound perceptions, i.e. notes, tones, noises etc., have, besides their quality of 'pitch' (high, low, hollow, brilliant etc.) a second quality, their *loudness* (or sound level). It corresponds to the *brightness* in the perception of light by the eyes. The loudness cannot be physically measured, i.e. it cannot be determined as a multiple of some physical unit, no more than brightness or any other perceptional quality of our sensory organs. However, with our ears, we can *perceive* two *irradiance strengths* (i.e. the incident radiation power $\dot{W}$/receiver area A), even from quite different *sound* sources, as being equally loud.

As is well known, the eye is capable of a corresponding comparison. It can *perceive* two irradiance strengths of light, even when they originate from quite different types of *light* sources, to be equally bright. The irradiance strengths as judged by the *eye* are termed *illumination* levels (cf. Vol. 2, Chap. 29).

This ability of our sensory organs forms the basis for setting up a phonometry (sound measurement) and a photometry (light measurement) for technical, practical purposes. Both make it possible to determine two spatially or temporally separated irradiance strengths

$$I = \frac{\text{Incident radiation power } \dot{W}}{\text{Receiver area } A}; \quad \text{unit: W/m}^2 \qquad (12.53)$$

so that they can be reproduced at an arbitrary place and time.

In *phono*metry, one does not try to quote the irradiance strength (intensity) I in absolute units, e.g. in W/m^2, but rather as a relative quantity in multiples of an agreed-upon, small *reference* irradiance strength $I_{\min} = 2 \cdot 10^{-12}\,\text{W/m}^2$. (This corresponds roughly to the detection threshold of the human ear at the frequency $\nu = 10^3\,\text{Hz}$, typically for a younger person, ca. 20 years old). The ratio $I/I_{\min}$ in practice would yield numbers between 1 and 10^{12}. This large range is made more compact by using decimal logarithms. The *loudness* L is defined as

$$L = 10 \cdot \log \frac{I}{I_{\min}} \qquad (12.54)$$

($I_{\min}$ depends on the frequency as seen in Fig. 12.93; see below).

The loudness is thus a pure number. One combines it with the number 1 as multiplier and gives this number the name *phon* (cf. the end of Sect. 12.24[C12.27]). One thus says for example that the loudness of everyday speech corresponds to 50 phon. This means that everyday speech has an irradiance strength as perceived by the ear of $I = 10^5 \cdot I_{\min} = 2 \cdot 10^{-7}\,\mathrm{W/m^2}$ $(10 \cdot \log 10^5 = 50)$. – The range of loudness stretches from 0 phon (the level of the barely perceptible) up to 120 phon (the noise level in a boiler factory or next to an airplane, sometimes termed the threshold of pain for the ears).

C12.27. The unit *phon* is not included in the SI (cf. Comment C2.14.). It is a unit of the subjective quality "loudness", used e.g. by the American National Standards Institute and the international standards organization ISO (see http://www.iso.org/iso/home.html). It is defined as "the dB sound pressure level (see Comment C12.28.) of a 1 kHz tone which sounds just as loud".

From the defining equation (12.54), it follows that if the irradiance strength as perceived by our ears is multiplied by a factor of 10, then the loudness increases by 10 phon, since $10 \cdot \log 10 = 10 \cdot 1 = 10$. – Examples: *One* very softly ticking clock has a loudness of 10 phon, ten such clocks together have $10 + 10 = 20$ phon. – *One* roaring motorcycle has a loudness of 90 phon, ten of them together have $90 + 10 = 100$ phon.

The *spectral sensitivity distribution* of the ear is illustrated by Fig. 12.93. The ordinate values show both the *sound pressure amplitude* Δp_0 and also the irradiance strength I. The values of the irradiance strength or intensity apply to a cross-sectional area of a free sound field, not disturbed by a human head, for a wave which is incident perpendicular to the face of the hearer. With increasing loudness, the spectral sensitivity distribution changes; the curves become flatter. With a further increase, instead of hearing, the observer feels pain. In the range of their greatest sensitivity, our ears react to

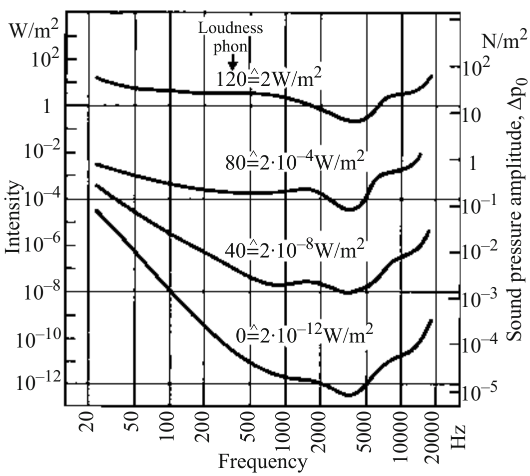


Figure 12.93 The curves of spectral sensitivity of the ear for different values of the loudness. Along each curve, the loudness is perceived as the same as for a pure comparison tone ($\nu = 1000\,\mathrm{Hz}$) which is produced with an irradiance strength (or sound pressure level, i.e. a physical stimulus!) corresponding to the given loudness. The bottom curve is the detection threshold of the ear of a young person (20 years old). With increasing age, the curves rise more sharply above 2000 Hz[C12.28].

C12.28. Today, loudness is usually quoted in dB (Eq. (12.51)). For this measure, instead of the individual curves as in Fig. 12.93, which depend on the age of the hearer, fixed conventional frequency dependencies (so-called 'A-weighting', 'B-weighting', etc.) are used, and they are indicated by adding corresponding letters, e.g. dB(A). For more details on A-weighting, see for example https://en.wikipedia.org/wiki/A-weighting changes in the air pressure of $\Delta p_0 = 10^{-5}\,\mathrm{N/m^2}$, i.e. a value which is 10^{-10} times smaller than the normal air pressure!

Figure 12.93 shows vividly just how astonishingly adaptable our ears are: In the frequency range of its greatest sensitivity, the ear – just like the eye – can deal with variations in the irradiance strength I over a range of $1 - 10^{12}$. Both of these sensory organs behave, put succinctly, as though they were measuring instruments with logarithmic scales. Both are wonderfully adapted to their respective purposes: The ear for example can hear nearly as well in a region of weak, diffracted, reflected or scattered sound waves as when the waves are freely propagating.

12.30 The Human Ear

The most essential part of our auditory canal is the "inner ear", which is a bony labyrinth containing a spiral organ (the cochlea) lodged in the petrous bone. The mechanical (sound) waves are directed to it via two paths: 1. through the eardrum (tympanum) and the bones of the middle ear (ossicles); and 2. through the soft tissue and the bones of the head. The first path is dispensable; one can still hear without an eardrum and the ossicles.

The eardrum and the ossicles fulfill merely the following purpose: The inner ear is filled with a watery liquid, the endolymph, and its density is about 800 times greater than that of air. As a result, the reflectivity R between the air and the inner ear would be nearly 1.0 (cf. Eq. (12.50). Now, however, the sound wave should reach the endolymph with its given energy density δ, without hindrance and without reflection losses. This is accomplished by the eardrum and the lever system of the ossicles[13]. According to this concept, the eardrum and the ossicles would be unnecessary in mammals that live exclusively in water (dolphins and whales). Indeed, none of these animals has an outer ear. The auditory canal, the eardrum and the ossicles have degenerated to mere remnants.

Nerve excitation in the *eye* takes place in the mosaic-like retina. It reacts practically to only *one* octave of the electromagnetic spectrum. Nevertheless, we can distinguish a multitude of colorful and un-colorful hues. Furthermore, the acuity of the images that we see is not compatible with the quality of the lens of the eye. Both effects cannot be understood physically. They can be produced only if central processes within the brain play an important participatory role.

The nerve excitation of the inner *ear* takes place in CORTI's organ, within the cochlea. The basilar membrane is a part of this organ. To first order, it can be described as a delicate separating wall between two rigid tubes (the scala tympani and the scala vestibuli). In

[13] At the hearing threshold and the most favorable frequency, $\nu \approx 3000\,\mathrm{Hz}$, the amplitude of the motions of the eardrum is only about $6 \cdot 10^{-10}\,\mathrm{m}$; it is thus only a few atomic diameters. In the inner ear, the amplitudes are still smaller by at least a factor of 30.

humans, it is 34 mm long, and its width increases from the beginning of the tubes to their ends from 0.04 to 0.5 mm. HELMHOLTZ interpreted this membrane as a tiny vibrating-reed frequency meter, a *spectral apparatus* which is characterized by its particular simplicity (Sect. 11.12). He thought of the membrane not as a homogeneous band, but rather as a dense series of stretched strings, grown together at their ends (in place of the vibrating reeds or leaf springs). Their resonance frequencies were presumed to decrease along the length of the membrane, i.e. with increasing width. In reality, the basilar membrane lacks such a structure, but nevertheless it retains the properties of a spectral apparatus. Its function can be demonstrated using a model which was designed for that purpose. It fulfills the necessary condition for demonstration models: The processes to be observed take place sufficiently slowly.

The model is sketched in Fig. 12.94. Two metal channels with a rectangular cross-section are fitted at their sides with glass walls, and between them in the middle is a highly elastic partition, the artificial basilar membrane. Its width is delimited by a wedge-shaped metal frame. The width of this frame increases from left to right. For reasons of "mechanical similarity", the membrane must have elastic properties which cannot be realized with a solid body; instead, we have to make use of an interface between two liquids of differing density and surface tension, e.g. benzene above and water below (with an added salt, $MgCl_2$, to increase its viscosity). – The "stirrup" (one of the ossicles) at the lower left end can move the "oval window" back and forth in a sinusoidal motion by means of an eccentric, at frequencies between 1 Hz and 8 Hz. The vibrating oval window becomes the source of a wave group which travels to the right along the "basilar membrane".

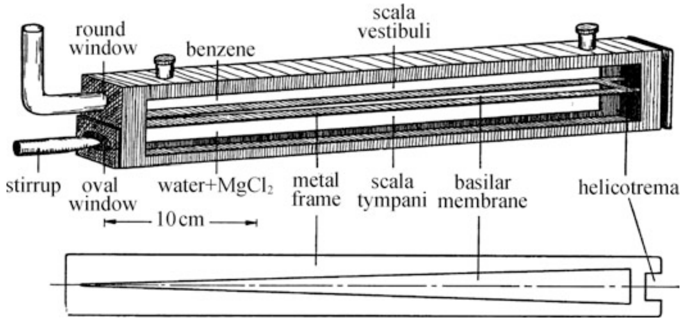


Figure 12.94 A linear model of the cochlea in the inner ear. The brass plate at the right end is removable and has a cork gasket. Instead of the "round window", a glass tube, bent upwards, is used. The artificial basilar membrane is 31 cm long and its width increases from 1 to 18 mm. The "stirrup' is caused to vibrate by an eccentric attached to an electric motor. In the human ear, the stirrup is the last of the three ossicles, the small bones which act as levers to reduce the displacement between the eardrum and the oval window[C12.29].

C12.29. Similar investigations were reported by H.G. DIESTEL (*Acustica* **4**, 489 (1954)). There, he mentions an unpublished thesis by H. DORENDORF, in which similar model experiments were described (1st Physical Institute, Göttingen (1950)). It may be that the measurements quoted here are at least partially based on this thesis (which has in the meantime been lost).

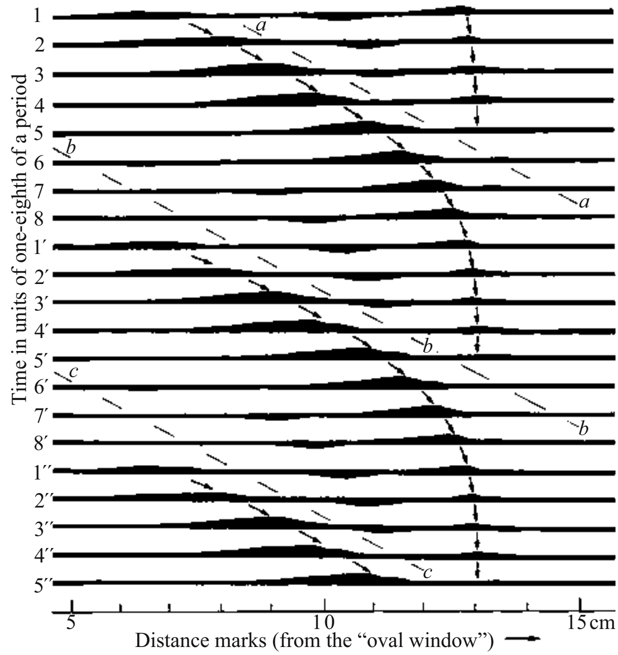


Figure 12.95 Wave groups which are travelling along the "basilar membrane" of the ear model after being sinusoidally excited at $\nu = 4\,\mathrm{Hz}$. To provide a clearer overview, the waves have been emphasized by filling with a black color; in the original photos, only their sharp outlines were visible. Wave trains without this subsequent filling-in can be seen in Fig. 12.96.

The eight upper lines in Fig. 12.95 (numbered 1 to 8) show photos ('instantaneous images') taken at equal time intervals during one oscillation period. At the right, above the dashed straight line *a*-*a*, we can see the remainder of the wave group from the previous period. From line $1'$ on, at the left of the dashed straight line *b*-*b*, the next wave group appears and follows a similar path. Below the line *c*-*c*, we see the beginning of still another wave group. – Thus, in spite of the sinusoidal excitation of the oval window, the basilar membrane does not carry a sine wave, but rather a peculiarly-shaped wave group. In this example ($\nu = 4\,\mathrm{Hz}$), it reaches its maximum amplitude at the 10 cm distance mark. Further to the right, it decays rapidly. In the process, its group velocity decreases. At the 13 cm distance mark, it has become practically zero. Further to the right, the "basilar membrane" remains completely at rest.

Now we come to a second, decisive experimental result: The position or distance mark where the wave group attains its maximum amplitude depends uniquely on the frequency of the excitation. This can be seen in the original photos in Fig. 12.96: The wedge-shaped membrane thus acts as a *spectral apparatus*; not identically to a vibrating-reed frequency meter, but in a similar manner.

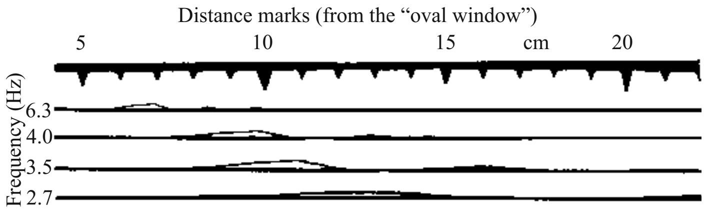


Figure 12.96 'Instantaneous images' of wave trains on the model basilar membrane at the moments when they reach their maximum amplitudes, for four different frequencies of the sinusoidal excitation. The wave groups remain similar to each other. They are simply stretched longer with decreasing frequency. The group in the second line corresponds to the group in the fourth line in Fig. 12.95; its wave crest has however not been filled in with black color in this image.

What it accomplishes can be described physically only as a *preliminary decomposition*[14]. The high resolving power which is in fact observed in human hearing cannot be understood physically. As in the case of the eye, we find also for the ear that central processes, which take place in the brain, must play a major role. These remain today outside the realm of our scientific understanding[C12.30]. The solutions to the great problems in biology and neurophysiology will probably be found only in some still-distant future.

C12.30. For further reading on the subject of hearing, see e.g. E.A. Lopez-Poveda, A.R. Palmer and R. Meddis, *The neurophysiological bases of auditory perception*. New York: Springer (2010); ISBN 978-1-4419-5685-9.

Exercises

12.1 A locomotive is approaching an observer, who is not moving, with a velocity of $u = 72\,\text{km/h}$, and emits a whistle tone of frequency $\nu = 500\,\text{Hz}$. What frequency ν' is heard by the observer? (Sect. 12.2)

12.2 A plane wave on water strikes a wall at perpendicular incidence. In the wall are two slits with a spacing a, whose width is small in comparison to the wavelength of the water wave. Describe the interference maxima at a distance $y \gg a$ for small deflection angles. (Sect. 12.3)

12.3 A planar sound wave strikes a screen that contains four small apertures. They are arranged at the corners of a square whose sides have a length a. Where are the interference maxima with the smallest distance from the symmetry axis? (Sect. 12.3).

[14] A preliminary decomposition selects from a mixture of waves of various lengths those that belong together in broad regions. One often uses "filters" for this purpose, e.g. in optics, a glass filter that passes only red light.

12.4 A planar sound wave is incident at right angles on a diffraction grating with narrow slits. On a screen at a distance a from the grating, the interference maxima have a spacing b. How large must the lattice constant D be, i.e. the spacing between two neighboring slits in the grating? (Sect. 12.15)

12.5 Determine the longitudinal sound velocity c_l in a glass rod of density $\varrho = 2.6\,\mathrm{g/cm^3}$ and modulus of elasticity $E = 6.5 \cdot 10^4\,\mathrm{N/mm^2}$. (Sects. 8.3 and 12.17)

12.6 The sound pulse of an echo-sounder travels in a time $\Delta t = 0.15\,\mathrm{s}$ from the surface of the sea to the sea bottom and back. The velocity of sound in water is $c = 1440\,\mathrm{m/s}$. Determine the depth d of the water. (Sect. 12.17)

12.7 In order to measure the velocity u of an auto that is travelling towards an observer, a radio wave of frequency ν is reflected from the auto. The frequency of the reflected wave is increased due to the DOPPLER effect and is superposed with the incident wave, so that the observer measures beats with a frequency ν_B. Determine the velocity u. The frequency is $\nu = 1 \cdot 10^9\,\mathrm{Hz}$ and $\nu_B = 56\,\mathrm{Hz}$ (Radio waves and electromagnetic waves in general are treated in detail in Vol. 2, Chap. 12. Here, we need only know that in air, they propagate at the velocity of light, $c = 3 \cdot 10^5\,\mathrm{km/s}$). In deriving the answer, it is useful to begin with the mathematical superposition of two travelling plane waves, moving in opposite directions; Comment C11.4 can be helpful for this. (Sect. 12.18)

12.8 Determine the angle α_{max} of the prism in Fig. 12.57 on which the incident sound wave undergoes total reflection, using the index of refraction which was found experimentally in Sect. 12.20. (Sect. 12.20)

12.9 A sound wave of wavelength $\lambda = 1.03\,\mathrm{cm}$ is incident on a cubic lattice with a lattice constant of $d = 3\,\mathrm{cm}$, as shown in Fig. 12.66. Find the angle under which diffraction maxima occur for the orders $m = 1, 2, 3$ and 4. (Sect. 12.20)

12.10 Two water waves of wavelengths λ and $\lambda + \mathrm{d}\lambda$ have indices of refraction n and $n + \mathrm{d}n$. Compute the group velocity c^* with which the pulse formed by superposition of the two waves propagates. (Sect. 12.22)

12.11 The frequency range of the human singing voice stretches from about 80 Hz (bass) up to 800 Hz (soprano). Find the range of wavelengths. The velocity of sound is 340 m/s. (Sect. 12.25)

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_12) contains supplementary material, which is available to authorized users.

Thermodynamics III

Fundamentals 13

13.1 Preliminary Remarks. Definition of the Concept 'Amount of Substance'

Thermodynamics is of fundamental importance for all branches of science and technology. Its most important laws are valid for all natural phenomena. Unfortunately, in contrast to the other areas of physics, its qualitative structure cannot be recognized directly from straightforward, clear-cut experiments; instead, long and tedious series of measurements are usually required. In thermodynamics, in contrast to electromagnetism, there are no perfect insulators. This often makes experimental setups complicated, and a quantitative evaluation of their results is time-consuming and tedious.

As a collective term for solid bodies and for definite amounts of liquids and gases, we will use the concept *amount of substance*. (This rather awkward term is part of the International System of Units (SI). In German, it is more compactly expressed as *Stoffmenge*.) If the matter is present only in the solid form, or is only liquid or gaseous, then we say that it is in a *single phase*. If the matter, e.g. water, is present both in liquid form and as a gas, then it has two phases. All substances have, in addition to their mass M, a volume V, a temperature T, and they are subject to some pressure p, for example the ambient atmospheric pressure. The three quantities V, p and T are called the *simple state variables* or *quantities of state*.

All substances are composed of atoms. The amounts encountered in everyday life contain an extremely large number of individual atoms. For example, the *number density*

$$N_{\mathrm{V}} = \frac{\text{Number } N \text{ of molecules in the volume } V}{\text{Volume } V} \tag{13.1}$$

of air at room temperature and pressure is $N_{\mathrm{V}} = 2.5 \cdot 10^{25}/\mathrm{m}^3$. Thus, in each cubic millimeter of air, there are typically $2.5 \cdot 10^{16}$ molecules. *Thermodynamics in the end deals with the behavior of enormous numbers of individual particles*. It says nothing whatever about the behavior and the fate of single individuals. We cannot specify at any given time for any one of them its exact position or velocity, neither its magnitude nor its direction. Pressure and temperature cannot even be defined for a single molecule. We can measure only the state which characterizes the ensemble of all the molecules.

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_13

Part III

In order to permit a simple description of relationships which depend to first order on the number of particles involved, the concept of the *amount of substance* was defined as a physical quantity in 1971, as an additional base quantity with the base unit "mole" (symbol: mol). 1 mol of a substance contains exactly as many particles as there are atoms in 12 grams of ^{12}C. The amount of substance n is thus a measure of the number N of particles:

$$n = \frac{N}{N_A} \quad \text{or} \quad N = n \cdot N_A \,.$$

Here, N_A is the AVOGADRO constant:

$$N_A = 6.022 \cdot 10^{23}\,\mathrm{mol}^{-1} \,.$$

This definition of the amount of substance[C13.1] means that it is to be used in physical-quantity equations just like other physical quantities, in particular for quantitative statements always as a numerical value *and* a unit (see Sect. 3.3 and comment C2.2 in Sect. 2.2).

Quantities which refer to the amount of substance are termed "molar", e.g. the molar mass $M_m = M/n$ (unit: kg/mol) and the molar volume $V_m = V/n$ (unit: m^3/mol).

C13.1. POHL did not include this definition of the amount of substance in his textbooks, although he was in other cases very open to accepting new definitions of quantities and units. Here, however, he "saw no advantage in it", as he wrote in the preface to the 12th edition, and he continued to use the mole as an "individual unit of mass". – We have nevertheless decided to introduce this convention into the text, as it has now been in effect for more than 40 years. We note however that this quantity is still not familiar to many physicists and is not very carefully defined in most textbooks. One finds again and again for example the incorrect statement that n is "the number of moles"!

13.2 The Definition and Measurement of Temperature

In geometry, we measure *one* quantity as the base quantity, namely *length*. In kinematics, a second quantity is needed, the *time*; and for dynamics, a third quantity, the *mass*. In thermodynamics, a fourth quantity is required, the *temperature*.

In the skin on the surface of our bodies and in some mucous membranes, besides pressure and pain receptors, we have another sort of sensory organs. One type of these reacts to external stimuli with the perception of *warm*, the other with the perception of *cold*. Guided by these sensory organs, we can order objects in a series according to their ability to call forth a feeling of 'warm' or 'cold'. The origin of these perceptions is called the *temperature* of the objects. The temperature, qualitatively defined in this way, is useful also as the "cause" of numerous other phenomena, including many that are independent of our perceptions. Changes in the temperature also cause changes in

1. the dimensions of objects. With increasing temperature, metal wires become longer and stretched rubber bands become shorter (Fig. 13.1); bimetallic strips become curved (Fig. 13.2) and gases expand[C13.2] (**Video 13.1**).
2. the absorption of light. Thus, for example, HgI_2 appears reddish at lower temperatures and yellow at higher ones (> 131 °C).

C13.2. This thermal (length or volume) expansion is described by the *thermal expansion coefficient* $\alpha = \frac{1}{l}\frac{\Delta l}{\Delta T} \approx \frac{1}{3}\frac{1}{V}\frac{\Delta V}{\Delta T}$. It can itself depend upon the temperature. For iron, for example, at room temperature, $\alpha = 1.23 \cdot 10^{-5}\,\mathrm{K}^{-1}$. For rubber, α is negative.

Video 13.1:
"Model experiment on thermal expansion and evaporation"
http://tiny.cc/nggvjy
The thermal expansion of a solid body as well as its melting and evaporation are demonstrated in a model experiment.

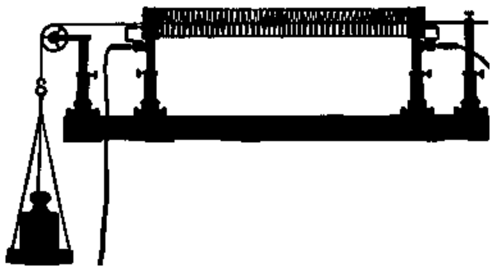


Figure 13.1 A stretched rubber band (length $\approx$ 50 cm) draws together by about 3 cm when it is heated to ca. +90°C. We must increase its load from 2 up to 2.2 kg in order to restore its original length.

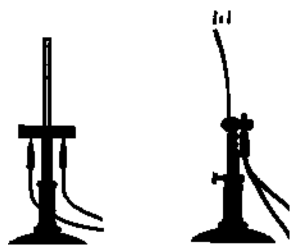


Figure 13.2 Bending of an electrically-heated bimetallic strip (at the left seen from the front, with the strip lying in the plane of the page; at the right seen in profile). It consists of two layers of sheet metal, made of a nickel-iron alloy, which are welded together; one of the strips contains 6 percent by weight of manganese. The strip has been slit so that it can be heated by an electric current.

3. the electrical resistance of metals, which increases linearly with temperature at not-too-low temperatures.
4. the electric voltage between two different metals which are in contact (Fig. 13.3: the thermocouple).

This list could be continued almost indefinitely: The majority of all physical and chemical phenomena show some dependence on the temperature. Each of these could be used to define a *measurement procedure* for the temperature and to construct a measuring instrument, called a *thermometer*, for the temperature.

In daily life, we often use the volume change of a *liquid*, as in a mercury thermometer, with a temperature scale suggested by the Swedish mathematician and surveyor A. CELSIUS in 1742. It contains one hundred divisions between the temperature of melting ice and that of boiling water, and is familiar today to every school child. Hg thermometers can be used between +800 and −39 °C. For lower temperatures down to −200 °C, thermometers filled with *pentane* are used.

Hg thermometers for use up to 300 °C are evacuated; at higher temperatures, vaporization of the Hg is prevented by filling the thermometer tube above the liquid with nitrogen gas at pressures up to 100 times atmospheric pressure (10^7 Pa). – Addition of a small fraction of thallium allows mercury to remain liquid down to −59 °C, so that it is usable for temperature measurements down to this temperature.

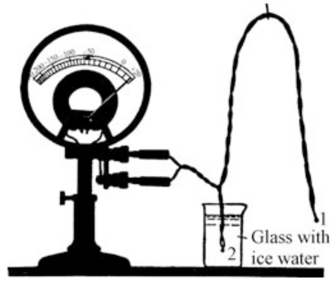


Figure 13.3 An electrical thermometer for demonstration experiments. It consists of two wires, one made of silver and the other of Constantan, which are soldered together at the points 1 and 2 and connected to a voltmeter. The contact point 1 is brought into thermal contact with the object to be measured, while point 2 is held at a reference temperature of 0 °C, for example using an ice-water bath[C13.3]. Compare also Fig. 15.6.

C13.3. The two thermocouples are connected in opposition, so that their difference voltage is measured (as in Fig. 15.6).
Today, for most demonstration experiments, electrical resistance thermometers with digital temperature indication are employed; they are more convenient to use.

Besides liquid thermometers or outside their limiting temperature ranges, electrical thermometers are often used. These can be thermocouples as shown in Fig. 13.3, or also resistance thermometers. For even higher temperatures, optical temperature measurements using "radiation thermometers" play an important role (see Vol. 2, Chap. 28).

All thermometers are calibrated today using legally determined and accurately reproducible temperature values known as "fixed points". These fixed points (melting points, vapor pressures, etc.) have been developed with a considerable amount of tedious effort. The following aspects are important for this work:

A quantitative definition of the temperature using a Hg or other liquid-expansion thermometer is – in spite of its great practical applicability – not completely satisfactory. This can readily be seen by referring to Fig. 13.4. It shows the scale of a Hg thermometer of the usual technical design, and to its left, the scale of an alcohol thermometer which has been calibrated with its aid. In the range shown, only the divisions on the scale of the Hg thermometer are uniform, while those of the alcohol thermometer are not. The temperatures defined by the volume expansion of liquids are seen to depend on the arbitrary choice of the substances (e.g. Hg), and on the type of glass used for the tube.

The temperature defined with gas thermometers (Fig. 13.5) is, for sufficiently low gas densities, practically independent of the nature of the gas used. A conceptually fully-satisfactory definition of the temperature should however be not only practically, but also fundamentally independent of the substance used in the thermometer. This goal was attained by the "thermodynamic temperature scale" (see Sect. 19.6). The gas thermometer belongs among the most important measurement instruments which can be used to establish a practical temperature scale based on the thermodynamic scale.

The currently best representation of the true thermodynamic temperatures T is presumed to be the "International Temperature Scale of 1990" (ITS-90), which contains fixed points ranging from 0.65 Kelvin

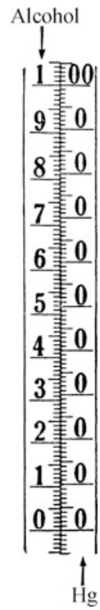


Figure 13.4 *Right:* The scale of a mercury thermometer with uniform divisions between 0 °C and 100 °C; at *left:* the divisions of an alcohol thermometer which has been calibrated using the mercury thermometer

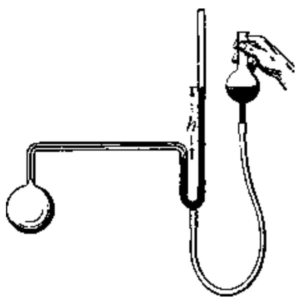


Figure 13.5 Schematic of a constant-volume gas thermometer. By adding mercury, the volume of the gas can be held constant and the pressure can be read off the height h of the mercury column. From the pressure, the temperature of the gas can be calculated according to Sect. 14.6 (cf. Sect. 14.9, Point 3)

up to the highest temperatures that can be measured by applying PLANCK's radiation law[C13.4]. The unit 1 kelvin (K) has been defined as the 273.16-th part of the temperature of the triple point of water. In daily life, we use the CELSIUS temperature scale. Its relation to the KELVIN scale is

$$\frac{t}{°\mathrm{C}} = \frac{T}{\mathrm{K}} - 273.15$$

(273.15 K is the melting temperature of ice under atmospheric pressure).

Science and technology have spoiled us by providing readily usable and accurate thermometers, no less than by e.g. the construction of precise clocks and electrical multimeters. Nevertheless, the fundamental question of the measurement procedure and the calibration of the instruments should not be neglected; otherwise, one could easily overlook the great efforts that have been expended in the past to develop reliable measurement technology.

C13.4. A more detailed treatment can be found in: W. Blanke, *Physik in unserer Zeit* **22**, 13 (1991). English: see bipm.org. In the year 2000, a further international temperature scale (PLTS-2000) was adopted, which extends the scale to the range of very low temperatures (from 1 K down to 0.9 mK).

13.3 The Definitions of the Concepts of Heat and Heat Capacity

The concept of 'temperature' by itself is not sufficient to describe all the processes associated with *changes in temperature*. This can already be seen from a simple example, namely the equalization of the temperature between two bodies of different composition and initially at different temperatures (Fig. 13.6, top). – Let the temperatures of the two bodies be T_1 and T_2, and their masses M_1 and M_2. The two bodies are brought into close thermal contact. Neither chemical transformations nor phase changes are assumed to take place, i.e. solids are presumed to remain solid, liquids remain liquid, etc. After a certain time, a 'mixing temperature' T, which lies between T_1 and T_2, will be observed. It is however not simply a mean value; the relation

$$M_1T_1 + M_2T_2 = (M_1 + M_2)T$$

does *not* hold. Rather, we require two factors c_1 and c_2 to describe the process, and we must write

$$M_1c_1T_1 + M_2c_2T_2 = (M_1c_1 + M_2c_2)T \tag{13.2}$$

or, rearranged,

$$c_1M_1(T_1 - T) = -c_1M_1(T - T_1) = c_2M_2(T - T_2)\,,$$

$$-c_1M_1\Delta T_1 = c_2M_2\Delta T_2\,. \tag{13.3}$$

For the product

$$Q = cM\Delta T \tag{13.4}$$

we use the name *heat*, and instead of Eg. (13.3), we write

$$-Q_1 = Q_2\,.$$

In words: *The heat $-Q_1$ which is given up by the warmer body 1 when thermal contact is established is equal to the heat Q_2 which is taken on by the cooler body 2* (G.W. RICHMANN, 1711-1753)[C13.5].

C13.5. RICHMANN proposed this rule in 1747/48 ("RICHMANN's rule"). In 1753, he was fatally injured while carrying out an experiment with lightning.

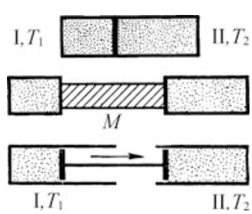


Figure 13.6 *Top:* The flow of heat from a hot to a cool gas container when direct contact is established. *Center:* The same process when the two containers are connected by a metal rod M. *Bottom:* Performing mechanical work transfers energy from a gas under high pressure to a gas at lower pressure; if initially $T_1 = T_2$, then after the transfer, $T_2 > T_1$.

The word *heat* is unfortunately used in different ways. Only in special cases does it refer to a particular form of energy, just like the words potential, kinetic, electrical and magnetic energy. Most often, the word 'heat' characterizes that special type of energy which can be *transferred* from a particular quantity of matter to another quantity of matter: This is a transport process which is driven by a *temperature difference* alone, *without any other contributing factors*. We refer to such a transport process for short as *thermal*. – In the special case mentioned above, one refers to heat as *the kinetic part of the internal energy* (cf. Sect. 16.1).

Heat transport can take place in various ways: either through *conduction*, when the two quantities of matter or their containers are in direct contact or via a thermally conducting material (Fig. 13.6, top and center); through *radiation*, if they are separated by a vacuum (or in practice often by air) (see Vol. 2, Chap. 28); or through *convection* (see Sect. 17.6).

The italicized words "without any other contributing factors" will be illustrated by comparing two examples: In Fig. 13.6 (top), a container of warm gas *I* is brought into direct contact with a container of cool gas *II*. Instead of the direct contact through the walls of the containers, a metal rod *M* (a "good conductor of heat") could also be used to connect the two containers; cf. Fig. 13.6 (center). In both cases, the temperature T_1 is initially higher and the temperature T_2 lower. *Heat* is conducted from the container *I* to container *II*.

In Fig. 13.6 (bottom), each of the gas containers has a sliding wall in the form of a piston. The two pistons are connected rigidly to each other by a glass connecting rod (a "thermal insulator"). The pressure of the gas in container *I* is higher than in container *II*. After some sort of catch (not shown) is released, the two pistons will move to the right. In this process, T_1 will decrease and T_2 will increase. Energy is removed from the left-hand container, because the *gas* it contains is performing work; the energy is added to the right-hand container, since its *piston*, in moving to the right, performs work on the gas in this container. Here, a "process of work" is taking place, i.e. a "transfer of work", which also leads to changes in the temperatures. This example illustrates that heat and work (or energy) are essentially similar. – Likewise, when a quantity of matter is heated through friction, mechanical work is being performed[1].

In practice, heat transfer processes play an important role both in the kitchen and in the laboratory. We could think for example of an electric immersion heater in water. The container is presumed to be *thermally insulating*, i.e. it *should allow no thermal transport due to temperature differences between its contents and the surroundings*. Thermally well-insulated vessels have double walls made of a poorly-conducting material such as glass. The space between the walls is

[1] Or else electrical work, when an electric current is passed through a resistive conductor and thereby heats it through 'electrical friction' (cf. Fig. 13.2, and Vol. 2, Sect. 1.12).

evacuated to prevent heat transfer by *convection*, and the walls are 'silvered' (with a thin film of silver, copper or aluminum) to reduce heat transfer by *radiation* as far as possible. We will often make use of such containers ("Thermos bottles"), which can be found today in almost every household.

The essential similarity of *heat* and *energy* was recognized as early as 1842 by the physician ROBERT MAYER. In retrospect, we may never be able to fully appreciate the achievements of scientific pioneers. Today, the similarity of heat to the other forms of energy has long since been considered to be "self-evident"; it has become a matter of course. The unit of heat Q is thus a unit of energy, e.g. joule (J) or watt-second (W s) (see Sect. 5.2)[C13.6].

C13.6. The heat unit "calorie", historical but still used in daily life, is given by 1 cal ≈ 4.2 J.

Now that we have introduced the quantity of heat Q, the factors c also acquire their physical significance. They refer to the *specific* heat, i.e. heat *referred to* some other quantity. Suppose for example that an electric heater produces the energy $E = IUt$ in a substance which contains the amount of substance n with the mass M. The energy transfer is thermal, i.e. the energy is delivered in the form of heat Q. Q increases the temperature of the substance by ΔT. The quotient $Q/\Delta T$ is called the *heat capacity*. Then, as the *specific heat capacity* (often simply called the "specific heat"), we define

$$c = \frac{Q}{M \cdot \Delta T}\,, \tag{13.5}$$

and as the *molar heat capacity*

$$C = \frac{Q}{n \cdot \Delta T}\,. \tag{13.6}$$

Table 13.1 lists some values[C13.7]. They are valid in every case only within a limited temperature range.

C13.7. "Molar" quantities are referred to the *amount of substance* n (cf. Sect. 13.1). Making use of the "molar mass":
$M_m = \dfrac{\text{Mass } M}{n}$,
specific quantities, i.e. quantities referred to the mass of a substance, can readily be converted to *molar* quantities referred to the amount of substance and *vice versa*. For the values in Table 13.1, this can readily be verified. – The numerical value of the molar mass was previously termed the "molecular weight" (see also comment C16.1).

Table 13.1 Specific and molar heat capacities and heats of melting of some solids and liquids

Substance	Molar mass	Specific and molar heat capacity at 20 °C		Specific and molar latent heat of melting	
	M_m $\left(\frac{\text{kg}}{\text{kmol}}\right)$	c $\left(\frac{\text{kW s}}{\text{kg} \cdot \text{K}}\right)$	C $\left(\frac{\text{kW s}}{\text{kmol} \cdot \text{K}}\right)$	l_f $\left(\frac{10^5\ \text{W s}}{\text{kg}}\right)$	L_f $\left(\frac{10^6\ \text{W s}}{\text{kmol}}\right)$
Aluminum	26.98	0.897	24.2	4.036	10.9
Copper	63.54	0.385	24.5	2.047	13.0
Lead	207.2	0.128	26.5	0.247	5.12
NaCl	58.45	0.85	49.7	5.17	30.2
Benzene	78.11	1.69	132	1.26	9.84
Water	18.02	4.16	75	3.34	6.02

13.4 Latent Heat

(Josef Black, 1762)[C13.8]. In our experiments thus far, the substances have undergone no kind of transformation. Solid bodies remained solid, liquids remained liquid and gases remained gaseous. The composition of materials also remained unchanged, both their chemical composition as well as their crystal structures. We now relax this restriction, and allow *phase transitions* to take place. Then a substance can accept or release energy in the form of heat *without* changing its temperature. In this case, the heat transferred is called *latent*. We offer two important examples:

C13.8. J. Black (1728–1799) carried out investigations on thermal equilibrium and discovered, independently of J.C. Wilcke, the concepts of latent heat and specific heat.

1. *Specific heat of vaporization and condensation.* In Fig. 13.7, we see a container which is partly filled with water and then evacuated. A manometer is attached to the container so that we can observe the pressure inside it, and also an adjustable spring valve. In addition, the container is equipped with an electric heater, insulated from the outside world.

After switching on the heater current, we observe that the water becomes warmer, and with its increasing temperature, its vapor pressure also increases[C13.9]. The vapor is continuously in contact with the liquid water and is in *equilibrium* with it. The resulting partial pressure of the water vapor is termed the *saturation pressure* or the *vapor pressure*. Numerical values can be found in Fig. 14.3. At a certain pressure p, the spring valve opens and the vapor escapes continuously. From this moment, the temperatures of both the water and of the vapor remain constant, and the two temperatures remain equal. – Deduction: The vapor which is escaping must be continually replaced; water must be constantly passing from the liquid to the vapor phase. The heat Q which is being passed into the water is consumed by the process of evaporation without any increase in the temperature, i.e. it is taken up as *latent heat*. The amount of water which has evaporated, measured in terms of its mass M, is found to be pro-

C13.9. *Vapor* refers here to the gaseous phase of water.

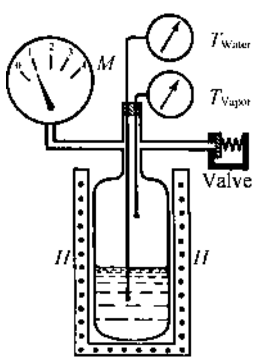


Figure 13.7 The measurement of the heat of vaporization (H = electric heater inside the insulation). The manometer M indicates normal air pressure when its connecting tube is open to the room air. It thus measures the entire pressure of the *vapor*, *not* just its excess pressure over the normal air pressure.

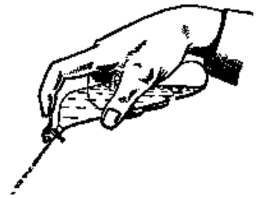


Figure 13.8 A cooling bottle with liquid ethyl chloride (C_2H_5Cl)

portional to the heat Q which has been added. Thus, we form the quotient

$$l_v = \frac{\text{Added heat } Q}{\text{Mass } M \text{ of the evaporated liquid}}, \qquad (13.7)$$

and call it the *specific latent heat of vaporization*[C13.10]. – In the next chapter, Fig. 14.3 shows measurements of the vaporization of water at temperatures between 0 °C and 374 °C.

C13.10 POHL used the symbol 'r' here; we have used the more relevant 'l', which stands for "latent heat". The symbol l_v refers to the (specific) latent heat of *vaporization*, i.e. referred to the mass of the body; compare also Sects. 14.3 and 14.4. The subscript 'v' denotes "vaporization". The symbol 'L' refers to the *amount of substance n* and is then termed the *molar* latent heat (similarly to heat capacities; see previous section). It may also refer to the latent heat of an *entire body*, e.g. the heat of melting of an ice cube, and this should be specifically mentioned. In Eq. (13.8), POHL originally used the symbol χ for the latent heat of melting (or fusion). We again use 'l' or 'L' here, with a subscript f for "fusion".

Every liquid which is evaporating takes up heat from its surroundings. Numerous types of cooling apparatus are based on this fact. In the laboratory, the cooling bottle sketched in Fig. 13.8 is often used. It contains liquid ethyl chloride (boiling temperature = 13 °C, vapor pressure at 18 °C = $1.24 \cdot 10^5$ Pa). The liquid is sprayed out of the small nozzle under its own vapor pressure. The surface struck by the spray has to provide the heat of vaporization, so that it is cooled. In this way, one can readily obtain temperatures below 0 °C in the laboratory. In medicine, this method is used to provide a local anesthesia by cooling. (Demonstration: Spray a piece of black paper, breathe on it and observe how it becomes frosty.)

The heat required for vaporization can be regained by condensing the vapor back to a liquid. Apart from its sign, the specific latent *heat of condensation* is equal to the specific *heat of vaporization*. For demonstration experiments, water vapor (steam) is passed into a thermos bottle filled with cold water. It condenses there, and warms the water in the process. From the mass of the condensed water and the temperature increase, we can calculate the value of the heat of condensation.

2. *Specific heat of melting and crystallization.* The specific heat of melting is determined in essentially the same way as the specific heat of vaporization. We determine the mass M of the substance which has been melted by applying the quantity of heat Q to it. Then we define the *specific latent heat of melting* to be the quantity

$$l_f = \frac{\text{Added heat } Q}{\text{Mass } M \text{ of the melted substance}}. \qquad (13.8)$$

Table 13.1 contains examples, along with the correspondingly-defined molar heat of melting $L_f = Q/n$, again at normal atmospheric pressure.

The latent energy contained in the liquid can be completely regained when it solidifies. The *heat of crystallization* is, apart from its sign, identical to the *heat of melting*. For a demonstration experiment, sodium thiosulfate ($Na_2S_2O_3 \cdot 5\,H_2O$), used as a fixing agent in photography, is well suited as the substance to be melted or solidified.

The melting point of this salt lies at +48.2 °C. The melt can be strongly supercooled. It can be maintained for days at room temperature. If the melt is “seeded” with a small crystal, the process of crystallization begins, accompanied by a considerable rise in the temperature. This process can be used to evaporate ether and make it visible at some distance by igniting the ether vapor. – Technically, this transition can be used to manufacture heating pads. The salt is filled into a rubber pouch and melted by immersing in hot water. It will then maintain its temperature for some time at +48 °C (the “hold point”).

3. *Heat of transition.* In Fig. 13.9, a piece of carbon-containing sheet iron (0.9 percent by weight of C) is heated by means of an electric current until it is glowing yellow. After the current is switched off, it cools rapidly and darkens. As it passes through a temperature of $t \approx 720\,°\mathrm{C}$, it again flashes brightly: at this temperature, it undergoes a phase transition – delayed by supercooling – from the form of iron called the γ phase into a less energetic mixture of carbon-free γ iron and Fe_3C (cementite). At this transition, a considerable amount of heat is released.

Without the supercooling, the temperature decrease would simply be brought to a standstill for a short time: the “hold point” is characteristic of a *phase transition*.

Summary: Adding heat to a system can not only increase the temperature of some substance, but also can induce a transition between phases at constant temperature in the interior of the material. In both cases, energy is stored within the substance. All forms of energy stored within a material are termed “internal energy” U. It is thus qualitatively distinguished from the potential and kinetic energies which the substance may possess *as a whole*. More details will be given in Sect. 14.3.

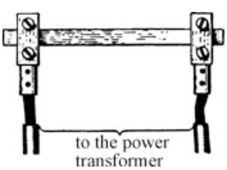


Figure 13.9 Demonstrating the heat of transition

Exercise

13.1 Calculate the influence of thermal expansion on the movement of a pendulum clock whose pendulum is made of steel, if the temperature increases by 20 K. For simplicity, the pendulum may be treated as a mathematical gravity pendulum. The linear coefficient of expansion α is $1 \cdot 10^{-5}\,\mathrm{K}^{-1}$ (Sect. 13.2).

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_13) contains supplementary material, which is available to authorized users.

The First Law and the Equation of State of Ideal Gases

14

14.1 Work of Expansion and Technical Work

Our next goal is to obtain a quantitative summary of the law of energy conservation, that is the First Law of thermodynamics. This and the following section will provide preparatory discussions.

In the mechanics of solid bodies, we defined the work W as the product "force in the direction of the motion times distance moved", i.e. $W = \int F\,\mathrm{d}s$ (Sect. 5.2). The force F can itself be expressed as the product "pressure p times area A" when it is exerted by a liquid or a gas. Then for the work, we obtain $W = -\int pA\,\mathrm{d}s$ or, since $A\,\mathrm{d}s$ = volume element ($\mathrm{d}V$)[C14.1],

$$W = -\int p\,\mathrm{d}V\,. \tag{14.1}$$

Just as with previous examples, we want to make this expression for the work clear by providing a sketch; it can be seen in Fig. 14.1. A working substance which is contained in a cylinder is pressing against a piston. It displaces the piston to the right, increasing its volume in the process, and thus performs mechanical work. The motion is supposed to take place so slowly that no local non-uniformities of pressure, density or temperature can occur. The pressure does not remain constant during the motion of the piston, as shown by the curve 1 … 2 in the figure. The work of expansion $-\int_1^2 p\,\mathrm{d}V$ corresponds to the shaded area *below* the expansion curve.

In technological applications, all machines work in *periodic cycles*. They can perform work only by making use of a *flowing* working substance. For this case, the concept of *technical* work W_{techn} has been defined. It is explained in Fig. 14.2. The upper images show the cylinder of a machine with an inlet and an outlet valve and a piston.

In the first time interval (stroke) of the cycle, the working substance flows at constant pressure into the cylinder. It occupies its initial volume by pushing the piston out to the position 1. In the process, it performs the work of displacement $-p_1V_1$ on the piston, i.e. the

C14.1. Determining the *signs* of energetic quantities is based on the following convention: Energy which is *input* into a particular system, e.g. the working substance in Fig. 14.1, is taken to be positive, while the energy that is *output* is negative. Referred to the working substance, the *work of expansion* which it performs

$$W = -\int_1^2 p\,\mathrm{d}V$$

is negative (both factors in the integral are positive), while the *work of compression*

$$W = -\int_2^1 p\,\mathrm{d}V$$

which is performed *on* it is positive, since owing to the reversed limits of the integral, $\mathrm{d}V$ now has the opposite sign. ▷

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_14

C14.1 *Continued*
The signs have been consistently defined here, to conform to this convention and to agree with other textbooks. The opposite signs which were used by POHL in some cases in earlier editions, which corresponded to the technical thermodynamics conventions of the time, have all been removed from this edition. This modification corresponds to a suggestion made by K. HECHT, a friend and former student (Dr. rer. nat. 1930) of the author, who pointed out in a detailed letter as early as 1985 that sign changes were needed.

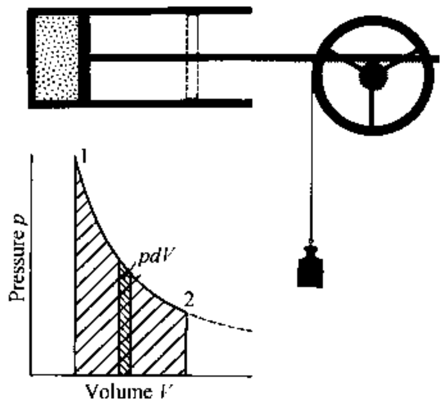


Figure 14.1 p-V diagram for defining the work of expansion, $-\int_1^2 p\,\mathrm{d}V$: It is the shaded area enclosed below the expansion curve. The working substance in this example performs not only lifting work, but also work of acceleration (cf. Fig. 18.2)

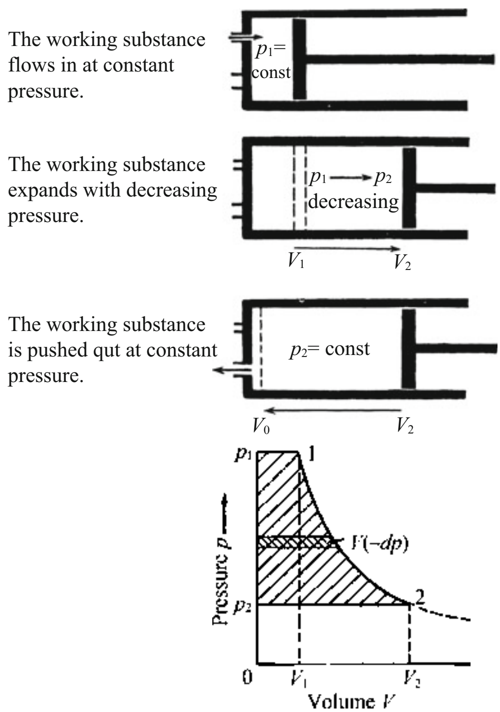


Figure 14.2 p-V diagram for defining technical work $W_{\text{techn}} = \int_1^2 V\,\mathrm{d}p$: the shaded area enclosed *alongside* the expansion curve. The working substance flows through the machine from a pressure vessel with a high, constant pressure p_1, e.g. a steam boiler, and is received by a vessel with a low, constant pressure p_2, e.g. a condenser, or into the free atmosphere.

working substance gives up energy. In the second interval, the inlet valve is closed and the working substance expands further, pushing the piston to position 2; its pressure decreases from p_1 to p_2. In this process, the working substance performs the work of expansion $W = -\int_1^2 p\,\mathrm{d}V$ on the piston. In the third interval, the outlet valve is open and the working substance is pushed out of the cylinder at the constant pressure p_2 by the piston. It regains the work of displacement p_2V_2 from the piston.

The working substance thus performs two portions of work on the piston, namely $-p_1V_1$ and $-\int_1^2 p\,\mathrm{d}V$. While it is flowing out of the cylinder, it receives the work p_2V_2. Thus, in sum, the working substance performs the following *technically useful* (or for short *technical*) work on the piston:

$$W_{\text{techn}} = -p_1V_1 - \int_1^2 p\,\mathrm{d}V + p_2V_2 = \int_1^2 V\,\mathrm{d}p\,. \tag{14.2}$$

	$-p_1V_1$	$-\int_1^2 p\,\mathrm{d}V$	p_2V_2	$\int_1^2 V\,\mathrm{d}p$
corresponds in Fig. 14.2 to:	vertical rectangular area $O{-}p_1{-}1{-}V_1$	area $V_1{-}1{-}2{-}V_2$ *under* the expansion curve	horizontal rectangular area $O{-}p_2{-}2{-}V_2$	shaded area $p_1{-}1{-}2{-}p_2$ *alongside* the expansion curve

So much for this specific example based on a cylinder and piston. In an analogous manner, we quite generally distinguish two different cases:

1. A confined working substance expands and performs work on the external system[C14.2]:

$$\textit{Work of expansion}\ W = -\int_1^2 p\,\mathrm{d}V\,. \tag{14.3}$$

C14.2. Don't let yourself be confused by the different signs of W and W_{techn}. Both of these expressions are negative; they thus describe work which is *performed by* the working substance on the external system.

2. A working substance *flows through* some arbitrary machine, increases its volume in the process from V_1 to V_2, and reduces its pressure from p_1 to p_2. It performs the

$$\textit{Technical work}\ W_{\text{techn}} = \int_1^2 V\,\mathrm{d}p \tag{14.4}$$

on the external system.

The relation between these two types of work can be seen from Eq. (14.2); it is

$$W_{\text{techn}} = W - p_1V_1 + p_2V_2\,. \tag{14.5}$$

In ideal gases (Sect. 14.6), at constant temperature, $p_1V_1 = p_2V_2$. In this case, there is no difference between work of expansion W and technical work W_{techn}.

14.2 Thermal State Variables

For every amount of substance n, we can specify its volume V, its pressure p, and its temperature T. These readily measurable quantities are called the *simple state variables*, as we have already mentioned in Sect. 13.1.

The defining characteristic of a state variable is that it is independent of the process or the "path" taken by previous changes of state. This independence is not the case for other important quantities, e.g. the work $W = -\int p\,\mathrm{d}V$. This can be seen in an example: In Fig. 14.1, the work performed is represented by the shaded area in the p-V diagram. This area however depends on the "path", i.e. in this example on the shape of the curve which leads from state 1 to state 2. – There is a unique relation between the state variables when only one phase is present; it is called the *thermal equation of state*. It is particularly simple in the case of an ideal gas. – When any two of the three state variables are known, the thermal equation of state determines the third, independently of any changes of state that may have taken place in the meantime. The only prerequisite is that none of the changes of state may change the chemical composition or other properties, e.g. the microcrystalline structure of the substance[C14.3].

C14.3. The amount of substance n or the number of particles N of the atoms or molecules involved is also assumed here to remain constant. When it is not, an additional state variable, the *chemical potential* μ must be taken into account (see e.g. Baierlein, Ralph (April 2001), "*The elusive chemical potential*", *American Journal of Physics* **69** (4), pp. 423–434).

In addition to the simple thermal state variables mentioned, there are a number of other state variables or *state functions*. Among these, we will encounter the internal energy U, the enthalpy H, the entropy S, and the free energy F. A small number of state variables is always sufficient to specify all of the observable and measurable quantitative relations in a (macroscopic) thermodynamical system.

14.3 The *Internal Energy U* and the First Law

We return to considering Fig. 14.1. Suppose that a quantity of heat Q is *applied to* the confined gas, e.g. by an electrical heater which is not shown in the figure. Some part of this heat can be removed to the external system as work W. Think of a steam engine, or of increasing the area of a liquid surface (cf. surface work, Sect. 9.5) or of the output of electrical energy, e.g. by a thermocouple. The rest of the energy which was input as heat can be stored in the inner parts of the system as *internal energy* U, thus increasing it by the amount ΔU. In the form of an equation[C14.4], we have:

C14.4. The formulation of the First Law (Eq. (14.6)) does *not* depend on whether the energies considered are input into a system or output by it. This applies also to the special case shown in Eq. (14.7)! The remarks given below the equation are only exemplary.

General

$$Q + W = \Delta U \tag{14.6}$$

For the special case of work of *expansion*

$$\begin{array}{ccccc} Q & + & \left(-\int p\,\mathrm{d}V\right) & = & \Delta U \\ \left.\begin{array}{c}\text{Heat input}\\ (\textit{not}\\ \text{a state variable})\end{array}\right\} & + & \left\{\begin{array}{c}\text{Energy}\\ \textit{out}\text{put as external work}\\ (\textit{not}\text{ a state variable})\end{array}\right\} & = & \left\{\begin{array}{c}\text{Increase of}\\ \text{the internal energy}\\ \text{(state variable).}\end{array}\right. \end{array} \tag{14.7}$$

Equation (14.6) is called the *First Law of thermodynamics*. As long as processes proceed without any temperature changes of the substances involved, the sum of potential and kinetic energies in mechanics remains constant. The same is true in electrodynamics for the electrical and the magnetic energies and for combinations of these four forms of energy. For these energies, the "law of energy conservation" has been thoroughly verified. The First Law of thermodynamics includes an additional energy form in these sums, which can be *thermally* taken up by a substance from its surroundings or given up to the surroundings, and in the process can be consumed completely or partially to change the internal energy U of the substance by an amount ΔU.

The internal energy has thus far been only qualitatively introduced in Sect. 13.4. The First Law permits us to define its changes *quantitatively* (corresponding to changes in the potential energy in mechanics). The First Law also implies the assertion that the internal energy is a *state function*. It declares that: Given a system[1] in a particular state 1, characterized by state variables p_1, T_1 . . ., through inputs and outputs of heat Q and external work W (of arbitrary kinds), the system will pass in sequence through the states 2, 3, . . . Finally, it will arrive back at its initial state 1. Then, one will find experimentally – without exception – that the sum of all the energies thermally input to and output from the system is equal to the sum of all the work input to and output by the system. During all the changes of state of the system, no energy is lost or gained. The internal energy in state 1 at the end is exactly the same as it was at the beginning. It is determined only by the state variables (p, T, . . .).

We first apply Eq. (14.7) to a system which consists only of a single chemically uniform substance. We know the value of its *heat of vaporization* $Q = n \cdot L_v$. A liquid is presumed to evaporate at a constant pressure, called its *saturation pressure* (Fig. 13.7). The measured heat of vaporization can be decomposed into two terms according to Eq. (14.7) in the following way:

$$\left\{\text{Heat input } Q\right\} + \left\{\begin{array}{c}\text{Work of expansion}\\ W = -p(V_\text{vapor} - V_\text{liquid})\end{array}\right\} = \left\{\begin{array}{c}\text{Increase } \Delta U\\ \text{of the internal energy}\\ \text{in the transition liquid}\\ \rightarrow \text{vapor.}\end{array}\right\}$$

The increase of the internal energy is $\Delta U = U_\text{vapor} - U_\text{liquid}$. It arises in particular from an increase in the potential energy of the molecules,

[1] As the "system" in physical terms, one means all the substances and objects which are included in the considerations at hand.

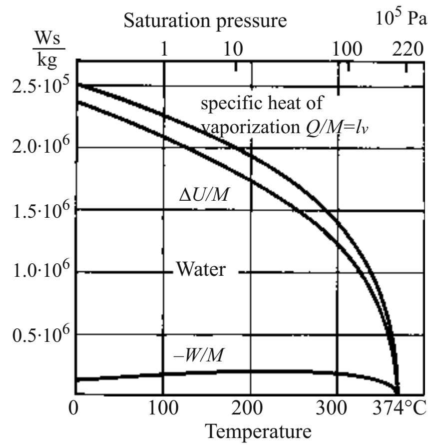


Figure 14.3 The specific latent heat of vaporization l_v of water at different temperatures, and its decomposition into two components, the increase ΔU of the internal energy, and the work of expansion $-W$. All three quantities are referred to the mass of the vapor produced. Thus, the *specific* quantities, referred to the mass M, are shown.

which mutually attract each other. The work of expansion has to be performed against the saturation pressure of the vapor, in order to make room for more vapor that is continually being produced. – Figure 14.3 shows this decomposition for the vaporization of water in the temperature range from 0 to 374.2 °C. In this range, the saturation vapor pressure increases up to $228 \cdot 10^5$ Pa. Above 374.2 °C, $\Delta U = 0$; vapor and liquid become indistinguishable. This temperature is called the “critical temperature” of water.

14.4 The State Function *Enthalpy, H*

In many applications of thermodynamics, we need to make use of a *flowing* working substance, as mentioned above. All such cases can be reduced to the scheme sketched in Fig. 14.4: A working substance flows *out* of a vessel I through a machine M and *into* a second vessel II. The two loaded pistons indicate that *constant pressures* are maintained in the two vessels.

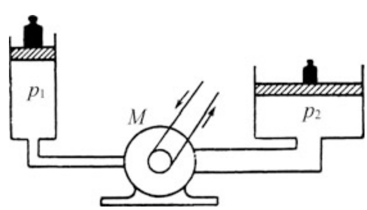


Figure 14.4 The performance of work by a flowing working substance. The meaning of M is explained in the text (fine print). A working substance flowing through the machine has a volume V_1 and the pressure p_1 before entering M; on exiting M, it has a volume V_2 and the pressure p_2.

> M could for example be a steam engine of arbitrary construction, or a pneumatic hammer. Both supply technical work W_{techn} to the external world. M could also be a compressor which inputs the technical work W_{techn} into a quantity of working substance. M could be a mixing machine which increases the temperature; the energy required is input as work. M could also be a heating or cooling apparatus, which adds or removes heat. Finally, various possibilities can be combined; one could for example equip a compressor with a cooling system.

For the treatment of flowing working substances, we introduced the concept of *technical* work. From its defining equation (14.4), it follows that

$$W = W_{\text{techn}} + p_1 V_1 - p_2 V_2 \,. \qquad (14.5)$$

We insert this value into Eq. (14.6), i.e. $Q + W = U_2 - U_1$, and obtain

$$Q + W_{\text{techn}} = U_2 - U_1 - p_1 V_1 + p_2 V_2 \qquad (14.8)$$

or

$$Q + W_{\text{techn}} = (U_2 + p_2 V_2) - (U_1 + p_1 V_1)\,. \qquad (14.9)$$

U, p and V are *state variables*. As a result, the sums in parentheses are also state variables. These sums have been given the name *enthalpy* H; thus

$$\begin{array}{ccccc} H & = & U & + & pV \\ \text{Enthalpy} & = & \text{Internal energy} & + & \text{Work of expansion} \end{array} \;. \qquad (14.10)$$

The enthalpy is a much-used energetic state variable or *state function*. It is applicable to *flowing* working substances in cases where one would use the internal energy U for *confined* working substances. With the enthalpy H and the technical work $W_{\text{techn}} = \int V\,\mathrm{d}p$, Eq. (14.9) takes on the form

$$\begin{array}{ccccccc} Q & + & \int V\,\mathrm{d}p & = & \Delta H \\ \text{Heat input}\} & + & \left\{\begin{array}{c}\text{Technical work performed}\\ \text{on the external system(!)}\end{array}\right\} & = & \left\{\begin{array}{l}\text{Enthalpy}\\ \text{increase.}\end{array}\right. \end{array} \qquad (14.11)$$

Example of an application: If a vapor is produced at its saturation pressure, then p is constant (Fig. 13.7). At constant p, we have $\int V\,\mathrm{d}p = 0$. Therefore, Eq. (14.11) gives $Q = \Delta H$; in words: The heat input for the vaporization is equal to the *increase in the enthalpy* of the substance resulting from the vaporization. Thus, this energy which is required for vaporization (the latent heat L_v) is also referred to as the *enthalpy of vaporization*; and the same applies to the heat of fusion L_f or *enthalpy of fusion*.

14.5 The Two Specific Heats, c_p and c_V

If we know the internal energy U and the enthalpy H, we can now cast the concept of *specific heat capacity* into a physically impeccable form. – Up to now, we have defined the specific heat of a substance by means of the equation

$$c = \frac{\text{Heat } Q \text{ added}}{\text{Mass } M \cdot \text{Temperature increase } \Delta T} . \tag{13.5}$$

The heat put into the substance, e.g. from an electric heater, is distributed quite differently when the system is held at constant volume than when it is at constant pressure. At constant volume, enforced by sufficiently rigid container walls, the temperature rises and thus only the internal energy U of the substance is increased. However, at constant pressure, the substance can expand during the temperature rise. In addition to the increase of its internal energy, there is also work of expansion. In other words: At constant pressure, instead of an increase of the internal energy U, the enthalpy $H = U + pV$ is increased.

As a result, we need to define two different specific heats: first, a specific heat c_V at constant volume, i.e.

$$c_\mathrm{V} = \left(\frac{\text{Increase } \Delta U \text{ of the internal energy}}{\text{Mass } M \text{ of the substance} \cdot \text{Temperature increase } \Delta T} \right)_{V=\text{const}}$$

or

$$c_\mathrm{V} = \frac{1}{M} \left(\frac{\partial U}{\partial T} \right)_{V=\text{const}} . \tag{14.12}$$

Secondly, a specific heat c_p at constant pressure, i.e.

$$c_\mathrm{p} = \left(\frac{\text{Increase } \Delta H \text{ of the enthalpy}}{\text{Mass } M \text{ of the substance} \cdot \text{Temperature increase } \Delta T} \right)_{p=\text{const}}$$

or

$$c_\mathrm{p} = \frac{1}{M} \left(\frac{\partial H}{\partial T} \right)_{p=\text{const}} . \tag{14.13}$$

The difference of the two specific heats is found to be

$$c_\mathrm{p} - c_\mathrm{V} = \frac{1}{M} \left[p + \left(\frac{\partial U}{\partial V} \right)_{T=\text{const}} \right] \left(\frac{\partial V}{\partial T} \right)_{p=\text{const}} . \tag{14.14}$$

Derivation
Start with the definition of the enthalpy, $H = U + pV$. Then, instead of Eq. (14.13), we can write:

$$c_\mathrm{p} = \frac{1}{M} \left[\frac{\partial U}{\partial T} + p \left(\frac{\partial V}{\partial T} \right) \right]_{p=\text{const}} . \tag{14.15}$$

In general, the internal energy U of a body or an amount of substance depends both on T and also on V. Therefore, we obtain

$$\mathrm{d}U = \left(\frac{\partial U}{\partial T}\right)_{V=\text{const}} \mathrm{d}T + \left(\frac{\partial U}{\partial V}\right)_{T=\text{const}} \mathrm{d}V \qquad (14.16)$$

and from this,

$$\left(\frac{\partial U}{\partial T}\right)_{p=\text{const}} = \left(\frac{\partial U}{\partial T}\right)_{V=\text{const}} + \left(\frac{\partial U}{\partial V}\right)_{T=\text{const}} \left(\frac{\partial V}{\partial T}\right)_{p=\text{const}} . \qquad (14.17)$$

Combining Eqns. (14.12), (14.15) and (14.17) yields Eq. (14.14).

So much for the now flawless definitions. For a reasonable *comparison* of the specific heats of various substances, one requires their *molar* heat capacities (Eq. (13.6) from Sect. 13.3). Then, for different materials, we are dealing with the same amounts of substance n or with the same numbers N of particles.

Figure 14.5 shows typical examples of the molar heat capacities C_p for some solid materials. Only at high temperatures do they approach a constant value. With decreasing temperature, they are strongly reduced.

The two specific heat capacities of gases, c_p and c_V, play an important role. Unfortunately, only one of them, namely c_p, the specific heat at constant pressure, can be measured reliably. – The fundamentals of the measurement procedure are explained in Fig. 14.6[C14.5]. A steady flow of gas passes through a coiled tube S within a *calorimeter* K; this is an apparatus for determining heat capacities, here a water calorimeter, in which the heat input ΔQ is obtained from the temperature rise

C14.5 The temperature of the calorimeter increases only slightly above room temperature during the experiment. Therefore, the cooling rate of the hot gas in the calorimeter remains practically unchanged, and (T_1-T_2) quickly becomes constant. Thereafter, both ΔQ and M increase linearly with time.

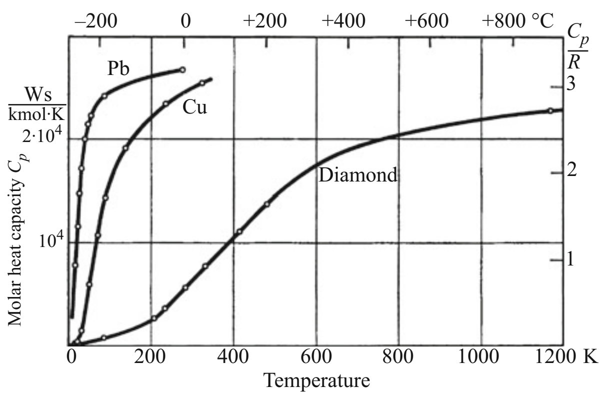


Figure 14.5 The temperature dependence of the molar heat capacity C_p for three solids. If the measured points are plotted against a rescaled temperature axis, T/Θ, then with a suitable choice of Θ, one can make the curves nearly identical. This 'universal curve' is called the DEBYE function[C14.6]. The temperature parameter Θ is the DEBYE temperature of each of the solids (Pb: $\Theta = 88$ K; Cu: $\Theta = 315$ K; Diamond: $\Theta = 1860$ K) (see also Sect. 16.4, final part. Here, R is the universal gas constant; cf. Eq. (14.22))

C14.6. P. Debye, Ann. Phys. **39**, 789 (1912).

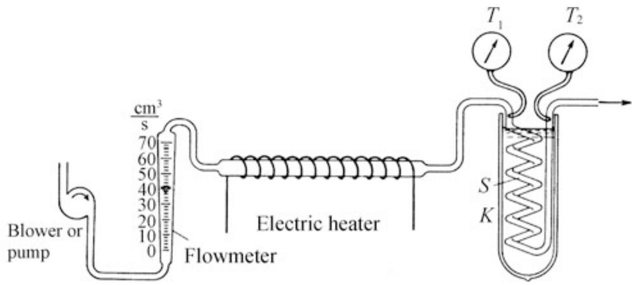


Figure 14.6 Scheme of the measurement of the specific heats of gases at constant pressure. The flowmeter operates according to the "Rotax" principle: In a glass tube whose diameter increases conically, there is a float with short, propeller-like wings. The float rises higher with increasing gas flow rate (gas volume/time). The coiled tubing S enhances heat exchange with the water in the calorimeter K, which is in a double-walled vessel like a Thermos bottle (see Sects. 13.3 and 17.9). Not shown is the thermometer for measuring the temperature rise of the water

Part III

of the water and the known heat capacity of the calorimeter vessel. The temperature of the gas is measured before and after it passes through the calorimeter; likewise the mass M of the flowing gas. The heat deposited in the calorimeter is then $\Delta Q = \Delta H = c_p M(T_1 - T_2)$. Measurements of this kind are suitable as practical laboratory exercises; but as lecture demonstration experiments, they have a boring effect. Table 14.1 contains some specific heats that were measured in this way.

"Measurements of this kind are suitable as practical laboratory exercises; but as lecture demonstration experiments, they have a boring effect."

Measurements of c_V, the specific heat at constant *volume*, are an awkward matter. The heat taken up by the walls of the gas container is more than that absorbed by the gas confined in it. The corrections are in general larger than the quantities to be measured. This can be avoided only by limiting the measuring time (to less than 10^{-2} seconds). As a result, usually an indirect route is followed for determining c_V. One measures the ratio $\kappa = c_p/c_V$ (or, equivalently, C_p/C_V) and computes from it the specific heat c_V using knowledge of c_p. The values listed in Table 14.1 were obtained in this manner.

There are several useful procedures for measuring $\kappa = c_p/c_V$, known as the *adiabatic exponent* (Sect. 14.9, Point 4). It is however expedient to discuss them later (Sect. 14.10).

14.6 The Thermal Equation of State of Ideal Gases. Absolute Temperatures

Thermodynamics received a considerable clarification through the investigation of gases, beginning with the thermal equation of state for

Table 14.1 Heat capacities of various gases

Gas		Mass density ϱ at 0 °C and 10^5 Pa	Molar mass	Specific and molar heat capacities at 20 °C				$\kappa = \frac{c_p}{c_V}$
				c_p	c_V	C_p	C_V	
		$\left(\frac{\text{kg}}{\text{m}^3}\right)$	$\left(\frac{\text{kg}}{\text{kmol}}\right)$	$\left(\frac{\text{kW s}}{\text{kg} \cdot \text{K}}\right)$		$\left(\frac{\text{kW s}}{\text{kmol} \cdot \text{K}}\right)$		
He	Mon-atomic	0.179	4.003	5.23	3.21	20.94	12.85	1.63
Ar		1.784	39.94	0.523	0.317	20.94	12.69	1.65
H_2	Di-atomic	0.0899	2.016	14.3	10.1	28.83	20.45	1.41
O_2		1.429	32.0	0.918	0.655	29.37	20.98	1.40
air		1.293	29	1.005	0.717	29.14	20.78	1.40
CH_4	Poly-atomic	0.7168	16.04	2.22	1.697	35.59	27.21	1.31
NH_3		0.771	17.03	2.16	1.655	36.78	28.18	1.31
CO_2		1.977	44.01	0.837	0.647	36.83	28.48	1.29

ideal gases. – Up to now, we have encountered the "ideal gas law" only for the special case of constant temperature. It states: For an ideal gas at constant temperature, the quotient '(pressure)/(mass density)' or the product '(pressure) times (specific volume)' is constant. Or, in the form of an equation,

$$\frac{p}{\varrho} = pV_s = \frac{pV}{M} = \text{const.} \qquad (9.11 \text{ to } 9.13)$$

(p = pressure, M = mass of the gas confined in the volume V, $\varrho = M/V$ = mass density of the gas, and $V_s = 1/\varrho = V/M$ = specific volume of the gas).

For air at 0 °C, one finds experimentally the quotient of pressure/mass density to be

$$\left(\frac{pV}{M}\right)_{0\,°\mathrm{C}} = 7.84\,\frac{\text{W s}}{\text{kg}}\,. \qquad (14.18)$$

This quotient of pressure/mass density has been measured over a wide range of temperatures, not only for air, but for many other gases. Some of the results are summarized in Fig. 14.7 (upper part). For all the gases, one observes straight lines when the values are plotted against the temperature. The *slopes* of these lines differ from one gas to another, but the extrapolation of all the lines to lower temperatures cuts the axis of temperature at the same point, namely at −273.2 °C.

In the lower part of the figure, instead of the quotient pV/M, the ordinate represents the quotient pV/n (n is the *amount of substance*). This produces an essential simplification: Now, the *slope* of the lines is the same for all gases. A single straight line can be used to fit the measured points for all the different gases. Its intersection with the abscissa at −273.2 °C remains the same. Thus, this temperature, −273.2 °C, has been found to be a distinguished value with a universal significance; it is called the 'absolute zero' of temperature.

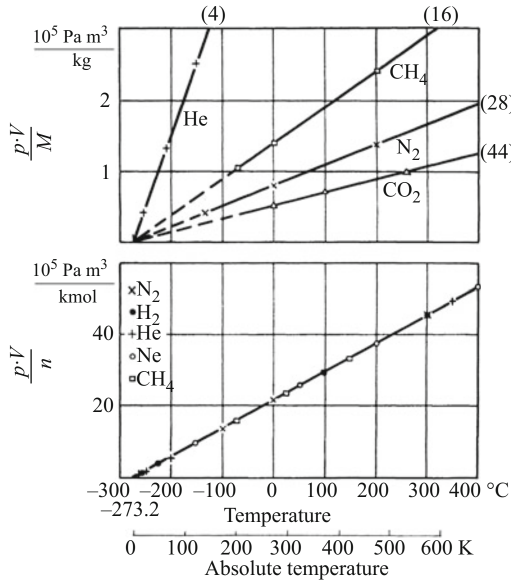


Figure 14.7 The equation of state of ideal gases and the definition of the absolute temperature. The small numbers in parentheses at the margins of the upper graph are the numerical values of the molar masses of the gases, M_m in g/mol (previously, they were called 'molecular weights', and were quoted as dimensionless numbers) (1 atm $\approx 10^5$ Pa.)

The experimental results shown in Fig. 14.7 allow us to define a temperature scale in "absolute" terms, i.e. *without negative values and practically independent of materials properties*. There are many possibilities for this definition. The one in use today goes back to Lord KELVIN: The point on the line in Fig. 14.7 (lower part) where the quotient pV/n was measured at the temperature of melting ice is associated with a temperature of 273.15 K. Any other value would have been equally permissible. But the value chosen by KELVIN has a great advantage: The units of the KELVIN scale are exactly the same as those of the CELSIUS scale. As we have already mentioned (Sect. 13.2), there is a simple expression relating the temperatures as measured on these two scales:

$$\frac{T}{\text{K}} = \frac{t}{{}^\circ\text{C}} + 273.15\,.$$

The KELVIN scale ('absolute temperatures') is likewise shown in Fig. 14.7. Temperatures measured on this scale give a simple form to the equation of state of ideal gases:

$$pV = nRT \qquad (14.19)$$

(n = amount of substance of the gas confined in the volume V; a numerical example is given in **Exercise 14.1**).

Other formulations are in common use and are expedient:

$$pV_\mathrm{m} = RT \tag{14.20}$$

and

$$p = \frac{\varrho RT}{M_\mathrm{m}} \tag{14.21}$$

($V_\mathrm{m} = V/n$ = molar volume, $M_\mathrm{m} = M/n$ = molar mass with M = the mass of the gas, ϱ = mass density)[C14.7].

C14.7. For the definition of "molar" quantities, see Comment C13.7.

The proportionality factor R is called the *universal gas constant*. It can be determined experimentally from the slope of the lines in Fig. 14.7. The result is

$$R = 8.31\,\frac{\mathrm{W\cdot s}}{\mathrm{mol\cdot K}}\,. \tag{14.22}$$

Equation (14.19), the equation of state of ideal gases, also contains the *amount of substance* n of the molecules confined within the volume V[C14.8]. If we insert $n = N/N_\mathrm{A}$ (Sect. 13.1) into Eq. (14.19) and at the same time use the abbreviation $R/N_\mathrm{A} = k$, then we obtain an additional formulation of the equation of state of ideal gases, namely

C14.8. The volume of an ideal gas with the amount of substance $n = 1$ mol can be readily computed from Eq. (14.19); under so-called "standard conditions" (or "normal conditions", i.e. $T = 273$ K and $p = 1.013\cdot 10^5$ Pa), it has the value $V = 22.4$ liter. This is the *standard molar volume* $V_\mathrm{m} = 22.4$ liter/mol. – thus, as long as one can apply Eq. (14.19), the amount of substance n or the molar mass $M_\mathrm{n} = M/n$ can be obtained very simply, by measuring the mass of the gas in a certain volume under standard conditions.

$$pV = NkT \tag{14.23}$$

(N = the number of molecules confined in the volume V).

The new constant k which appears here is thus

$$k = \frac{R}{N_\mathrm{A}} = \frac{8.31\,\mathrm{W\cdot s}}{\mathrm{mol\cdot K}}\cdot\frac{\mathrm{mol}}{6.022\cdot 10^{23}}$$

or

$$k = 1.38\cdot 10^{-23}\,\frac{\mathrm{W\cdot s}}{\mathrm{K}} \tag{14.24}$$

(N_A = AVOGADRO's constant = $6.022\cdot 10^{23}\,\mathrm{mol}^{-1}$).

This universal constant k is called BOLTZMANN's constant (the motivation for this name is given in Sect. 18.4).

14.7 Addition of Partial Pressures

According to Eq. (14.23), the pressure p in the volume V at a given temperature is independent of the *type of molecules* and depends only on their number N. This leads us to DALTON's law of the addition of partial pressures. We explain it by referring to Fig. 14.8:

Two different gases (which do not react chemically with each other) are confined in two equal-sized chambers at the pressures p_1 and p_2. A piston pushes the gas from the right chamber through a valve O into the left chamber, while the temperature is being held constant. – Result: In the left chamber, we now have a pressure of $p = p_1 + p_2$. The two pressures p_1 and p_2 add as "partial pressures" to give the total pressure p.

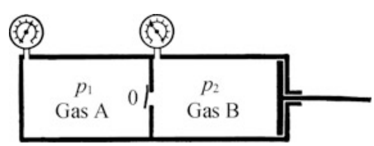
An example of DALTON's law: At the temperature of the human body, i.e. +37 °C, the air pressure at sea level (p = 1013 hPa) in the *lungs* of a person is composed of the following partial pressures[2]:

Gas	Nitrogen	Oxygen	Carbon dioxide	Water vapor
Partial pressure =	757	140	53.3	62.7 hPa

At an altitude of 22 km, the atmospheric pressure is only 62.7 hPa (see Fig. 9.32). At body temperature, the vapor pressure of water alone is just as high. As a result, the partial pressures of the other gases in the lungs drop to zero. The lungs of a person are thus filled only with water vapor and breathing is no longer possible. At still lower external pressures, the human body begins to boil, i.e. the vapor pressure of the water it contains becomes higher than the external air pressure.

Boiling means that bubbles of vapor form within a liquid. It occurs when the vapor pressure of the liquid becomes equal to the external pressure which acts on it, for example the pressure of the surrounding atmosphere. This, together with DALTON's law, leads to two surprising demonstrations:

Figure 14.8 Schematic of the addition of partial pressures



[2] The air in the lungs thus has a considerably higher concentration of CO_2 than the air outside. The ratio of the partial pressure of CO_2 to that of O_2 is nearly 0.4. At high altitudes, a person breathes faster and more deeply. Nevertheless, this ratio increases still further with increasing altitude, since the body produces just as much CO_2 at high altitudes as it does at sea level. One therefore cannot calculate the composition of air in the lungs at various altitudes merely according to physical criteria.

1. At normal atmospheric pressure, water boils at 100 °C, and carbon tetrachloride (CCl_4) at 76.7 °C. – We arrange these nearly mutually-insoluble liquids in two layers, with the heavier CCl_4 below, and heat them in a water bath; then boiling begins at the interface already at a temperature of 65.5 °C! – Reason: At this temperature, water has a vapor pressure of 256 hPa and CCl_4 has a vapor pressure of 757 hPa. The sum of these two partial pressures, according to DALTON's law, gives a total vapor pressure of 1013 hPa, equal to the ambient atmospheric pressure, so that the formation of bubbles and boiling can begin.

2. We immerse a test tube filled with air, with its open end downwards, in a flat dish full of diethyl ether ($(CH_3CH_2)_2O$). Immediately, air bubbles out of the opening! It is displaced by the partial pressure of the ether vapor (≈ 580 hPa at 20 °C).

14.8 The Caloric Equations of State of Ideal Gases. GAY-LUSSAC's Throttle Experiment

Besides the simple state variables p, V and T, we have defined other state variables (or *state functions*, also called *thermodynamic potentials*; for example the internal energy U and the enthalpy H). They in turn are functions of p, V and T. This dependence can be represented by *caloric* equations of state. In these, one of the three simple state variables can always be substituted by the other two. In general, caloric equations of state thus contain *two* simple state variables.

For example, in order to describe the dependence of the internal energy on the temperature, we make use of Eq. (14.17). It contains the dependence of the internal energy of an amount of substance n on its volume at a fixed temperature, which is the same before a process begins and when it is completed (independently of whatever changes take place during the process). We thus need the quantity

$$\left(\frac{\partial U}{\partial V}\right)_{T=\text{const}} \quad \text{i.e. the limiting value of} \quad \left(\frac{\Delta U}{\Delta V}\right)_{T=\text{const}} .$$

It has to be determined experimentally. For the measurement of ΔU, we can use the relation

$$\Delta U = Q + W . \tag{14.6}$$

For measurements on gases, we employ the apparatus which is sketched in Fig. 14.9: Two steel cylinders, *I* and *II*, are in a water-bath calorimeter (the thermometers, thermal insulation and stirrer are not shown in the figure). *I* contains air at high pressure (e.g. $152 \cdot 10^5$ Pa; under these conditions, air is still nearly an ideal gas), while *II* is evacuated. When the throttle valve between the two cylinders is

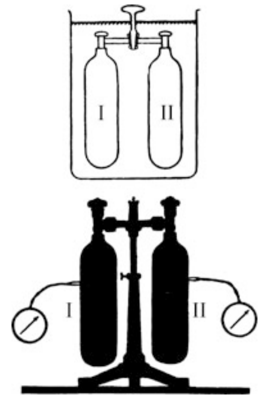


Figure 14.9 The throttle experiment of L.J. GAY-LUSSAC (1807): At constant temperature, the internal energy of an ideal gas is independent of its pressure and density. (Above: schematic; below: a demonstration experiment. Each cylinder has a volume of $V = 2$ liter, a mass of $M = 4.52$ kg and a heat capacity of 2093 W s/K)

Part III

opened, the pressure and density of the air decrease without any work being performed on the outside world (for short: external work). Such a decompression is called a *throttled release*. – When $W = 0$, Eq. (14.6) is simplified to $\Delta U = Q$. In words: A confined gas in a calorimeter is decompressed through a throttle. The gas is supposed to be at the same temperature after the decompression as before. In order for this to hold, a quantity of heat Q must be added to or removed from the calorimeter. Then we have $Q = \Delta U$, i.e. Q is equal to the change in the internal energy U of the gas decompressed through the throttle.

C14.9. This experiment is not shown as a demonstration. It is illustrated in the upper part of Fig. 14.9. Here, the whole setup is inside *one* calorimeter, which indicates no temperature change during the experiment, $\Delta T = 0$. In the actual demonstration experiment (lower part of the figure), we still see the same result due to its symmetric construction.

Experimentally, one finds $Q = 0$, and therefore $\Delta U = 0$[C14.9]. This means that the internal energy of the air did not change during its decompression. *The internal energy U of an ideal gas is independent of volume, pressure, and density at constant temperature*. As a formula:

$$\left(\frac{\partial U}{\partial V}\right)_{T=\text{const}} = 0\,. \qquad (14.25)$$

To understand the demonstration experiment, it is best to follow the process in detail. The cylinders themselves serve as calorimeters, each equipped with an electric thermometer. When the throttle valve is opened, the air in cylinder *I* is decompressed and expands. It produces a jet and performs work of acceleration. The equivalent quantity of heat Q is taken up from the walls of cylinder *I*, its temperature decreases by $-\Delta T_I$. The kinetic energy of the air jet is converted within cylinder *II* by turbulence and internal friction into internal energy. The temperature of *II* thereby increases by ΔT_{II}. In practice, one finds $-\Delta T_I = \Delta T_{II}$, in this example ≈ 7 K. The result is that the quantity of heat which is taken up by the air from cylinder *I* is equal to the quantity of heat which is released in cylinder *II*. Thus, also in this demonstration experiment, the net quantity of heat Q exchanged with the air is overall zero.

From the throttle experiment of GAY-LUSSAC, we can draw two conclusions:

1. In ideal gases, the internal energy U contains no *potential* energy which depends upon the spacing of the molecules. Therefore, in describing *ideal* gases, we can neglect the *forces* between the molecules and treat them as vanishingly small.

2. In ideal gases, the internal energy U depends only on the *temperature*, and thus not on two, but only on *one* of the simple state variables. As a result, in Eq. (14.12), we can leave off the condition $V = \text{const}$; similarly, we can omit $p = \text{const}$ in Eq. (14.13), since pV also depends only on T. Then we obtain

$$Mc_V = \frac{\partial U}{\partial T} \qquad \text{or} \qquad U = Mc_V T + U_0 \tag{14.26}$$

and

$$Mc_p = \frac{\partial H}{\partial T} \qquad \text{or} \qquad H = Mc_p T + H_0 \,. \tag{14.27}$$

Every energy is defined only with respect to an arbitrarily-chosen zero point; think for example of the potential energy of a stone which has been lifted to a certain height. Thus, we can agree to take U_0 and H_0, the zero points of the internal energy of an ideal gas and its enthalpy, to be their values at the temperature of absolute zero[3]. Then for ideal gases, we obtain the simple caloric equations of state

$$\text{Internal energy} \qquad U = Mc_V T \,, \tag{14.28}$$

$$\text{Enthalpy} \qquad H = Mc_p T \,, \tag{14.29}$$

or, for the corresponding molar quantities,

$$\text{Internal energy} \qquad U = nC_V T \,, \tag{14.30}$$

$$\text{Enthalpy} \qquad H = nC_p T \,. \tag{14.31}$$

It is important not to overlook the essential assumption underlying these expressions: In integrating the starting equations (14.26) and (14.27), c_p and c_V were taken to be *constant*.

The enthalpy H and the internal energy U differ by the quantity pV (Eq. 14.10). For an ideal gas containing an amount of substance n, we have $pV = nRT$. Then we find

$$n(C_p - C_V)T = nRT \,,$$

or

$$C_p - C_V = R \,. \tag{14.32}$$

[3] The magnitude of the constants U_0 (or H_0), i.e. the energy of a substance at absolute zero, is well known today. It is equal to the mass of the substance multiplied by the square of the velocity of light. U_0 is therefore very large; for example, for 1 mol of hydrogen, it is equal to $1.8 \cdot 10^{14}$ W s.

In words: *For every ideal gas, the difference of its two molar heat capacities is equal to the universal gas constant.*

Numerical example for nitrogen

$$C_p - C_V = 29.14 \frac{\mathrm{W\,s}}{\mathrm{mol\,K}} - 20.79 \frac{\mathrm{W\,s}}{\mathrm{mol\,K}}$$
$$= 8.35 \frac{\mathrm{W\,s}}{\mathrm{mol\,K}} \approx R\,.$$

14.9 Changes of State of Ideal Gases

In addition to the thermal and caloric equations of state for ideal gases, we must consider as a third case the equations for *changes* of state. Such changes can in general be represented in a p-V diagram, and the equations for the changes of state give the relation between two of the simple state variables. These are presumed to change *uniformly* within the volume of gas under consideration. There should be no local differences in temperature, pressure or density, such as might occur when the changes of state are very rapid. Unfortunately, these equations are sufficiently simple only in the limiting case of completely ideal gases. For these, one usually distinguishes five types of state changes:

1. *Isothermal changes of state*. They occur when the *temperature is held constant*. We have already encountered their equation under the name BOYLE's Law:

$$\frac{pV}{M} = \text{const}\,. \tag{9.11}$$

From it, by differentiating, we obtain

$$\frac{\mathrm{d}p}{\mathrm{d}V} = -\frac{p}{V} \tag{14.33}$$

and the isothermal compressibility

$$\frac{\mathrm{d}V}{V} \cdot \frac{1}{\mathrm{d}p} = -\frac{1}{p}\,. \tag{14.34}$$

The origin of pressure from the random thermal motion has also been treated in previous sections. The graphs of Eq. (9.11) are hyperbolas. One of these curves, which are called *isotherms*, is drawn in Fig. 14.10. A transition, e.g. from state 1 to state 2, that is an *isothermal expansion*, produces the *external work W*. *In this process, the internal energy U of the gas remains unchanged*. Therefore, the work W which is performed on the external system must be *replaced* by

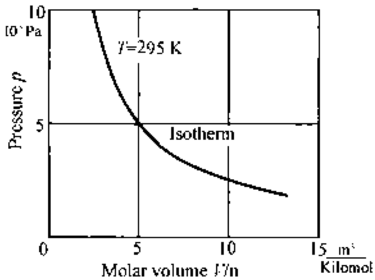


Figure 14.10 An isotherm at 22 °C

input of a quantity of heat Q. Quantitatively, both for the work of expansion and for the technical work, we find[C14.10]

$$W = W_{\text{techn}} = -Q = -nRT \ln \frac{V_2}{V_1} = -nRT \ln \frac{p_1}{p_2} . \quad (14.35)$$

C14.10. In general, the external work W and the technical work W_{techn} are different quantities. In the special case of isothermal changes of state of ideal gases treated here, however, they are equal. Both the internal energy U and the enthalpy H remain constant during such processes.

Derivation

$$W = -\int_1^2 p\,\mathrm{d}V, \quad p = \frac{nRT}{V}, \quad W = -nRT \int_1^2 \frac{\mathrm{d}V}{V} = -nRT \ln \frac{V_2}{V_1} .$$

In such a process then, the entire quantity of heat Q which is input to the amount of substance n of the ideal gas is converted into work. Conversely, in an isothermal compression, the work performed on the gas is completely converted into heat and given off by the gas.

2. *Isobaric changes of state*. They take place at *constant pressure*. Their equation is

$$\frac{T}{V} = \text{const.} , \quad (14.36)$$

i.e. the volume V increases proportionally to the temperature T (Fig. 14.11). The transition from the state 1 to the state 2 or *vice versa* is represented in a p-V diagram by a straight line parallel to the *abscissa*. In an isobaric *expansion*, for example, the amount of substance n of the gas performs the work

$$-W = p(V_2 - V_1) = nR(T_2 - T_1) . \quad (14.37)$$

During the expansion, the enthalpy of the gas increases by $\Delta H = nC_\text{p}(T_2 - T_1)$, and the corresponding quantity of heat must be input to the gas. The ratio of work performed and heat input is

$$\frac{-W}{Q} = \frac{nR(T_2 - T_1)}{nC_\text{p}(T_2 - T_1)} = \frac{R}{C_\text{p}} = \frac{C_\text{p} - C_\text{V}}{C_\text{p}} , \quad (14.38)$$

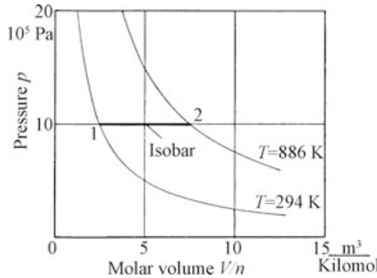


Figure 14.11 An isobar between two (lightly drawn) isotherms

or, with $\kappa = C_\mathrm{p}/C_\mathrm{V}$,

$$\frac{-W}{Q} = \frac{\kappa - 1}{\kappa} . \tag{14.39}$$

In an isobaric volume *decrease*, the corresponding quantity of heat must be *removed*.

3. *Isochoric changes of state*. These take place at *constant volume*. Their equation is

$$\frac{T}{p} = \text{const} . \tag{14.40}$$

Pressure and temperature in an isochoric change of state are proportional to each other. The transition, e.g. from a state 2 to the state 1 is represented by a straight line parallel to the *ordinate* of the p-V diagram (Fig. 14.12). Heat must be input to the gas; it is converted entirely to internal energy, giving an increase of

$$\Delta U = nC_\mathrm{V}(T_2 - T_1) . \tag{14.41}$$

Work is *not* performed, since the volume remains constant.

On cooling (the transition going from state 1 to the state 2), the internal energy is decreased and the corresponding quantity of heat must be removed.

Video 14.1: "Adiabatic changes of state"
http://tiny.cc/yggvjy
Using an air pump which has a small heat capacity, the temperature changes during compression and expansion are demonstrated qualitatively.

4. *Adiabatic changes of state* (**Video 14.1**). These occur *without exchanging heat with the surroundings* (i.e. thermally isolated), that is with $Q = 0$. They play an important role in physics and technology. On expansion of the gas, its pressure drops not only due to its *volume*

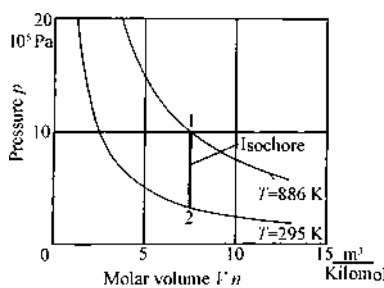


Figure 14.12 An isochore between two (lightly drawn) isotherms

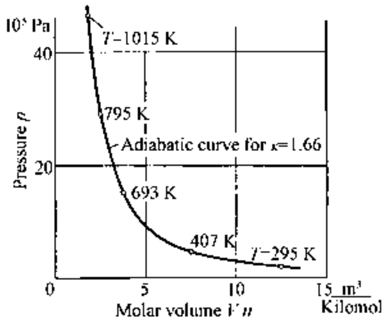


Figure 14.13 Adiabatic curve for a monatomic gas with $\kappa = 1.66$ (Table 14.1)

increase, but also due to the associated *cooling* of the gas. The corresponding curves are called *adiabatic curves* (Fig. 14.13) and fall off more *steeply* than a hyperbola. Their equation is given by[4]

$$pV^{\kappa} = \text{const} \tag{14.42}$$

(POISSON's Law).

Derivation

We refer to Fig. 14.14. We can replace the adiabatic expansion by an expansion from 1–3 at constant pressure (isobar) and a pressure reduction from 3–2 at constant volume (isochore). Along the path 1–3, that is the isobaric volume increase, the quantity of heat $\mathrm{d}Q_{1\text{-}3} = nC_\mathrm{p}\mathrm{d}T_{p=\text{const}}$ must be input to the gas to keep its pressure constant. Along the path 3–2, that is the isochoric pressure decrease, the quantity of heat $\mathrm{d}Q_{3\text{-}2} = nC_\mathrm{V}\mathrm{d}T_{V=\text{const}}$ must be removed from the gas to keep its volume constant. The sum of these two quantities of heat must be zero, since all together, we want to reproduce an adiabatic change of state, in which no heat is exchanged with the surroundings. We thus arrive at

$$C_\mathrm{p}(\mathrm{d}T)_{p=\text{const}} = -C_\mathrm{V}(\mathrm{d}T)_{V=\text{const}}\,. \tag{14.43}$$

The two temperature changes result from the thermal equation of state for ideal gases, i.e. from $pV = nRT$. We find

$$(\mathrm{d}T)_{p=\text{const}} = \frac{p\,\mathrm{d}V}{nR} \quad \text{and} \quad (\mathrm{d}T)_{V=\text{const}} = \frac{V\,\mathrm{d}p}{nR} \tag{14.44}$$

or

$$\frac{(\mathrm{d}T)_{V=\text{const}}}{(\mathrm{d}T)_{p=\text{const}}} = \frac{V\,\mathrm{d}p}{p\,\mathrm{d}V}\,. \tag{14.45}$$

Continuing, with Eq. (14.43) we obtain

$$\frac{\mathrm{d}p}{\mathrm{d}V} = -\frac{C_\mathrm{p}p}{C_\mathrm{V}V} = -\kappa\frac{p}{V}\,. \tag{14.46}$$

In words: *Along the adiabatic curves, the differential pressure change is by a factor* κ *greater than along an isotherm* (Eq. (14.33)).

[4] Instead of the volume V, one can also use the molar volume $V_\mathrm{m} = V/n$, if the amount of substance n of the confined gas is separated out of the constant.

Part III

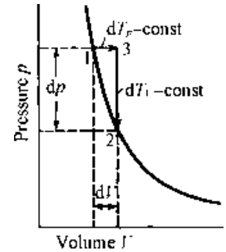


Figure 14.14 The derivation of the adiabatic exponent

From Eq. (14.46), we obtain by integrating:

$$\ln p + \kappa \ln V = \ln[\text{const}]$$

or

$$pV^{\kappa} = \text{const}\,. \tag{14.42}$$

Additional equations which are important for adiabatic changes of state can be found in the following Point 5.

5. *Polytropic changes of state*. They take place when *the thermal insulation is not sufficient to guarantee an adiabatic change of state*. During the expansion, the pressure falls due to the increasing volume and the associated cooling. Because of the insufficient thermal insulation, this cooling is however less than in a perfectly adiabatic expansion. As a result, the curve referred to as a *polytropic* (Fig. 14.15) falls off less steeply than an adiabatic curve. Its equation is

$$pV^{\alpha} = \text{const.} \tag{14.47}$$

Thus, when the thermal insulation is not perfect, one can*not* set the exponent α equal to κ; instead, a *smaller* value must be used, called the polytropic exponent. Then, instead of e.g. Eq. (14.46), we have: Along a polytropic curve, the differential pressure change is by a factor α greater than along an isotherm.

Making use of the equations

$$\text{und} \quad \left.\begin{array}{rcl} p_1V_1 &=& nRT_1 \\ p_2V_2 &=& nRT_2 \end{array}\right\}, \tag{14.19}$$

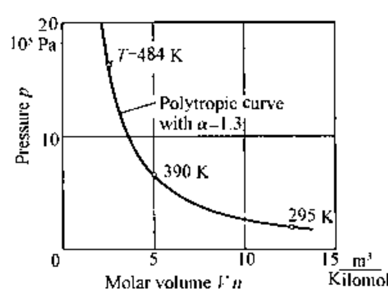


Figure 14.15 A polytropic curve for a diatomic gas ($\kappa = 1.4$)

along with Eq. (14.47), we obtain the following relations which are useful for applications:

$$\frac{T_2}{T_1} = \left(\frac{V_1}{V_2}\right)^{\alpha-1} = \left(\frac{p_2}{p_1}\right)^{\frac{\alpha-1}{\alpha}} \tag{14.48}$$

and for the *external work* which is performed during the expansion:

$$\begin{aligned} W &= -\frac{p_1 V_1}{\alpha - 1}\left[1 - \left(\frac{p_2}{p_1}\right)^{\frac{\alpha-1}{\alpha}}\right] \\ &= -n\frac{R}{\alpha - 1}(T_1 - T_2) = -\frac{p_1 V_1 - p_2 V_2}{\alpha - 1}\,. \end{aligned} \tag{14.49}$$

The technical work W_{techn} is in this case larger than W by a factor α:

$$W_{\text{techn}} = -\frac{\alpha}{\alpha - 1} p_1 V_1 \left[1 - \left(\frac{p_2}{p_1}\right)^{\frac{\alpha-1}{\alpha}}\right]. \tag{14.50}$$

For adiabatic *changes of state, in all of these equations we must replace α by $\kappa = c_\text{p}/c_\text{V}$.* Then, for the external work performed during an adiabatic expansion, we derive from Eq. (14.49), second term:

$$W = -nC_\text{V}(T_1 - T_2)\,. \tag{14.51}$$

Derivation
of Eqns. (14.49) and (14.50):

$$W = -\int_1^2 p\text{d}V = -\int_1^2 \text{const}\, V^{-\alpha}\text{d}V = -\text{const}\,\frac{V_2^{1-\alpha} - V_1^{1-\alpha}}{1-\alpha}\,. \tag{14.52}$$

Furthermore, from Eq. (14.47), we set the const $= p_1 V_1^\alpha = p_2 V_2^\alpha$, and from Eq. (14.19) we set $pV = nRT$, and rearrange to obtain Eq. (14.49). We then arrive at Eq. (14.50) by starting with Eq. (14.49) and making use of Eq. (14.5).

14.10 Applications of Polytropic and Adiabatic Changes of State. Measurements of $\kappa = c_\text{p}/c_\text{V}$

The changes of state described in Sect. 14.9 are important for numerous applications. Here, we can give only a few examples:

1. The *measurement of a polytropic exponent* α. In Fig. 14.16, air is confined in a glass vessel (V = several liters) at a small overpressure

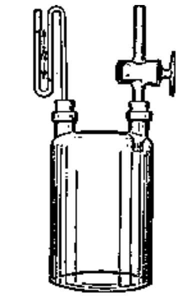


Figure 14.16 The measurement of a polytropic exponent α

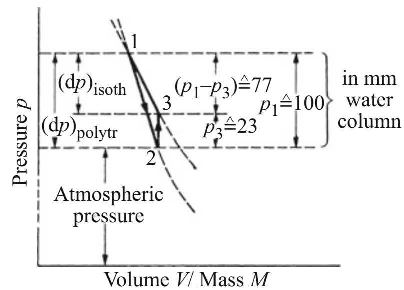


Figure 14.17 Measurement of the polytropic exponent as in Fig. 14.16 (Pressure units: 1 mm water column $\approx$ 10 Pa)

$p_1 = 10^3$ Pa ($\cong h = 100$ mm water column). The valve is opened and then immediately closed when the overpressure has been completely released. The expansion is *polytropic* (from Point 1 to Point 2 in Fig. 14.17), since the thermal insulation offered by a glass vessel is not perfect. The air inside is not as strongly cooled as it would be by an adiabatic expansion, i.e. with perfect thermal insulation. Nevertheless, less air leaves the vessel than if the expansion had been *isothermal*. As a result, the pressure *increases* (along the isochore from Point 2 to 3) as the air gradually warms back to room temperature. Again, an overpressure p_3 results; in the example, $p_3 = 230$ Pa ($\cong h = 23$ mm water column). We could have reached Point 3 immediately through a slow *isothermal* expansion. We would simply have to allow exactly the same amount of air to escape the vessel as in the case of the rapid polytropic expansion.

The pressure changes are small compared to the ambient air pressure (atmospheric pressure, i.e. 10^4 mm water column). Both the polytropic curves and the isotherms in Fig. 14.17 can therefore be approximated as short *straight* lines. From this figure, we read off

$$\begin{aligned} &\text{the polytropic pressure drop} \quad (-\mathrm{d}p)_{\mathrm{polytr}} &= \; p_1 \\ &\text{the isothermal pressure drop} \quad (-\mathrm{d}p)_{\mathrm{isoth}} &= \; p_1 - p_3\,. \end{aligned}$$

As found by computing $\mathrm{d}p/\mathrm{d}V$ in Eqns. (14.19) and (14.47), the ratio of the two pressure drops is the polytropic exponent α that we are seeking, and thus

$$\frac{\mathrm{d}p_{\mathrm{polytr}}}{\mathrm{d}p_{\mathrm{isoth}}} = \alpha\,.$$

In the example, $\alpha = \frac{100}{100-23} = 1.3$. The air thus expands in Fig. 14.17 with the polytropic exponent $\alpha = 1.3$.

2. *Measurement of the adiabatic exponent* $\kappa = c_p/c_V$ *from the velocity of sound*. With 'perfect' thermal insulation, the expansion in Fig. 14.17 is *adiabatic*. The measured exponent α must then become equal to the *adiabatic* exponent $\kappa_{air} = 1.40$. Indeed, it has often been attempted to measure κ in this manner. However, it is not simple to eliminate all possible perturbing heat inputs. – This is more readily accomplished with very *rapid* expansion processes. They can be found for example in *sound waves*, both propagating and standing waves. One can obtain κ with good precision from measurements of the velocity of sound. For the sound velocity in gases, we found

$$c_{\text{sound}} = \sqrt{\frac{K}{\varrho}}\,. \tag{12.46}$$

Here, ϱ is the density of the gas and K its modulus of compression:

$$\frac{1}{K} = -\frac{\mathrm{d}V}{V}\frac{1}{\mathrm{d}p}\,. \tag{14.53}$$

For an adiabatic expansion ($\alpha = \kappa$), we have (Eq. (14.46))

$$\frac{\mathrm{d}V}{\mathrm{d}p} = -\frac{1}{\kappa}\frac{V}{p}\,,$$

and thus

$$K = \kappa \cdot p\,. \tag{14.54}$$

Inserting Eq. (14.54) into Eq. (12.46) yields

$$c_{\text{sound}} = \sqrt{\kappa \cdot \frac{p}{\varrho}}\,. \tag{14.55}$$

Numerical example
At 18 °C and $p = 10^5$ Pa (1000 hPa), air has a density of $\varrho = 1.215\,\text{kg/m}^3$. The velocity of sound is measured to be $c_{\text{sound}} = 340$ m/s; from this we find $\kappa = 1.40$. The velocity of sound can best be determined using standing waves of a known frequency. ("KUNDT's dust figures", Sect. 11.7).

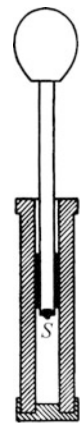
The velocity of sound c_{sound} *decreases with decreasing temperature*. This can be seen by rearranging Eq. (14.55) using the equation of state of ideal gases, (14.21)[C14.11]:

$$c_{\text{sound}} = \sqrt{\kappa \cdot \frac{R}{M_m} \cdot T}\,. \tag{14.56}$$

C14.11. The velocity of sound c_{sound} in an ideal gas is thus independent of its density ϱ!

3. *The production of high temperatures through polytropic compression*. Before European matches were introduced, the *fire pump* was in

Figure 14.18 A Malaysian fire pump, replicated here as a glass model. Instead of the tinder S, we could attach a piece of cotton wool dipped in carbon disulfide or diesel oil to the bottom of the piston. Then the heat of compression would cause the air-vapor mixture to burst into flames.



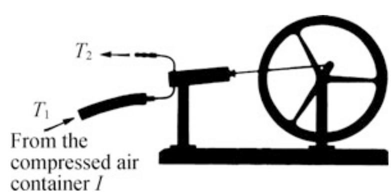
widespread use along the Malaysian Archipelago, especially in Borneo; it is often called a 'pneumatic lighter' (Fig. 14.18): A piston is thrust into a wooden cylinder. The air in the cylinder is heated, igniting a piece of tinder attached to the bottom of the piston. Today, diesel engines make use of the same process to ignite the fuel-air mixture which has been injected into the cylinder.

14.11 Pneumatic Motors and Gas Compressors

"The non-physicist often thinks that a cylinder of compressed air is an energy storage system, similar for example to the wound-up spring of a pocket watch. This interpretation is false".

These are not only technically important machines (e.g. the pneumatic jackhammer), but they are also very instructive from the point of view of physics. The non-physicist often thinks that a cylinder of compressed air is an energy storage system, similar for example to the wound-up spring of a pocket watch. This interpretation is false, for the following reason: Air is a nearly ideal gas, and therefore, the energy of a quantity of air at a constant temperature is independent of its pressure and density. Thus, when compressed air performs work by expanding isothermally, it gives up none of its own internal energy, but instead takes the energy from some other source. A compressed-air motor, as we shall see, is simply a machine which converts heat into work.

Figure 14.19 A compressed-air motor. For isothermal operation, the cylinder is surrounded by an electric heating mantle. **(Video 14.2)**



Video 14.2:
"Compressed-air motor"
http://tiny.cc/0ggvjy

Isothermal operation is exemplified by the cylinder of a small machine as shown in Fig. 14.19, when it is surrounded by an electric heating mantle whose heater current is controlled so that the air flowing into and out of the cylinder is kept at the same temperature, i.e. $T_2 = T_1$. Then the expansion of the air is *isothermal*, and we can apply Eq. (14.35) for the technical work performed:

$$W_{\text{techn}} = -Q\,.$$

(See the derivation in Sect. 14.9.)

In words: In isothermal operation, the heat Q which is input is completely converted into work, and the machine represents the ideal limiting case of 100 % mechanical performance[C14.12]. If the heating mantle is left off, the machine operates in a polytropic mode; in the limiting case of good thermal insulation it is adiabatic, and thus without heat exchange with the surroundings. In this latter limiting case, we find for the technical work

$$W_{\text{techn}} = \Delta H$$

(follows without computation from Eq. (14.11) for $Q = 0$).

C14.12. The "mechanical performance" quoted here does *not* contradict the Second Law. It should not be confused with the *efficiency* of heat engines. POHL discusses the compressed-air motor in this connection in more detail in Sect. 19.7.

In words: When a compressed-air motor is operating adiabatically, the work performed is equal to the decrease in enthalpy of the compressed air. The air leaves the machine cooler than when it enters; that is, $T_2 < T_1$, but this loss of enthalpy will be replaced afterwards by heat exchange from the surrounding atmosphere, so that T_1 is restored. In the adiabatic mode, the compressed air must merely lend some energy *in advance*.

The enthalpy reduction in a thermally insulated motor can be used to cool gases, e.g. for the liquefaction of helium. One then refers to 'cooling by performing work' or an 'expansion machine' (G. CLAUDE (1870–1960); cf. Sect. 15.6).

The operation of a compressor is just the reverse of that of the motor. When its cooling is not adequate, the machine driving the compressor must increase the enthalpy of the compressed air while compressing it, and the additional work performed for this is lost afterwards when the compressed air cools uselessly back to ambient temperature in its pressure vessel.

Exercises

14.1 The molar mass of oxygen is $M_\text{m} = 32\,\text{g/mol}$. $n = 2\,\text{kmol}$ of O_2 gas ($\hat{=}\ M = 64\,\text{kg}$) is at a temperature of 27 °C, i.e. $T = 300\,\text{K}$, in a volume $V = 300\,\text{liter} = 0.3\,\text{m}^3$. What is the pressure of the O_2? (Sect. 14.6)

Part III

14.2 A gas cylinder of volume $V_1 = 10\,\mathrm{l}$ contains oxygen at a pressure of $15 \cdot 10^5\,\mathrm{Pa}$. Another cylinder of volume $V_2 = 40\,\mathrm{l}$ contains nitrogen at a pressure of $8 \cdot 10^5\,\mathrm{Pa}$. What is the total pressure when the two cylinders are connected to each other? (Sect. 14.7)

14.3 To what fraction of its original volume must air be compressed in order to reach a temperature of $500\,°\mathrm{C}$? Assume that the process is polytropic, with a polytropic exponent of $\alpha = 1.36$ (Sect. 14.9).

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_14) contains supplementary material, which is available to authorized users.

Real Gases

15

15.1 Phase Changes of Real Gases

We now know the equation of state of ideal gases[C15.1], and that knowledge suffices to permit us to derive the equations of the various types of state transformations (isothermal, adiabatic etc.) without new experimental results. For *real* gases, there is no generally-applicable equation of state; therefore, to quantitatively describe their changes of state, we must rely on observations. The most important of these is the experimental determination of the *isotherms* of real gases. The isotherms exhibit the same qualitative behavior in all cases. For carbon dioxide (CO_2), they can be exhibited in a demonstration experiment with a modest effort. Figure 15.1 shows the setup, and Fig. 15.2 summarizes the results in a p-V_m diagram whose coordinate axes have been appropriately chosen.

C15.1. The terms "ideal gas" and "real gas" have nothing to do with the gas itself. As long as a gas obeys the equation of state of ideal gases, it is termed "ideal". If its behavior deviates from the predictions of that equation, e.g. in the neighborhood of its boiling point, then it is termed "real".

Part III

At temperatures above +80 °C, the isotherms are still *hyperbolic*. They can still be described by the equation pV_m = const. At +40 °C, already a strong deformation of the curve is seen. At +31 °C, the isotherm exhibits a *point of inflection* with a horizontal tangent: In the neighborhood of this *critical point K*, the pressure is independent of the volume of the confined gas. The state variables at this point are termed *critical*. For CO_2, the critical temperature is

$$T_{cr} = 304\,\mathrm{K}\ (31\,°\mathrm{C}),$$

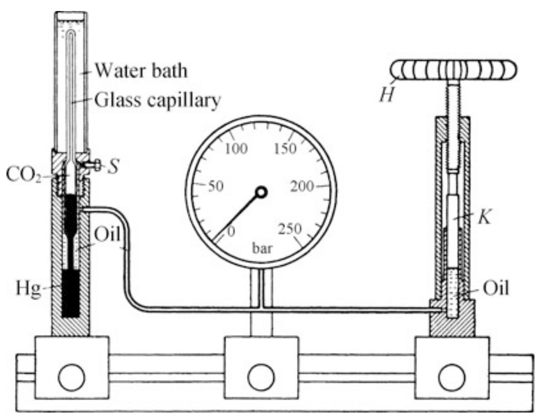


Figure 15.1 The investigation of phase changes (semi-schematic). S is used to fill the apparatus (1 bar = 10^5 Pa).

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_15

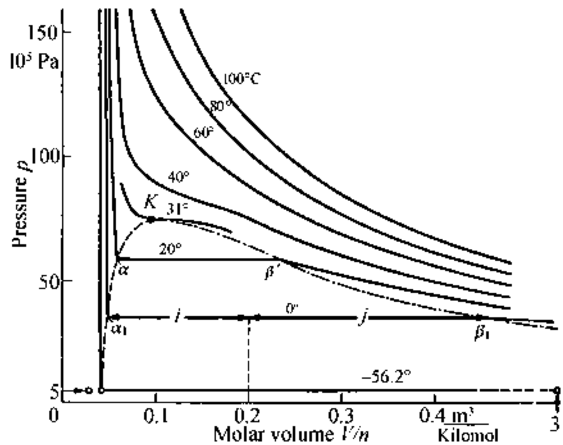


Figure 15.2 A p-V_{m} diagram of carbon dioxide with appropriate axes (Thomas Andrews, chemist in Belfast, 1813–1885). ($V_{\mathrm{m}} = V/n$ is the molar volume.) At 0 °C, the liquid has a molar volume of $V_{\mathrm{m,l}} = 0.048\,\mathrm{m^3/kmol}$ (abscissa value of point α_1) and the gas has a molar volume of $V_{\mathrm{m,g}} = 0.46\,\mathrm{m^3/kmol}$ (abscissa value of point β_1).

the critical pressure is

$$p_{\mathrm{cr}} = 73.6 \cdot 10^5\,\mathrm{Pa},$$

and the critical molar volume is

$$V_{\mathrm{m,cr}} = 0.096\,\mathrm{m^3/kmol}.$$

Examples for other substances can be found in Table 15.1.

Below the critical temperature, the phenomena become completely different. Let us follow the isotherm at +20 °C, beginning at a large volume, that is at the lower right: initially, the pressure increases up to the value $58.1 \cdot 10^5$ Pa at the point-β. When the volume is decreased further, the pressure remains constant (along the segment-$\beta - \alpha$ of the curve). Along this segment, the form of the carbon dioxide changes: An increasing fraction is separated from the rest by an *interface* or surface, i.e. it becomes *liquefied*. At the point α, all of the substance is in the liquid phase and an interface is no longer present. Continuing to decrease the volume requires that we apply an enormous pressure: Liquid CO_2 is considerably less compressible than gaseous CO_2 (steep rise of the isotherm!).

All the other isotherms below the critical point K exhibit a similar behavior. The end points of their straight, horizontal sections are delimited at the left by a dashed curve, and at the right by a dot-dashed curve, defining the *envelope of the coexistence region*. The two curves meet at the critical point K. They show the limits of the region in which the liquid and the gaseous phases can exist together. At the left of the dashed curve, there is only liquid; at the right of the dot-dashed curve, there is only gas. Above the critical point K, *distinguishing between gas and liquid is no longer meaningful*.

For the straight-line segments (the "average lines") of the isotherms, the abscissa values at their end points, e.g. α and β', give the molar volumes V_m of the liquid and the gaseous fractions (a numerical example is given in the caption of Fig. 15.2).

For each filling and temperature of a container (in the example sketched for $V_m = 0.2\,m^3/kmol$ and $0\,°C$), the ratio of the lengths j/i gives the ratio of the amount of substance as liquid/amount of substance as gas[C15.2]. – At the *critical filling*, where $V_{m,cr} = 0.096\,m^3/kmol$ and $T = 0\,°C$ (cf. Fig. 15.4),

89 % of the substance $\widehat{=}$ 45 % of the volume is liquid
11 % of the substance $\widehat{=}$ 55 % of the volume is gaseous.

The left-hand (dashed) limiting curve ends at a pressure of $5 \cdot 10^5$ Pa on a point marked by a small circle in the figure. Below this pressure, CO_2 is a solid. At the same pressure, a second circle is drawn in to the left, and a third is far to the right outside this region on the isotherm corresponding to $-56.2\,°C$ (at $V_m = 3\,m^3/kmol$). We will return to the significance of these circles when we deal with the triple point.

For water, today still the most important working substance, some special symbols and terms are in use. Water vapor in a state *outside* the limiting curves is called *superheated* steam, and *on* the limiting curve it is called *dry saturated* steam; *within* the limiting curve, it is termed *wet steam*.

Wet steam is a mixture of water vapor and fine water droplets. It appears to the eye as a white fog or as a white cloud. Superheated and saturated water vapor are just as invisible as for example the air in a room. They lack the fine suspended water droplets which scatter light (see Vol. 2, Chap. 26). – The non-physicist usually thinks of water vapor as this visible wet steam (fog).

Technically, the term *specific steam content* of wet steam is used. This refers to the ratio

$$x = \frac{\text{Mass of the dry saturated steam}}{\text{Mass of the steam and the water droplets suspended in it}} = \left(\frac{i}{i+j}\right) \qquad (15.1)$$

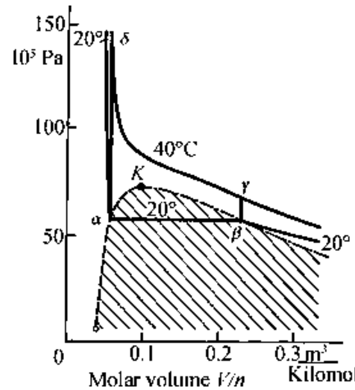
in Fig. 15.2. Along the left-hand limiting curve, $x = 0$; on the right-hand curve, $x = 1$.

C15.2. Derivation: The total amount of substance n is the sum of the fractional amounts of substance of the gas, n_g, and the liquid, n_l. At a molar volume of V_m, $\alpha_1 < V_m < \beta_1$ (Fig. 15.2), a part of the substance is gaseous, $n_g = xn$ ($0 < x < 1$) and the rest is liquid, $n_l = (1-x)n$. x is thus the molar fraction of the total amount of substance that is *gaseous*. Then we have for the combined molar volume: $V_m = x\beta_1 + (1-x)\alpha_1$ or $x = (V_m - \alpha_1)/(\beta_1 - \alpha_1)$, since α_1 and β_1 are the molar volumes of pure liquid and pure gas, respectively. For the example shown in Fig. 15.2, we find $x = i/(i+j)$ and $1 - x = j/(i+j)$, and thus $n_l/n_g = j/i$.

15.2 Distinguishing the Gas from the Liquid

The isotherms of CO_2 (Fig. 15.2) lead us to some important conclusions. Figure 15.3 shows only two of the isotherms, namely the

Figure 15.3 Distinguishing between gases and liquids: A cyclic process which includes the critical point K between two isotherms of CO_2 (the limiting curve at the left is dashed, at the right dot-dashed)



one at +20 °C and the one at +40 °C. In addition, the two limiting curves are drawn in and the region bounded by them is shaded. Within this region (the "coexistence region"), the gas and the liquid are *both* present. We begin with the state α and allow the volume to increase. The pressure remains constant, while an increasing fraction of the substance "vaporizes". At the state point β, all of the CO_2 has vaporized and no surface or interface remains. Now we keep the volume constant ($V_m = 0.227\,m^3$/kmol) and raise the temperature up to +40 °C (γ); thereafter, we compress the gas back to its original volume ($V_m = 0.057\,m^3$/kmol) isothermally (δ). In this process, the pressure increases up to about $150 \cdot 10^5$ Pa. From now on, we keep the volume constant and cool back to +20 °C, finally arriving back at our starting point, α. Result: We have not observed any *interface formation*, nor any fog, i.e. no precipitation of liquid CO_2 in the form of small suspended droplets. Nevertheless, all of the CO_2 has now again been liquefied. It exhibits a characteristic property of every liquid: It cannot be noticeably compressed even when the pressure (here at 20 °C) is increased to several hundred 10^5 Pa.

The whole closed path (cycle) can also be followed in the reverse direction, i.e. in the order α, δ, γ, β, α. Then we see no interface *disappear*, but nevertheless, beginning at the point β, we see a new interface form.

Result: In general, phase transitions take place with *discontinuous* changes of the physical properties, e.g. the transition (solid ↔ liquid): The two phases are separated over their entire range of existence by the limiting curve (Figs. 15.10 and 15.11). The phase transition (liquid ↔ gas), in contrast, takes place *continuously* below the critical temperature: In Figs. 15.10 and 15.11, the dashed curve *ends* at the critical point K.

A liquid cannot exist alone, i.e. in an otherwise empty space[1]. *The liquid surface is not a shell which secures it.* An interface forms sim-

[1] In outer space, liquids with a large mass could be held together by their mutual attraction (gravitation). But then they would always be surrounded by a gaseous atmosphere.

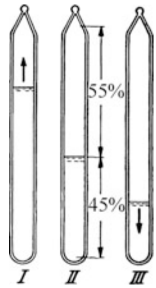


Figure 15.4 The liquid-gas phase transition in CO_2 with increasing temperature. In *I*, all of the substance is finally liquid (danger of explosion!), in *III* it is finally all gaseous. However, in *II*, the continuous transition at the critical temperature can be tracked; tube *II* illustrates the critical filling at 0 °C; cf. Sect. 15.1[C15.3].

C15.3. It is worth the effort to reconstruct these three phase transitions using the diagram in Fig. 15.2; this will lead to a more complete understanding of the diagram. The case of tube *I* at first seems to contradict all of our everyday experience!

ply as the boundary between two phases of the same substance. On its outer side, the *same* substance must be present in the form of *saturated vapor*, pure or mixed with another gas, for example ambient air. Only then is there an equilibrium, in which the same number of molecules move per unit time from the one phase to the other and *vice versa*. For water at room temperature, around 10^{22} molecules per second and cm^2 do this!! (cf. Sect. 16.1, Point 1). This statistical equilibrium prevents one phase from growing at the expense of the other.

At the end of Sect. 9.9, we encountered the diffusion boundary between two chemically different gases as a kind of interface or surface. With the same justification, we can now consider the surface of a liquid to be a *diffusion boundary*. It separates two chemically identical substances in physically different phases.

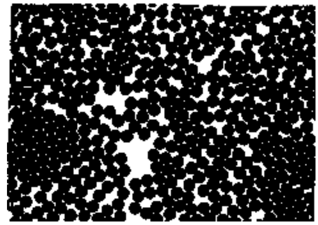
These results can be very clearly demonstrated with three equal-sized glass tubes filled with different amounts of CO_2 (Fig. 15.4). When the temperature is increased, the surface in tube *I* rises, in tube *III* it falls, and in tube *II*, at the critical temperature the surface approaches the center of the tube and vanishes there. That is, at the critical temperature, the molar volumes or specific volumes of the gaseous and liquid phases have become equal; the two phases are no longer distinguishable.

When the tubes are again cooled, the surface appears once again in the middle of tube *II*[2]. Its reappearance is announced by a *glimmering* layer of fog: In the statistical manifestation of the thermal motions, the phase transition appears first here, then there[3]. Sub-microscopic "nucleation centers" form through a local accumulation or accretion of molecules, and from them, tiny and at first volatile droplets are formed. Only at larger particle densities N_V of the droplets do they merge together to form a surface, i.e. a boundary interface between the two phases.

[2] Only there does the molar volume have precisely its critical value in the gravitational field of the earth; above the middle, it is larger; below, it is too small.
[3] At the critical point, $dp/dV = 0$ or $dV/dp = \infty$. That is, even minimal local variations in the pressure are sufficient to cause noticeable changes in the specific volume or its reciprocal, the density.

Part III

Figure 15.5 A two-dimensional model of the structure of a liquid with statistical density fluctuations



On approaching the critical temperature, there is certainly a continuous transition from the liquid to the *gas*. At lower temperatures, however, the liquid phase is much closer to the *solid* than to the gaseous phase. A liquid can almost be thought of as a very fine, microcrystalline powder: its micro-crystals have very short lifetimes, and the fragments of decaying micro-crystals reunite in an uninterrupted alternation to form new micro-crystals. – The following formulation puts this another way: *A liquid is a solid in a state of turbulence with very small but still crystalline turbulence elements*. As “higher-order individuals”, these are subject to mutual, progressive motions and rotations. In two dimensions, this can be convincingly simulated with steel-ball molecules in a flat dish. On shaking and stirring, we find “crystalline” regions that continually change their sizes and shapes, but always maintain hexagonal close packing (Fig. 15.5).

15.3 The VAN DER WAALS Equation of State for Real Gases

All of the isotherms shown in Fig. 15.2, with the exception of the straight-line segments in the coexistence region, can be represented to a good approximation by a third-order equation, the VAN DER WAALS equation of state. It is given by

$$\left(p + \frac{a}{V_\mathrm{m}^2}\right)(V_\mathrm{m} - b) = RT \tag{15.2}$$

($V_\mathrm{m} = V/n =$ molar volume of the gas).

In this equation, a and b are two constants which are characteristic of the particular type of molecules (i.e. the *substance*) that it describes. For every *ideal* gas, a *single* constant, namely R, suffices to determine its equation of state. For every *real* gas, by contrast, we require at least *three* constants. Some *numerical values* of a and b are set out in Table 15.1.

At the critical point, the isotherm in the p-V_m diagram runs parallel to the abscissa, and furthermore, it has a point of inflection there. Therefore, the

Table 15.1 Critical state variables and VAN DER WAALS constants for some real gases

		Critical state variables			VAN DER WAALS	
Substance	Molar mass	Temperature	Pressure	Molar volume	Constant	Constant
	M_m	T_{cr}	p_{cr}	$V_{m,cr}$	a	b
	$\left(\frac{kg}{kmol}\right)$	(K)	(10^5 Pa)	$\left(\frac{m^3}{kmol}\right)$	$\left(\frac{10^5\,Pa\,m^6}{kmol^2}\right)$	$\left(\frac{m^3}{kmol}\right)$
H_2	2.02	33	12.9	0.065	0.19	0.022
He	4	5.2	2.3	0.058	0.034	0.024
H_2O	18	647.4	221	0.055	5.54	0.031
N_2	28	126	34.1	0.090	1.36	0.039
O_2	32	154	50.4	0.075	1.37	0.032
CO_2	44	304	73.6	0.096	3.65	0.043
SO_2	64	430	78.5	0.096	6.84	0.056
Hg	200	≈ 1720	≈ 1080	≈ 0.040	0.82	0.017

following relations hold there:

$$\left(\frac{\partial p}{\partial V_m}\right)_{cr} = 0 \tag{15.3}$$

and

$$\left(\frac{\partial^2 p}{\partial V_m^2}\right)_{cr} = 0\,. \tag{15.4}$$

With these two conditions, from Eq. (15.2) we obtain

$$a = 3(V_m^2 p)_{cr} \tag{15.5}$$

and

$$b = \tfrac{1}{3}(V_m)_{cr}\,. \tag{15.6}$$

The values listed in Table 15.1 for the constants a and b have been chosen so that they fit the VAN DER WAALS equation of state (15.2) to the measured isotherms over the widest range possible. With the values of a and b fixed in this way, the condition (15.5) is well fulfilled, but the condition (15.6) only to a relatively poor approximation. For CO_2, for example, the measured value is $(V_m)_{cr} = 0.096\,m^3/kmol$; but in Table 15.1 we find $3b = 0.129\,m^3/kmol$. The VAN DER WAALS equation is just an approximation. Strictly considered, each different gas, due to the individual properties of its molecules, would require its own particular equation of state. We cannot demand too much of an equation that neglects these many individual properties! (**Exercise 15.1**)

The VAN DER WAALS equation of state differs from the simple ideal-gas equation by the additional terms a/V_m^2 and b. Their physical significance is readily seen. Let us start with the *internal pressure*, the term a/V_m^2. Under otherwise similar conditions, a lower pressure would be measured externally if the molecules were to experience a mutual attraction for each other. Therefore, in the equation of state for real gases, we must add a correction term to the measured pressure p when the molar volume V_m becomes small and the molecules are on average close together. The force that corresponds to the internal pressure a/V_m^2 acts in the same sense as the force applied externally by a piston (B in Fig. 9.24)[C15.4].

C15.4. Both forces are thus directed from the surface (container walls) into the interior of the gas.

Now to the term b. The equation $pV_\mathrm{m} = \text{const}$ was derived for a model gas. The volume in which the molecules can carry out their random thermal motions was taken to be the volume V of the whole container. When the molar volume of the *gas* becomes small, we can no longer neglect the proper volume of the *molecules themselves*, as we did for ideal gases. We must decrease the molar volume available to the *gas* by a quantity which is proportional to the molar volume $V_\text{m,molec.}$ of *its molecules themselves*. That is, instead of V_m, we write $V_\mathrm{m} - \text{const} \cdot V_\text{m,molec.}$ or, using the abbreviation $\text{const} \cdot V_\text{m,molec.} = b$: $V_\mathrm{m} - b$ (the factor 'const' is ≈ 4).

Part III

15.4 The Joule-Thomson Throttle Experiment

Now that we have knowledge of the van der Waals equation of state, we return to Gay-Lussac's experiment (Sect. 14.8).

From this fundamental experiment, it follows that we can expand a gas without changing its internal energy U. In this case, for Eq. (14.6), we find

$$\Delta U = Q + W = 0\,.$$

This can be the case when there is no exchange of thermal energy with the surroundings (adiabatic process, $Q = 0$) and furthermore no external work is performed (expansion through a throttle valve, $W = 0$)[4]. In this type of expansion, the temperature of the gas remains constant if it is an *ideal* gas.

For *real* gases, in contrast, under the same conditions there will be temperature changes ΔT. In principle, this can be demonstrated with the setup sketched in Fig. 14.9; but that has two disadvantages: First, it is relatively insensitive; and second, it is superfluous to first produce a gas jet with kinetic energy and then convert that energy by friction back into internal energy. In order to avoid both of these disadvantages, J.P. Joule and W. Thomson (later Lord Kelvin) replaced the expansion of a confined gas with its *internal energy* held constant by the expansion of a gas flow with its *enthalpy* H held constant, i.e. an *isenthalpic* process.

Their experimental arrangement, shown schematically in Fig. 15.6, corresponds to the general scheme of Fig. 14.4. In the Joule-Thomson experiment, the "machine" M is a reducing valve which prevents the formation of a jet, e.g. a narrow opening, or the fine channels of a porous disk. It is thermally well insulated. As a result,

[4] The other possibility, namely $Q = -W$, was already discussed in Sect. 14.11 (isothermally-operated compressed-air motor).

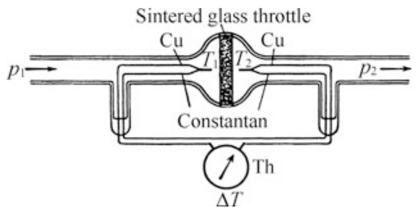


Figure 15.6 The expansion experiment of JOULE and THOMSON (1853). The throttle consists here of a porous sintered glass disk. It is fused into the glass walls of the apparatus. The two thermocouples are wired in opposition, so that the voltmeter *Th* indicates the *difference* ΔT of their two temperatures directly.

the gas cannot take up any heat from its surroundings; in Eq. (14.11), i.e. $Q + W_{\text{techn}} = \Delta H$, $Q = 0$. During the expansion, $W_{\text{techn}} = 0$ (no moving parts, no acceleration!), so that $\Delta H = 0$ and

$$H = U + pV = \text{const}\,. \tag{15.7}$$

As a result of the constancy of the enthalpy, the internal energy U, and with it the temperature, can increase or decrease simply by changing the quantity pV during the expansion. Some data from a measurement are shown graphically in Fig. 15.7. Usually, the expansion ($\Delta p < 0$, i.e. negative) causes a cooling effect ($\Delta T = T_2 - T_1 < 0$). However, in certain ranges of temperature, heating is observed, e.g. with air at a pressure of $220 \cdot 10^5$ Pa, above an "inversion temperature" of about 230 °C (Point b in Fig. 15.7)[C15.5]. The JOULE-THOMSON effect $\Delta T/\Delta p$ observed in particular cases must thus be the sum of two contributions, a cooling effect and a heating effect. Usually, the cooling predominates; but the heating effect can also be the larger of the two.

C15.5. The inversion temperatures of H_2 and He are ≈ 200 K and ≈ 50 K, respectively. In order to obtain cooling with the JOULE-THOMSON effect, these gases must therefore be precooled, for example by adiabatic expansion (see also Sect. 15.6).

In order to elucidate these two contributions, we start with the flow of an *ideal* gas. A gas containing the amount of substance n of mass M passes through the throttle. At the pressure p_1, we assume its volume to be V_1.

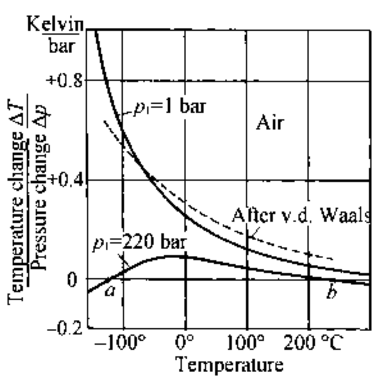


Figure 15.7 Measurements of the JOULE-THOMSON effect for air at various starting temperatures and pressures p_1. Outside the range a-b and at $p_1 = 220$ bar, heating occurs (1 bar $= 10^5$ Pa).

During its expansion, no work is performed; the prerequisite for performing *internal* work, i.e. an attractive force between its molecules, is lacking in an ideal gas, and there is no external work since the apparatus contains no moving parts and the molecules are not accelerated. The work of displacement performed *on the gas* at the left by the compressor, p_1V_1, is just the same as the displacement work performed *by the gas* at the right, $-p_2V_2$. Therefore, from Eq. (15.7), $U_1 = U_2$ and thus $T_1 = T_2$: The equation of state of *ideal* gases permits no change in the temperature on expansion without performing external work. The cooling or warming of *real* gases on expansion must therefore be connected with the additional *correction terms* in the VAN DER WAALS equation which describes them. The internal pressure a/V_m^2 which results from the mutual attraction of the molecules explains the *cooling* observed on expansion. As a result of the internal pressure a/V_m^2, the pressure of a *real gas* is *lower* than that of an ideal gas at the same number density N_V of its molecules. The greater the *compression*, the more the pressure is reduced relative to that of an ideal gas. Thus, the work of displacement p_1V_1 performed on the *compressed* real gas by the compressor is less than the work of displacement $-p_2V_2$ performed by the *expanded* gas. As a result, according to Eq. (15.7), $U_1 > U_2$ and $T_2 < T_1$. The gas leaves the throttle at a lower temperature, it is cooled.

This cooling can, however, be overcompensated by a heating effect which is due to the correction term b. *Due to the term* b, at the same number density N_V of the molecules, the pressure of a *real* gas is *higher* than that of an ideal gas. For the ideal gas, we have $p = \frac{1}{3}u^2/V_\mathrm{s} = \frac{1}{3}u^2 M_\mathrm{m}/V_\mathrm{m}$ (cf. Sect. 9.8); for a real gas, $p = \frac{1}{3}u^2 M_\mathrm{m}/(V_\mathrm{m} - b)$. Therefore, the work of displacement performed on the *compressed* real gas by the compressor, p_1V_1, is greater than the work of displacement performed by the *expanded* gas, $-p_2V_2$. Then from Eq. (15.7), $U_1 < U_2$ and $T_2 > T_1$. The gas leaves the throttle at a higher temperature, it has been heated. This heating may predominate over the cooling described in the previous paragraph.

The quantitative mathematical formulation of these considerations does not yield satisfactory results; in particular, it gives no dependence of the effects on the pressure, in stark contrast to the experimental results (Fig. 15.7). The VAN DER WAALS equation is, as we have already emphasized, only an approximation. We should not expect too much from its application.

15.5 The Production of Low Temperatures and Liquefaction of Gases

Some important *processes for the liquefaction of gases* are based on cooling by means of the JOULE-THOMSON effect, in particular for the liquefaction of air, hydrogen, and helium. Figure 15.8 illustrates a demonstration experiment showing the liquefaction of air. Dry air at a pressure of around $150 \cdot 10^5$ Pa flows through a tightly wound, multilayered copper coil within a transparent Thermos bottle ("Dewar vessel"). At its lower end is a fine nozzle, the *throttle*. The expanded and cooled gas can exit the Dewar vessel at its top. On the way out, it flows between the coils of the copper tubing and cools the gas which follows it, at the same time itself being warmed (SIEMENS[C15.6] coun-

C15.6. KARL WILHELM SIEMENS (Sir WILLIAM) (1823–1883), the brother of WERNER VON SIEMENS, studied chemistry, physics and mathematics in Göttingen.

Figure 15.8 A demonstration experiment showing liquefaction of air by the LINDE process. The cooled but not yet liquid fraction of the air flows upwards *between* the coils of the copper-tubing heat exchanger and escapes into the room. This precools the air which follows on its way into the vessel ("countercurrent heat exchanger"; see Sect. 17.6.). The copper tubing has an outer diameter of 2 mm and an inner diameter of 1 mm. The nozzle (throttle) consists of its flattened end. **(Video 15.1)**

Video 15.1:
"Liquefaction of oxygen"
http://tiny.cc/8ggvjy
The experiment becomes more straightforward when the flattened end of the copper tube in the video is replaced by an adjustable needle valve which can be controlled from above and which serves as throttle.

tercurrent process, Sect. 17.6). After a few minutes, the expanded gas has been cooled to its boiling temperature. Then its liquefaction begins at a constant temperature. One can observe a liquid emerging from the tube, at first as a fog, then as a continuous liquid jet; it quickly fills the lower part of the Dewar vessel. – In this process, only a small fraction x of the gas flowing into the apparatus is liquefied ($x \approx 0.1$). The greater portion ($1 - x \approx 0.9$) flows out of the vessel as gas, taking with it the enthalpy of condensation of the liquid produced.

We can consider the entire LINDE apparatus, that is the throttle and the countercurrent heat exchanger, to be an *isothermal* expansion machine: The gas flowing into the apparatus and the gas flowing out have the same temperature T. The gas flowing in, at a pressure p_1, brings its enthalpy $H_{\mathrm{T,p_1}}$ into the vessel, while the gas flowing out at the pressure p_2 takes its enthalpy $(1-x)H_{\mathrm{T,p_2}}$ with it. The enthalpy xH_{liquid} remains with the fraction x which is liquefied. Then we can write the enthalpy balance as

$$H_{\mathrm{T,p_1}} = (1-x)H_{\mathrm{T,p_2}} + xH_{\mathrm{liquid}}$$

and from it, we find the fraction x which is liquefied to be

$$x = \frac{H_{\mathrm{T,p_1}} - H_{\mathrm{T,p_2}}}{H_{\mathrm{liquid}} - H_{\mathrm{T,p_2}}}\,.$$

Numerical example for the liquefaction of air

$$\left(\frac{H}{M}\right)_{\substack{20\,^\circ\mathrm{C}\\ 200\,\mathrm{bar}}} = 4.634\cdot 10^5\,\frac{\mathrm{W\,s}}{\mathrm{kg}}\,,\quad \left(\frac{H}{M}\right)_{\substack{20\,^\circ\mathrm{C}\\ 1\,\mathrm{bar}}} = 5.028\cdot 10^5\,\frac{\mathrm{W\,s}}{\mathrm{kg}}$$

$$\left(\frac{H_{\mathrm{liquid}}}{M}\right)_{\substack{-193\,^\circ\mathrm{C}\\ 1\,\mathrm{bar}}} = 0.922\cdot 10^5\,\frac{\mathrm{W\,s}}{\mathrm{kg}}\,,\qquad \text{and thus}\quad x = 0.096 \approx 0.1\,.$$

Besides air liquefiers, there are today technically highly-developed liquefiers for hydrogen and helium. Those in use in modern laboratories yield many liters of liquid per hour[C15.7]. By reducing the pressure above liquid helium and thus lowering its boiling point, one can obtain $T \approx 1$ K. Temperatures between 1 and 0.4 K can be reached by

C15.7. See for example G.K. White, P.J. Meeson: *Experimental Techniques in Low-Temperature Physics*, 4th ed., Clarendon Press, Oxford (2002).

using the helium isotope[5] ^{3}He. It is circulated in a closed-cycle cryostat, alternating continuously between the liquid and the vapor phase (similarly to the cooling substance used in household refrigerators). – Still lower temperatures (down to 0.005 K) can be obtained not with ^{3}He *alone*, but rather by using mixtures of ^{3}He and ^{4}He isotopes[C15.8].

C15.8. The production of low temperatures by ^{3}He-^{4}He "dilution refrigeration", mentioned by POHL in the last editions, is very widespread today. More details can be found for example in the book by F. Pobell, *Matter and Methods at Low Temperatures*, Springer, 2nd ed. (1996).

15.6 Technical Liquefaction Processes and the Separation of Gases

The liquefaction of gases has considerable economic importance, as well as scientific interest. The various processes differ mainly in the way in which the gas is precooled. Often, precooling is accomplished by adiabatic expansion in a piston and cylinder machine or in a turbine: Then the work performed by the gas can be put to good use. – The last cooling stage prior to liquefaction is frequently implemented by using the JOULE-THOMSON effect, i.e. according to the scheme shown in Fig. 15.8. The yield of liquid air from all the different processes is roughly the same. It is about 1.33 liter/kWh (the maximum possible yield from an ideal process would be 5.3 liter/kWh).

Liquefied gases are indispensable cooling resources for technology and research today. Liquefaction on a technological scale is required among other things for the *separation* of gases, in particular for separating air into oxygen and nitrogen. Nitrogen is used for example in the synthesis of ammonia (fertilizers!) and oxygen for example in welding and for medical purposes. – The amount of work required for the separation of *air* is rather small in the ideal case, namely 0.014 kWh/m^3. It is required only in order to compress the two gases from their partial pressures up to atmospheric pressure.

Every gas separation is hampered by the thermal motions of the molecules. For this reason, one first cools the air until it liquefies and separates the gas mixture at a low temperature. A *transitory* cooling can in principle be carried out without expending work, if one uses a *countercurrent heat exchanger* to transfer the temperature (Sect. 17.6).

The actual separation process is known under the name of *rectification*. It is based on the fact illustrated in Fig. 15.9: At a given temperature, air (like many other mixtures of different substances) has a different composition in the liquid and the gas phases. In percentages of the amount of substance (mole percent), air at 83 K (its boiling point), for example, consists of:

$$\begin{array}{llll} \text{in the liquid phase:} & 65\,\%\,O_2 & \text{and} & 35\,\%\,N_2\,, \\ \text{in the gaseous phase:} & 37\,\%\,O_2 & \text{and} & 63\,\%\,N_2\,. \end{array}$$

[5] This rare isotope of helium became more widely available, and affordable, since the manufacture of its parent isotope, ^{3}H or tritium, was carried out on an industrial scale for military purposes. ^{3}He is currently again rather expensive.

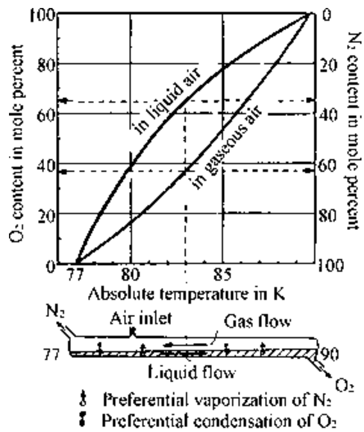


Figure 15.9 The separation of air into its component gases by 'rectification'. The gentle slope of the tube in the lower image indicates that the liquid current flow is maintained by gravity. Technical rectification columns are vertical, and the pure oxygen is drawn off as liquid at the bottom. The distribution of concentrations observed in such a plant is shown in the upper image; the difference in the vapor/liquid distributions of O_2 and N_2 near their boiling points is an example of their differing 'relative volatilities', often observed for distinct liquids, which can then be separated by *distillation*.

The essential point of a rectification process is shown in Fig. 15.9: currents of liquid and gas phases flow in intimate contact with each other in *opposite* directions through a tube *along which a temperature gradient is maintained*. The liquid flows in the direction of increasing temperature. Nitrogen, which boils at 77 K, evaporates preferentially from the liquid current, while oxygen, boiling point 90 K, condenses preferentially out of the gas phase. When the flow rate is sufficiently slow, an equilibrium is established at every point along the tube, corresponding to the local temperature at that point. For example, in a section of the tube which is at 83 K, the liquid and the gas phases have the different compositions marked by dashed arrows in the figure, as mentioned above.

In the technical implementation of rectification plants, care is taken in particular to provide a close contact and mutual mixing of the two oppositely-directed currents. The yield of pure *oxygen* from well-designed plants is about 2 m^3/kWh (much less than the ideally possible yield of 14 m^3/kWh).

15.7 Vapor Pressure and Boiling Temperature. The Triple Point

The p-V_m diagram of a substance (for example CO_2, as in Fig. 15.2) does not permit us to discern an important relation, namely the dependence of the vapor pressure on the temperature. This relation is best represented in a p-T diagram. An example for CO_2 can be seen in Fig. 15.10, and for water in Fig. 15.11. In both cases, the ordinate of the diagram is logarithmic, i.e. it shows increasing orders of magnitude.

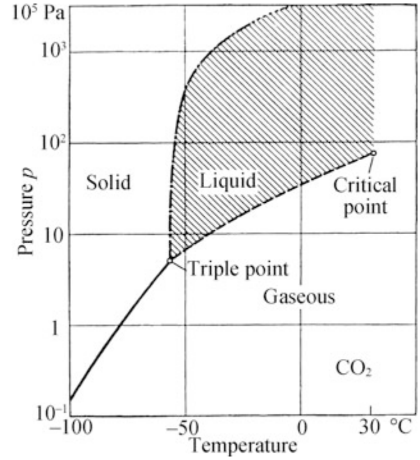


Figure 15.10 The phase diagram of CO_2. With a linear scale on the ordinate, instead of the logarithmic scale used here, the curves would rise very steeply.

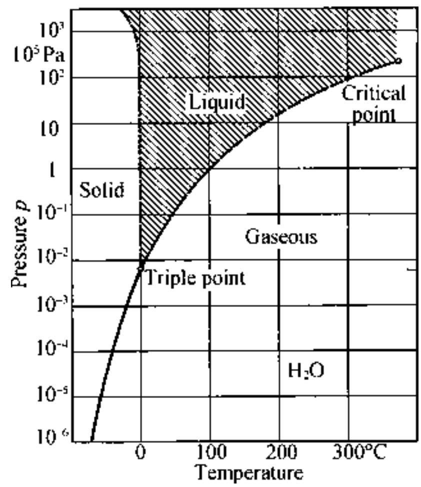


Figure 15.11 The phase diagram of water. At the triple point, all three curves intersect each other with *different* slopes. Compare Fig. 15.12.

These diagrams each contain three curves. Every point on a curve denotes a matched pair of the variables pressure and temperature. *Only* at these paired values can two phases of the substance *coexist* in a stable manner, that is in equilibrium with each other.

The dashed curve corresponds to the pressure which is required to liquefy the gas, known as the saturation pressure. The solid curve corresponds to the pressure of solidification of the gas, that is the formation of *frost*: the saturation pressure of the solid phase. Finally, the third, dot-dashed curve corresponds to the solidification of the liquid phase (**Video 15.2**).

Video 15.2:
"Liquid and solid nitrogen"
http://tiny.cc/ihgvjy
By "pumping off", liquid nitrogen can be cooled and frozen. Note that the density of the solid phase is greater than that of the liquid phase.

These two figures can also be rotated by 90° so that the pressure axis becomes the abscissa. Then for every value of the pressure, the dashed curve shows the corresponding *boiling point* of the liquid; the full curve gives the *sublimation temperature* of the solid phase; and the dot-dashed curve shows the *melting point* of the solid phase.

Part III

Below a pressure of $500 \cdot 10^5$ Pa, the melting temperature depends only weakly on the pressure. For CO_2, the melting temperature rises somewhat with increasing pressure; for water, it is lowered.

The three curves have one common point of intersection, the so-called *triple point*. The data for the triple point are:

$$\begin{aligned} &\text{for } CO_2: \quad T = 217\,\text{K}\,(-56.2\,°\text{C}), \quad p = 5 \cdot 10^5\,\text{Pa}\,, \\ &\text{for } H_2O: \quad T = 273.16\,\text{K}\,(0\,°\text{C}), \quad p = 611\,\text{Pa}\,. \end{aligned}$$

At the triple point – but only there – *all three* phases, solid, liquid, and gaseous, can exist together. They are in equilibrium; no one of the three phases can grow at the cost of the other two. In Fig. 15.2, we left the three points marked by small open circles without explanation. Their meaning is now clear: they correspond to the triple point. At this point, at a temperature of $T = 217$ K (-56.2 °C) and a pressure of $p = 5 \cdot 10^5$ Pa, the molar volume V_m

$$\begin{aligned} &\text{of solid } CO_2 \text{ is} \quad 0.034\,\text{m}^3/\text{kmol}, \\ &\text{of liquid } CO_2 \text{ is} \quad 0.041\,\text{m}^3/\text{kmol, and} \\ &\text{of gaseous } CO_2 \text{ is} \quad 3.22\,\text{m}^3/\text{kmol}\,. \end{aligned}$$

Away from the triple point, *at most two* phases can exist together; along the solid curve, only a solid phase and its saturated vapor. At pressures below 611 Pa, ice can no longer melt, it can only sublime (evaporate). Likewise, at atmospheric pressure (10^5 Pa), it is not possible to produce liquid carbon dioxide, only CO_2 snow (the well-known 'dry ice') at $T = 194$ K (-79.2 °C).

The production of dry ice is very simple (**Video 15.3**): The CO_2 cylinders which are commercially available have a pressure of about $50 \cdot 10^5$ Pa at room temperature; according to Fig. 15.10, they contain a mixture of liquid and gaseous CO_2. A thick cloth bag is attached to the cylinder valve and the valve is cautiously opened, so that gas from the cylinder flows out through the bag. As it flows out, the CO_2 gas forms a jet and performs work of acceleration (*external* work to accelerate the molecules in the jet). In addition, due to the JOULE-THOMSON effect, it also performs *internal* work, i.e. work against the mutual attractive forces between its molecules. For both reasons, the CO_2 cools until it reaches the temperature corresponding on the curve to the ambient atmospheric pressure of 10^5 Pa, i.e. -79°C (see Figs. 15.6 and 15.10, and Sect. 15.4).

Video 15.3:
"Solid carbon dioxide (dry ice)"
http://tiny.cc/lhgvjy
In this video, showing part of a lecture given by Prof. BEUERMANN, the preparation of dry ice is explained and demonstrated.

The contents of Figs. 15.10 and 15.11 form the basis of GIBBS' phase rule for a system containing a single substance: The number of freely disposable state variables is equal to 3 minus the number of phases which are in equilibrium. When *all three* phases are in equilibrium, *no* state variable is free; all three have their fixed values at the triple point. – When *two* phases of one substance are in equilibrium, only *one* of the two state variables p and T is freely disposable; the other is fixed on one of the curves in the p-T diagram. – In order to obtain only *one* phase of a substance, one can choose *both* state variables p and T arbitrarily; every pair p and T on the diagram is allowed; the values no longer have to remain on one of the curves.

15.8 Hindrance of the Phase Transition Liquid → Solid. Supercooled Liquids

The melting temperature of a crystalline solid at a given pressure is a characteristic and precisely-determined quantity for that particular substance. The melting temperature can*not* be exceeded without melting of the solid; that is, its outer layers become liquid. In the reverse direction, the situation is different: The melting point can be considerably *undershot* without passing through the phase transition liquid → solid: Liquids may be strongly *supercooled*.

If a boiling flask containing dust-free water is dipped into a cooling bath at −20 °C and shaken or stirred, avoiding splashing, the water can be readily cooled to around −10 °C. Smaller amounts of water, several tenths of a gram, can be supercooled down to −33 °C. The actual solidification temperatures are determined by the presence of various *submicroscopic impurities* which act as *nucleation centers*. The process of crystallization initiated at these centers always leads to the formation of hexagonal ice.

Tiny water droplets can even be cooled down to −72 °C without freezing. The water must be purified of impurities which could act as nucleation points for crystallization by repeated freezing and melting. At this temperature, the ice crystallizes in a *cubic* structure with a solidification point of −70 °C. – The vapor-pressure curve of a supercooled liquid is a continuation of the curve for the normal liquid without bends or kinks (cf. Fig. 15.12, an enlarged section from Fig. 15.11).

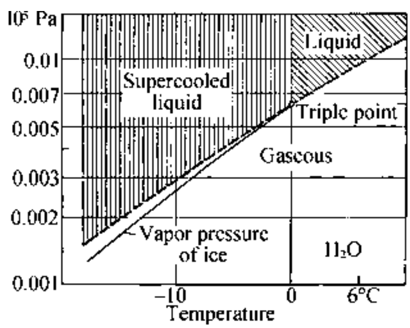


Figure 15.12 The vapor-pressure curve of supercooled water (*dashed*). For comparison, the vapor-pressure curve of ice is drawn in as a thin solid line.

15.9 Hindrance of the Phase Transition Liquid ↔ Gas: The Tensile Strength of Liquids

The phase transition *gaseous* → *liquid* can also be hindered by removing *nucleation centers*. Saturated vapors can be strongly supercooled, most simply by an adiabatic expansion. Here, too, the phase transition can be induced retroactively by introducing *nucleation centers*. Suitable for this purpose, among many other species, are *ions* of any type. At such nucleation centers, *surfaces* are formed from *supersaturated* vapors, leading to the precipitation of droplets as fog (This is applied for example in a *cloud chamber* used for the detection of ionizing radiation).

Furthermore, the phase transition from *liquid* → *gaseous* can also be strongly hindered by removing nucleation centers. For a demonstration, we fill a test tube which has been well cleaned using chromic acid with double-distilled water and heat it slowly in an oil bath. The water can reach a temperature of around 140 °C without boiling. It persists in a quiet state, with evaporation from its surface. Then, suddenly, within the water a turbulent transition to vapor sets in; the contents of the tube are ejected explosively. Boiling water can produce a genuinely dangerous situation. That is why, in practice, a *delay in boiling*, accompanied by 'bumping', should be avoided if possible. A simple protection lies in using 'dirty' vessels. A better method is to introduce small, angular objects ("boiling stones"), for example small shards of porcelain. The pores in their surfaces permit the stabilization of small gas bubbles (i.e. by preventing their reabsorption), which then serve as nucleation points.

To form such free surfaces within liquids, it is by no means always necessary to increase their *temperature*. One can also cause the liquid to *rupture*. This requires tensile forces of the order of $100 \cdot 10^5$ Pa (Sect. 9.5). Is there a prospect of increasing this value still further by removing perturbing nucleation centers? The VAN DER WAALS equation replies positively: Fig. 15.13 shows the isotherm for water at 300 °C in a p-V diagram. It was determined experimentally, like the two limiting curves $\alpha''K$ and $K\beta$. The isotherm $\eta\alpha'\beta'\zeta$ shows the normal process, where along the linear segment of the curve $\alpha'\beta'$, a *surface* is formed. In this region, *two* phases are present, and therefore, we can*not* apply the VAN DER WAALS equation here.

It is probable that the nucleation centers can be effectively removed so that the linear segment of the isotherms can be suppressed. If this is indeed possible, then only *one* phase will be present, and then we can apply the VAN DER WAALS equation also *between* the limiting curves. This is indicated in the curve segment $\alpha\gamma\varepsilon\delta\beta$. This represents the complete isotherm as calculated for 18 °C. It of course is an approximation, as is every application of the VAN DER WAALS equation. For example, the point α should lie on the left-hand limiting

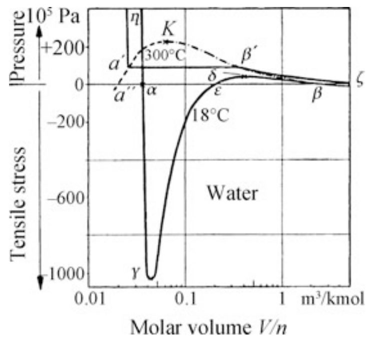


Figure 15.13 Computing the tensile strength of "nucleus-free" water by applying the VAN DER WAALS equation. – As a supplement to the text, we quote the following data: In "nucleus-containing" water, the linear segment $\alpha\beta$ of the 18° isotherm would fall practically along the axis of the abscissa, since the saturation pressure of water vapor at 18 °C is only 2.13 hPa. In order to show all the data in this figure, the abscissa has a logarithmic scale. Thus, the areas below the curve segment $\varepsilon\delta\beta$ and above the curve segment $\alpha\gamma\varepsilon$ are not equal, as they would be with a linear scale on the abscissa. – The curve segment $\gamma\delta$ cannot correspond to stable states. Along this segment, an increase in the molar volume leads to a decrease in the tensile strength.

curve at α''. – Nevertheless, one result is valid: The isotherm leads for certain values of the molar volume to *negative* pressures, that is to *tensile stresses* of over 10^8 Pa ($= 100\,\mathrm{N/mm^2}$ or ≈ 1000 bar). Therefore, nucleus-free water at 18 °C should exhibit a tensile or rupture strength of this order[C15.9] **(Video 9.4)**.

C15.9. Comparable to the tensile strengths of solids (see Table 8.2; see also Sect. 9.5)

The tensile strength of all liquids decreases with increasing temperature. It becomes zero at an upper limiting temperature. This experimental fact is also in agreement with VAN DER WAALS's equation: at $T = \frac{27}{32}\,T_{\mathrm{cr}}$, the isotherm intersects the V_{m} axis at a point γ, where the tensile strength thus vanishes.

Video 9.4:
"The tensile strength of water"
http://tiny.cc/3vqujy
(see Sect. 9.5).

Exercise

15.1 From the measurements shown in Fig. 9.23, it follows that air at temperatures above 0 °C and pressures up to 100 bar behaves like an ideal gas. This is to be quantitatively confirmed: a) Calculate the molar volume V_{m} of nitrogen at $T = 293$ K and $p = 2 \cdot 10^7$ Pa using the equation of state for ideal gases and compare it with the molar volume of crystalline (solid) nitrogen, with a density of $\varrho = 1.026\,\mathrm{g/cm^3}$ (this value was measured at $T = 21.15$ K). b) Show that the value of V_{m} calculated in a) also fulfills the VAN DER WAALS equation of state to a good approximation.

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_15) contains supplementary material, which is available to authorized users.

Heat as Random Motion

16

16.1 Temperature on the Molecular Scale

We have seen that the concept of random motion, for short "thermal motions", can be readily understood with the aid of model experiments. In their simplest versions, these replace the molecules by small, elastic steel balls (Sect. 9.7). Based on such model experiments, we derived the following equation for the pressure:

$$p = \tfrac{1}{3}\varrho\overline{u^2} \tag{9.14}$$

(ϱ = density of the gas, u = velocity of the molecular translational motions.)

For the experimentally-determined equation of state of ideal gases, we have already encountered the form

$$pV = NkT\,. \tag{14.23}$$

Furthermore,

$$\varrho = \frac{Nm}{V} \tag{9.15}$$

(N is the number of molecules in the volume V, m the mass of a single molecule, and k is the BOLTZMANN constant (Eq. (14.24))).

Combining Eqns. (9.14), (14.23) and (9.15) yields

$$m\overline{u^2} = 3kT\,. \tag{16.1}$$

$\overline{u^2}$ is the mean value of the squared velocity; the left side is thus twice the kinetic energy E_{kin} of *one* molecule. Then the energy can also be expressed as

$$E_{\mathrm{kin}} = \tfrac{3}{2}kT\,. \tag{16.2}$$

This equation states that *the average kinetic energy E_{kin} of each molecule of an ideal gas is proportional to the temperature and is independent of its chemical nature or its mass m*. Or conversely: *The temperature of a gas is determined by the kinetic energy E_{kin} of its*

Part III

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_16

molecules. – From Eq. (16.1), it follows that the velocity of the gas molecules has an average value of[1]

$$u_{\text{rms}} = \sqrt{\overline{u^2}} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3RT}{M_\text{m}}}\,. \tag{16.3}$$

($M_\text{m} = M/n =$ the molar mass[C16.1], where $M = Nm$ is the total mass of the molecules, and R is the universal gas constant (Eq. (14.22))).

C16.1. The "molar mass" M_m (see Comment C13.7 in Chap. 13) is usually denoted simply by M in the chemical literature. In order to avoid confusion with the mass in general, we will continue to use the index 'm' here. Remarkably, in most physics textbooks in use today, the molar mass is not mentioned at all!

We list here a few applications of these equations:

1. *The thermal motions of molecules during evaporation and expansion through a nozzle.* We refer to Sect. 9.8. – A gas volume V_1 contains N_1 molecules, each having a mass m. Then, in a time t, approximately $N = \frac{1}{6}\frac{N_1}{V_1} A\,ut$ molecules with the velocity u pass through an area A. The total mass Nm of these N molecules is $M = \frac{1}{6}\varrho A\,ut$, where $\varrho = \frac{N_1 m}{V_1}$ is the density of the gas. The equation of state of an ideal gas in the form of Eq. (14.21) at a pressure p gives for the density $\varrho = pM_\text{m}/RT$. Inserting these expressions into Eq. (16.3) yields the ratio $\frac{M}{A\,t} \approx 0.29\,p\sqrt{M_\text{m}/RT}$. A more precise calculation changes only the numerical factor; for the *mass-current density*, we obtain

$$\frac{M}{A\,t} = 0.4\,p\sqrt{\frac{M_\text{m}}{RT}} \tag{16.4}$$

($M =$ total mass of the molecules which pass through the area A within the time t).

Instead of the *mass* M, we could also refer to the *amount of substance* n of the molecules, or the *number of molecules* N, or their *volume* V. Then we obtain the corresponding current densities

$$\frac{n}{A\,t} = 0.4\,p\sqrt{\frac{1}{M_\text{m}RT}}\,, \qquad \frac{N}{A\,t} = 0.4\,p\,N_\text{A}\sqrt{\frac{1}{M_\text{m}RT}} \tag{16.5}$$

and

$$\frac{V}{A\,t} = 0.4\sqrt{\frac{RT}{M_\text{m}}} \tag{16.6}$$

(N_A is the AVOGADRO constant, $= 6.022 \cdot 10^{23}\ \text{mol}^{-1}$).

[1] For the velocity of sound c_sound, we have

$$c_\text{sound} = \sqrt{\kappa \cdot \frac{RT}{M_\text{m}}}\,. \tag{14.56}$$

Therefore, the velocity of the gas molecules is

$$u = c_\text{sound}\sqrt{3/\kappa} \approx c_\text{sound}\sqrt{2}\,.$$

Part III

Example
Water at 293 K (20 °C) and at its saturation pressure of $p = 2.32 \cdot 10^3$ Pa. $R = 8.31$ W s/(mol K), $M_m = 18$ g/mol. Inserting these values yields the particle current density

$$\frac{N}{A\,t} \approx \frac{10^{26}}{\mathrm{m}^2\,\mathrm{s}}.$$

At room temperature, every second around 10^{26} molecules escape from each m^2 of water surface, and the same number returns from the saturated vapor above the surface. – In 1 m^2 of surface area, only about 10^{19} water molecules are located within the surface layer (cf. Fig. 15.5). Therefore, the dwell time of an individual molecule at the surface is on the average only about 10^{-7} s[C16.2].

C16.2. This explains why water vapor in bubbles is quickly liquefied when the bubbles shrink, as we can conclude from the observations in **Video 9.4** (http://tiny.cc/3vqujy, see the explanation beside Fig. 9.16). We could also think of the bubbles which occur in cavitation and cause strong local heating, producing flashes of light (sonoluminescence, see Comment C9.12). This also leads us to suppose that the gas phase disappears rapidly.

Equation (16.6) leads to the ratio of the times t_1 and t_2 in which, at a given temperature and pressure, equal volumes of two different gases escape through an opening:

$$\frac{t_1}{t_2} = \sqrt{\frac{M_{m,1}}{M_{m,2}}} = \sqrt{\frac{m_1}{m_2}}. \tag{16.7}$$

From this relation, the molar masses M_m (or the molecular masses m) of different gases can be compared (R. Bunsen). Figure 16.1 shows a tried-and-tested experimental arrangement. The gas is confined at the left under the pressure p of a mercury column. At the upper end of the container is a small opening in a thin metal disk.

For demonstration experiments, we replace the small opening at the top of the tube by the porous wall of a clay cylinder (Fig. 16.2). At its bottom, a water manometer is connected to the cylinder, and a wide beaker is placed over it. Then for example hydrogen is blown into the beaker. The pressure within the clay cylinder rises sharply. The

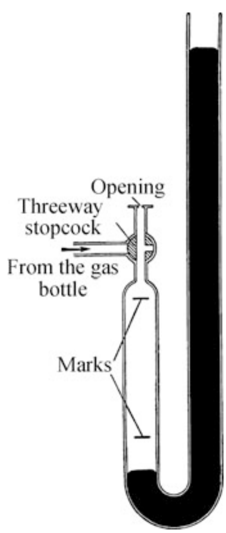


Figure 16.1 The comparison of molecular masses, after R. Bunsen. One measures the time in which the mercury (shown black in the figure) in the left-hand section of the tube rises from the lower mark to the upper one.

Figure 16.2 Diffusion through a porous clay cylinder

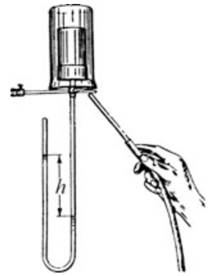
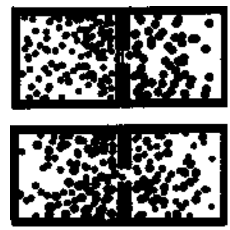


Figure 16.3 Demonstration of diffusion using two steel-ball model gases. Top: At the left, there are initially only small 'molecules', at the right only large ones. Bottom: The opening in the separating wall has been unblocked and diffusion has begun (see also **Video 16.1**)



Video 16.1:
"Model experiments on diffusion and osmosis"
http://tiny.cc/xhgvjy
With an apparatus whose principle is explained in Fig. 16.14, diffusion, osmosis and the BROWNian motion are illustrated with the model. The openings in the separating wall are large enough that the small balls can diffuse through, but not the large balls (a "semipermeable membrane").

reason: H_2 molecules diffuse in large numbers into the clay cylinder, about four times faster than the slower air molecules can diffuse out. After several seconds, the hydrogen flow is stopped and the beaker is removed. Very quickly, the overpressure in the clay cylinder is converted into a reduced pressure. The confined H_2 molecules diffuse more rapidly out of the cylinder than the air molecules which replace them can diffuse inwards.

Given the great importance of diffusion processes, a model experiment with our 'steel-ball gas' is quite appropriate. Figure 16.3 shows the setup, which we have already encountered in Fig. 9.24; however here, it is divided in the center by a wall with a small opening. Furthermore, on both sides there are oscillating pistons which maintain the artificial thermal motions. With static pictures in a book, we are limited to snapshots; they give only a pale impression of the lively effect of this demonstration experiment in action (**Video 16.1**).

2. *Temperature changes accompanying volume changes*. Every gas warms on compression and cools on expansion (Sect. 14.9). Reason: On expansion, the molecules are reflected from a wall which is moving away, and this reduces their velocities. On compression, the wall is moving inwards into the container, so that the molecules which are reflected from it experience an increase in their velocities. This can be clearly demonstrated with a single 'steel-ball molecule'. In Fig. 16.4, it is dropped from a height h onto a glass plate. It undergoes an elastic reflection and flies up again; at the same time, a second, smaller glass plate is moved downwards by hand. Now, the rising ball is reflected on a wall which is moving towards it; it flies back down with an increased velocity. This game is repeated several times, then the ball is allowed to fly up past the second glass plate and rises up to well above its original height h.

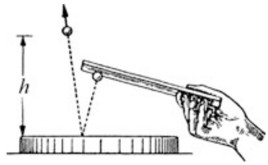


Figure 16.4 A model experiment on warming of a gas by compression (HARALD SCHULZE[C16.3])

C16.3. HARALD SCHULZE (1900-1997), student in the experimental physics lectures of the author. The experiment is mentioned already in the first edition of the "Thermodynamics" lectures (1941).

16.2 The Recoil of Gas Molecules Upon Reflection. The "Radiometer Force"

Figure 16.5 shows schematically a glass bulb which can be evacuated. It contains a small plate P, for example of aluminum foil or mica. The plate is mounted on a leaf spring and this serves as a force meter. – If one produces a temperature gradient between the two surfaces of the plate at a low gas pressure, then a force F acts on the plate in the direction of decreasing temperature. This phenomenon is called the *radiometer effect*. This somewhat misleading name is due to the fact that the temperature difference is usually produced by irradiating one side of the plate with light.

Especially suitable for demonstration experiments is the "light mill" or CROOKES' radiometer (Fig. 16.6). In it, four mica platelets, blackened with soot on one side, are mounted as the blades of a pinwheel. The wheel has a point and socket bearing at its hub, allowing it to rotate freely in a horizontal plane, as seen in Fig. 12.50. When irradiated with light, it always begins to rotate in the direction indicated

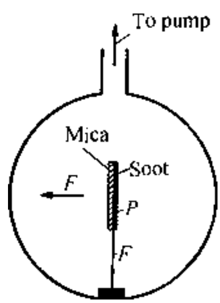


Figure 16.5 A schematic arrangement for detecting the radiometer effect, which results from the recoil of the reflected gas molecules. For quantitative measurements, a torsion band should be used instead of the leaf spring F shown here.

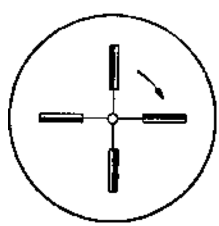
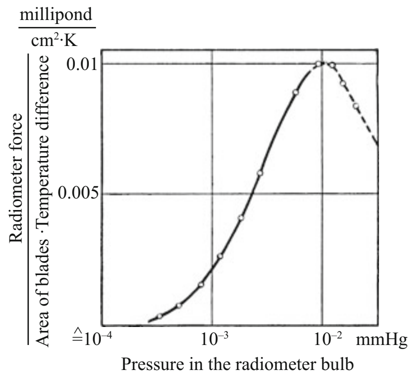


Figure 16.6 A horizontal section through a "CROOKES radiometer" (W. CROOKES[C16.4], 1874) **(Video 16.2)**

C16.4. Sir WILLIAM CROOKES (1832–1919). He studied gas discharges and discovered the element thallium.

Video 16.2:
"CROOKES radiometer (light mill)"
http://tiny.cc/qhgvjy

Figure 16.7 The dependence of the radiometer force on the gas pressure in the low-pressure range. A temperature difference is present between the two surfaces of the platelet. (Original measurement by W.H. WESTPHAL[C16.5], thus the old-fashioned units: 1 mmHg ≈ 1.33 hPa, 1 millipond = $9.81 \cdot 10^{-6}$ N).



C16.5. WILHELM HEINRICH WESTPHAL (1882–1978), from 1920 professor of physics at the University of Berlin, from 1935 at the Technical College in Berlin; author of several textbooks. See also W. Gerlach, *Z. Physik* **2**, 207 (1920).

by the arrow. Its rotational frequency increases with the intensity of the light.

In spite of its name, the radiometer effect has nothing to do with radiation, in particular not with the tiny radiation pressure of visible light. This can be most simply shown by using a sheet of mica which is blackened with soot on the side which faces *away* from the light source. The radiometer force is then directed *towards* the light source. Radiometers rotate even when the light is coming from all sides. The essential point is the difference in light absorption by the two surfaces of the blades; it produces a temperature gradient between the two sides and thus causes the radiometer force.

The radiometer force F depends in a characteristic way on the gas pressure p (Fig. 16.7). In the low-pressure range, the force increases proportionally to the pressure. In this range, the mean free paths of the gas molecules are large compared to the dimensions of the radiometer. The molecules undergo few collisions with each other; instead, they are reflected only by the radiometer blades and by the walls of the bulb. When the blackened surface has a higher temperature than the shiny surface, the molecules are reflected from it with a higher average velocity than from the cooler shiny surface. The reflected molecules produce a recoil on both sides; but it is greater on the warm side than on the cooler side. Thus, the resultant force F points in the direction of decreasing temperature. The force is proportional to the frequency of the collisions with the blade, and therefore to the gas pressure.

At higher gas pressures, the mean free path of the gas molecules is no longer long compared to the dimensions of the radiometer. Then, other phenomena can occur, and the force F *decreases* with increasing gas pressure (dashed segment of the curve in Fig. 16.7). The details would take us too far afield here. – The radiometer effects on small suspended particles or thin fibers are extremely strange. For example, small suspended carbon particles in the focus of the condenser of a projection lamp follow tiny spiral-helical orbits for hours; the orbits themselves form closed rings (photophoresis).

Part III

16.3 The Velocity Distribution and the Mean Free Path of the Gas Molecules

We now know two methods for determining the velocities of gas molecules experimentally (Sects. 9.8 and 16.1). Both methods are based on applications of momentum conservation and yield only average velocities. However, the distribution of the velocities around this average value can also be determined experimentally. To this end, one makes use of *molecular beams*. In Fig. 16.8, a small block of silver *Ag* is vaporized from an electrically-heated molybdenum crucible. The highly-evacuated glass vessel is equipped with two slits, *A* and *B*. They select a sharply bundled beam from the silver atoms which are flying in all directions out of the crucible. This beam collides with the cooled wall *W*. There, the atoms form a sharply-bounded, highly reflective spot. Its shape corresponds to the path of the beam shown as a dashed line in the figure. Above the second slit *B*, the glass walls of the vessel remain free of precipitated silver. To measure the velocities of the silver atoms, the whole apparatus is mounted on a rapidly-rotating turntable, whose rotation axis is perpendicular to the plane of the paper in Fig. 16.8. We then have the same situation as described in Chap. 7 for the measurement of the velocity of a bullet. The silver atoms are deflected to one side by the *Coriolis force*, and from the degree of deflection, we can compute their velocities (Sect. 7.3, Point 5).

Carrying out this measurement with a similar apparatus yields the results as shown in Fig. 16.9 for nitrogen at two temperatures: A distribution of the velocities over a wide range is observed.

The distribution over a wide range of velocities can be represented by the "distribution law" derived by MAXWELL. It gives the fraction dN/N of

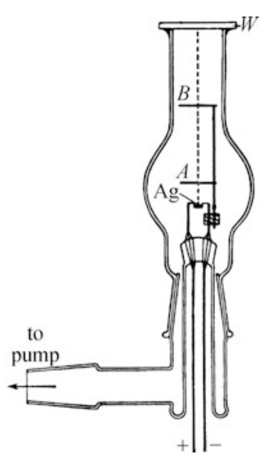


Figure 16.8 The production of molecular beams. 'Ag' is the silver which is being evaporated from a Mo crucible[C16.7].

C16.7. OTTO STERN used the experiment in Fig. 16.8 to determine the average thermal velocity of Ag atoms (*Zeitschrift f. Physik* **2**, 49 and **3**, 417 (1920)). The velocity *distribution* as in Fig. 16.9 was measured with a more advanced apparatus using Cs atoms, described in his Nobel lecture (1946) and by J. Estermann, O.C. Simpson, and O. Stern, *Phys. Rev.* **71**, 238 (1947).

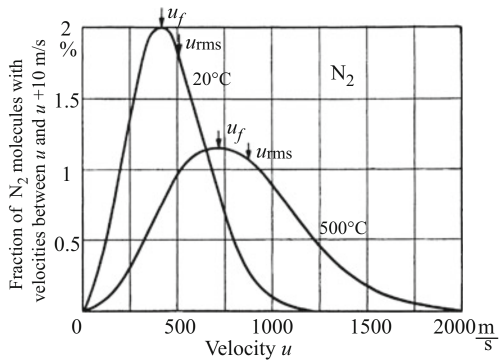


Figure 16.9 The velocity distribution of gas molecules using the example of nitrogen (molar mass $M_m = 28$ g/mol; thus at $T = 293$ K (20 °C), $u_f = 417$ m/s and the root mean squared velocity is $u_{rms} = 1.22\,u_f = 509$ m/s). The velocities are marked by arrows.

Part III

the molecules whose velocities lie between u and $(u + du)$. MAXWELL's *distribution law* is given by[C16.6]:

$$\frac{dN}{N} = \frac{4u^2}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{\frac{3}{2}} e^{-\frac{\frac{1}{2}mu^2}{kT}}\, du \tag{16.8}$$

C16.6. Often called the MAXWELL-BOLTZMANN distribution. For its derivation, see e.g.: W. Nolting, *Statistische Physik*, 3rd ed., Vieweg Braunschweig/Wiesbaden (1998), Exercise 1.3.7. English: see https://en.wikipedia.org/wiki/Boltzmann_distribution

A detailed examination of this equation shows that the maximum of the curves in Fig. 16.9 corresponds to the most *frequently-occurring* or most *probable* velocity:

$$u_f = \sqrt{\frac{2kT}{m}} = \sqrt{\frac{2RT}{M_m}}\,. \tag{16.9}$$

Taking the arithmetic mean of all the velocities, we obtain the *mean* (average) velocity

$$u_m = \frac{2}{\sqrt{\pi}} u_f = 1.13\,u_f\,. \tag{16.10}$$

It is thus somewhat higher than the most probable velocity.
Initially, we introduced the rms ("root-mean-square") value u_{rms} of the velocity, defined by the equation

$$u_{rms} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3RT}{M_m}}\,. \tag{16.3}$$

This rms velocity u_{rms} is thus higher by a factor of $\sqrt{\frac{3}{2}} = 1.22$ than the most probable velocity u_f, and $\frac{1.22}{1.13} = 1.08$ times higher than the mean velocity u_m. – The differences between the most probable, the mean, and the rms velocities are thus of no practical importance.

Our molecular picture of a gas is now nearly complete. Only the concept of the *mean free path* is still lacking. This is the name given to the straight-line path of a molecule between two successive collisions with other gas molecules. – Nitrogen, to mention one example, has a mean free path of $l \approx 6 \cdot 10^{-8}$ m at 0 °C and 1013 hPa, roughly 20 times larger than the average molecular separation under these conditions. Experiments for determining l will be described in Sect. 17.10.

16.4 Molar Heat Capacities in a Molecular Picture. The Equipartition Principle

The molar heat capacities of ideal gases can be understood in a molecular picture as follows (A. NAUMANN,[C16.8] 1867): We write the corresponding defining equations (14.12) and (14.13), as well as Eq. (14.32) for the two molar heat capacities

C16.8. ALEXANDER NAUMANN, Gießen, *Ann. der Chemie* **142**, 265 (1867); see also the book by the same author: *"Lehr- und Handbuch der Thermochemie"*, Verlag Vieweg (1882), Chap. 9. NAUMANN was not aware of the rotational degrees of freedom and interpreted their contribution to the specific heat as molecular vibrations ("heat of atomic motions"). – The rotational degrees of freedom (zero for single atoms, two for diatomic and three for polyatomic molecules) became understandable only after RUTHERFORD showed in 1911 that the masses of atoms are concentrated in their atomic nuclei! (see the footnote on the next page).

$$C_\mathrm{V} = \frac{1}{n}\left(\frac{\partial U}{\partial T}\right), \tag{16.11}$$

$$C_\mathrm{p} = \frac{1}{n}\left(\frac{\partial H}{\partial T}\right) = C_\mathrm{V} + R\,. \tag{16.12}$$

For an ideal *monatomic gas*, that part of the internal energy U which is due to random motion is mainly the *kinetic* energy E_kin of the *linear* motions of the molecules; briefly, their kinetic energy of translational motion. A single molecule makes a contribution of

$$E_\mathrm{kin} = \frac{3}{2}kT = \frac{3}{2}\frac{nRT}{N}\,. \tag{16.2}$$

A gas with N molecules or the amount of substance n gives the contribution

$$U = NE_\mathrm{kin} = \frac{3}{2}nRT\,. \tag{16.13}$$

Therefore, from Eqns. (16.11) and (16.12), we find

$$C_\mathrm{V} = \frac{3}{2}R\,, \quad C_\mathrm{p} = \frac{5}{2}R\,, \quad \frac{C_\mathrm{p}}{C_\mathrm{V}} = 1.67\,.$$

A monatomic molecule has three *degrees of freedom* (of translational motion). That means that its velocity along its linear path (translation) consists in general of three components, one in each of the directions of three-dimensional space. *Each one of these three degrees of freedom of a single molecule contributes the kinetic energy*

$$E_\mathrm{kin} = \frac{1}{2}kT = \frac{1}{2}\frac{nRT}{N} \tag{16.14}$$

(k = BOLTZMANN constant $= 1.38 \cdot 10^{-23}$ W s/K, $R = 8.31$ W s/(mol K));

Part III

Table 16.1 Molar heat capacities ($R = 8.31$ W s/(mol K))

Type of molecule	Examples	Molar heat capacities		
		C_p	C_V	$\kappa = C_p/C_V$
		at constant		
		Pressure	Volume	
monatomic	$\left\{\begin{matrix}\text{Hg vapor}\\ \text{noble gases}\end{matrix}\right\}$	$\frac{5}{2}R$	$\frac{3}{2}R$	1.67
diatomic	$\left\{\begin{matrix}H_2,\ O_2,\ N_2\\ \text{CO, HCl}\end{matrix}\right\}$	$\frac{7}{2}R$	$\frac{5}{2}R$	1.40
polyatomic	$\left\{\begin{matrix}CH_4,\ NH_3\\ CO_2\end{matrix}\right\}$	$4R$	$3R$	1.33

or, as the contribution of an amount of substance n of a gas containing N molecules, we find for its kinetic energy:

$$E_{\mathrm{kin,N}} = NE_{\mathrm{kin}} = \frac{1}{2}nRT\,. \tag{16.15}$$

A diatomic molecule has the shape of a dumbbell. It can rotate around two axes, each perpendicular to the other and to the long axis of the dumbbell[2]. This contributes two additional degrees of freedom. These new degrees of freedom are considered to be equivalent to the translations; this is called the *principle of statistical equipartition*. Thus a diatomic gas has all together five degrees of freedom. They contribute a kinetic part to the internal energy of a gas containing the amount of substance n, given by

$$U = \frac{5}{2}nRT\,. \tag{16.16}$$

As a result, from Eqns. (16.11) and (16.12), we find for *diatomic gases*:

$$C_V = \frac{5}{2}R\,, \quad C_p = \frac{7}{2}R\,, \quad \frac{C_p}{C_V} = 1.40\,.$$

Tri- and polyatomic molecules have three degrees of freedom for their rotations. Together with their three degrees of freedom for translation, one thus finds for *tri- and polyatomic gases*

$$C_V = \frac{3+3}{2}\cdot R = 3R\,, \quad C_p = 4R\,, \quad \frac{C_p}{C_V} = 1.33\,.$$

Table 16.1 summarizes these results for several gases. They agree well in general with the measured values as shown in Table 14.1.

[2] The rotation around the long axis itself need not be considered, since its moment of inertia is vanishingly small.

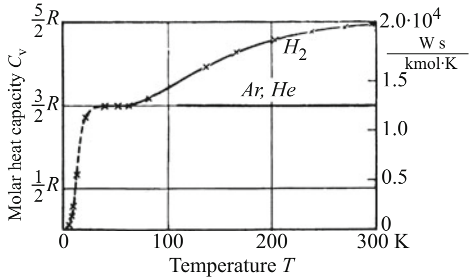


Figure 16.10 The molar heat capacity of hydrogen as a function of the temperature (R is the gas constant $= 8.31$ W s/(mol K)). For comparison, the results for two monatomic gases are also shown. In the dashed region, H_2 is liquid and solid (cf. Fig. 14.5).

In *solids*, the molecular building blocks have fixed rest positions. They cannot participate in progressive motions. Instead, the atoms can vibrate around their rest positions. For the kinetic energy of these vibrations, there are three degrees of freedom. By themselves, they would give $C_p = \frac{3}{2}R$[C16.9]. To maintain thermal equilibrium, in the solid we must add an equal amount of potential energy of vibration. This second energy contribution leads to $C_p = \frac{3}{2}R + \frac{3}{2}R = 3R$. This explains the limiting value of $C_p \approx 3R$ which is found for most solids built out of many individual atoms. (This is the rule of DULONG and PETIT[C16.10], cf. Fig. 14.5.)

C16.9. For solids, there is practically no difference between C_p and C_V, as long as nonlinear effects can be neglected, i.e. as long as the temperature is not too high ($T \leq 300$ K). Since measurements are always carried out at constant pressure, only C_p is given here in the text.

These successes gave experimental support to the equipartition principle. But in any case, it must be considered to be only an idealization of the *limiting case*, permissible in the range of high temperatures. This can be seen from the measurements given in Fig. 16.10. They refer to the molar heat capacity of a diatomic gas (H_2) at various temperatures. At high temperatures, C_V has a value of $\approx \frac{5}{2}R$; but as the temperature is lowered, it decreases and finally reaches a value of $\frac{3}{2}R$, i.e. the value expected for monatomic molecules. – Interpretation: As the temperature decreases, the rotations gradually come to rest; only translational motions remain, as for monatomic molecules. – At this point, we can go no further with the methods of "classical physics". We instead must turn to quantum mechanics[C16.11].

C16.10. A.T. Petit, P.L. Dulong, *Ann. Chim. Phys.*, 2nd Series, Vol. 10 (1819), p. 395.

C16.11. The same is true of the lattice vibrations in solids. For a discussion of the temperature dependence of the specific heat as seen in Fig. 14.5, see e.g. R.O. Pohl, *Am. J. Phys.* **55**, 240 (1987).

We summarize the essentials of Sects. 16.1 through 16.4: The ideal gases gave us the opportunity to interpret the state variables *temperature* and an important part of the *internal energy* in a graphic manner. The directly measurable state variables temperature and pressure arise from the random thermal motions of the molecules (Sect. 16.1). They are found to be statistical averages over an enormous number of individuals (molecules or atoms). We can make only *statistical* statements about the individual molecules. According to the equipartition principle, we can say: *On statistical average, each* *molecule makes the following contribution to the internal energy U for each of its degrees of freedom at a sufficiently high temperature T:*

$$E_{\text{kin}} = \frac{1}{2}kT \tag{16.14}$$

where k is the BOLTZMANN constant, $1.38 \cdot 10^{-23}$ W s/K. An experimental determination of k will be described in Sect. 16.6. The following section has the goal of preparing us for that task.

16.5 Osmosis and Osmotic Pressure

Osmosis originally meant "diffusion through porous membranes". If two substances are separated by a wall through which they can diffuse at different rates, then *temporarily*, a pressure difference will arise. This phenomenon is best known with two gases (Fig. 16.2). A similar experiment can also be carried out with two liquids. Example: We immerse a glass vessel filled with alcohol (CH_3CH_2OH) and closed at the bottom with the membrane from a pig's bladder into a dish containing pure water. The membrane will bulge outwards (I. A. NOLLET, 1748). In this case, also, the pressure difference is only *temporary*. Osmotic phenomena can also be observed when diffusion between a solution and the pure solvent takes place. To show this, we can make the wall or membrane *semipermeable*, i.e. penetrable by the solvent and impenetrable by the solute. Diffusion then leads to a *steady* pressure difference between the solution and the solvent. *This phenomenon is what today is exclusively called* osmosis.

Semipermeable barriers are realized in their most complete and manifold forms as *living* cell membranes. The best *artificial* production remains the membrane described in 1867 by MORITZ TRAUBE, made of *copper ferrocyanate*. Its preparation is simple: One adds for example a droplet of concentrated copper sulfate solution onto the surface of a dilute solution of yellow potassium ferrocyanate ($K_4Fe(CN)_6$). Then a skin-like bubble of copper ferrocyanate forms on the surface. *It quickly bulges out due to its uptake of water*; the surrounding solution becomes more concentrated due to the loss of this water, and sinks to the bottom of the container as a result of its greater density, forming visible striations (shadow projection; cf. Fig. 16.11).

With a suitable experimental setup, this puffing up takes place in a preferred direction. For example, one can throw a few small crystals of ferrous chloride ($FeCl_2$) onto the bottom of a cuvette filled with a potassium ferrocyanate solution (30 g $K_4Fe(CN)_6 \cdot 3\,H_2O$ in 1 liter of water). In the course of a half-hour, plant-like formations will grow from the crystals up towards the surface (Fig. 16.12).

Experiments of this kind and their quantitative extensions can be generically described on the basis of Fig. 16.13. There, we see two

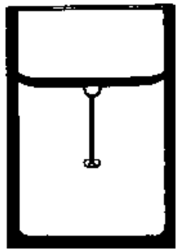


Figure 16.11 Puffing up of a membrane in a solution as a result of osmotic pressure. The bubble which is hanging under the surface of the solution appears bright in the shadow projection. At the lower end of the striation which is sinking downwards, we can see a vortex ring.

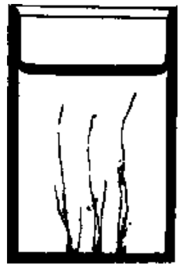


Figure 16.12 Osmotic pressure can produce plant-like formations

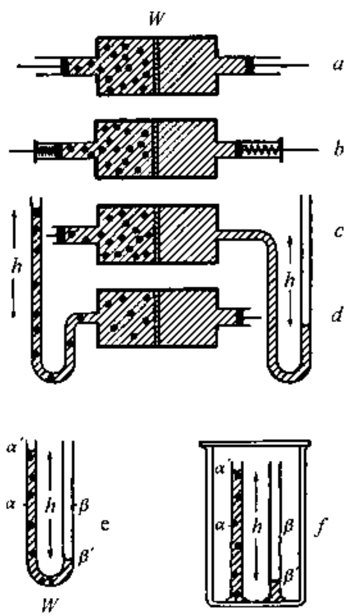


Figure 16.13 Volume increase of a solution due to osmotic pressure

chambers, separated by a semipermeable partition W. The chambers are each equipped with a piston of cross-sectional area A. They contain a solvent (shaded), e.g. water, and the left-hand chamber in addition contains dissolved molecules of a solute (black dots). This arrangement is *not* in equilibrium: Both pistons are moving to the left; the left-hand piston is being pushed out and the right-hand piston is being pulled in.

Explanation: *The dissolved molecules behave qualitatively like a gas*; their thermal motions produce a pressure, called the "osmotic pressure" p_{os}. This osmotic pressure pushes out the left piston and thereby increases the volume of the solution in the left chamber. *This volume increase by the osmotic pressure* is possible only when water from the right chamber can *flow through* the membrane, pulling the right-hand piston inwards.

The osmotic pressure p_{os} can be measured in two ways (Fig. 16.13, Part *b*): Either we let the left-hand piston *press* on a spring force meter, or the right-hand piston *pull* on a spring. In both cases, a force F results which hinders the motion of the left-hand piston[3] After either of the springs has been deformed by a sufficient amount, the volume increase of the solution stops, and $F/A = p_{os}$, so that a mechanical equilibrium has been established.

By the addition of springs, each of the two pistons has been converted into a *manometer*. – The simplest form of manometer is still a *liquid-column manometer*. In it, the piston is replaced by a free surface, and the force of the spring by the weight of the column of liquid. We could thus use either one of the setups sketched in Parts *c* and *d* in Fig. 16.13. In both, the final displacement h of the manometer indicates the osmotic pressure to be measured.

One can also combine the setups sketched in Parts *c* and *d* and design the chambers to have the same diameter as the manometer tubes. Then we arrive at a simple U-tube, which is divided by a semipermeable partition W at its lowest point. The solution in the left-hand chamber rises from α to α', while the water level on the right falls from β to β' (Part *e* of the figure).

In many cases which are of practical importance, the semipermeable partition is itself *movable*. Such cases can be readily demonstrated by a model experiment (Fig. 16.14). Little steel balls represent water molecules, while larger balls represent the solute molecules. The outer walls are vibrating pistons; they produce the random thermal motion of the model molecules. The partition between the two chambers has holes like a sieve through which the small 'molecules' can pass. A spiral spring provides a *rest position* in the middle for this "semipermeable" partition. If only small molecules are present, the partition remains in its center position (Fig. 16.14, left side). If larger solute molecules are added to the left-hand chamber, their "osmotic" pressure pushes the partition to the right (Fig. 16.14, right side): The left-hand chamber increases its volume due to the pressure of the larger molecules. In both images, one can begin with an arbitrary distribution of the small molecules, e.g. all of them in the left-hand or the right-hand chamber; they will always reproduce the same equilibrium distribution after some time.

[3] *Water has a sizeable tensile strength* (Sect. 9.5); it can thus be treated like a rigid connection between the two pistons. Compare also Sect. 15.9.

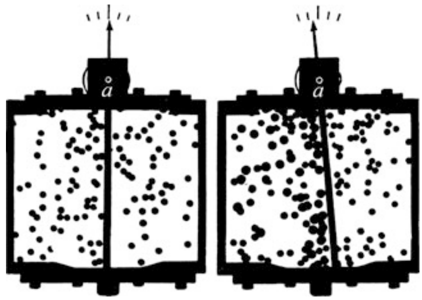


Figure 16.14 A model experiment showing the origin of osmotic pressure. On each side of the axle *a*, we can see the outer coil of a spiral spring. The motion of the semipermeable partition is damped by an oil-filled shock absorber (not shown) **(Video 16.1)**.

Video 16.1: "Model experiments on diffusion and osmosis" http://tiny.cc/xhgvjy

This model experiment explains for example the behavior of red blood cells in pure water. They puff up due to their semipermeable, elastic skins. Finally, the skin bursts. To prevent this, one must never replace serious blood losses by injecting pure water into a blood vessel; instead, a "physiological solution" must be used, which has the same osmotic pressure as that found in the interior of the red blood cells ($p_{os} = 7$ bar, corresponding to a NaCl solution with a concentration of $c = 0.16\,\mathrm{kmol/m^3}$)[C16.12].

C16.12. Correctly known as "normal saline" (see S. Awad, S.P. Allison, and N. Lobo Dileep (2008), "The history of 0.9 % saline", *Clinical Nutrition* (Edinburgh, Scotland) **27** (2), pp. 179–188). Every molecule of NaCl adds two ions to the solution, so that the total ion concentration is $0.32\,\mathrm{kmol/m^3}$. 0.9 % refers to the mass concentration, $9\,\mathrm{kg/m^3}$ NaCl in water (see also Comment C16.14.). For comparison: Seawater contains about 3 % NaCl.

The essential point in all osmotic phenomena is the volume increase of the solution. It always occurs when the solute molecules cannot leave the solution, but the solvent can enter it. This condition can be fulfilled even without a visible semipermeable partition; one can for example use a vacuum as a "semipermeable" region. This is the case e.g. in an *isothermal distillation*.

In Fig. 16.13, Part *f*, we see an evacuated vessel containing two measuring cylinders. The left one contains the solution, e.g. LiCl in water; the right cylinder contains the solvent, e.g. pure water. Initially, the two cylinders are filled to the same height; the liquid surfaces were at α and β. In the course of a few days, the volume of the solution increases, and the levels become different (a foolproof demonstration experiment!). Their final stationary height difference is denoted by h. Then the pressure equivalent to a liquid column of height h is equal to the osmotic pressure p_{os}, i.e.

$$p_{os} = h\varrho_S g \tag{16.17}$$

(ϱ_S = density of the solution, g = acceleration of gravity).

Interpretation: Above the two liquid surfaces is saturated vapor. It consists only of molecules of the solvent, but the vapor pressure p_S above the solution is *lower* than the vapor pressure p_0 above the pure solvent (demonstration experiment in Fig. 16.15). As a result, in Fig. 16.13, Part *f*, more water molecules strike the surface at α than can evaporate from it; that is, water is distilled from β to α.

When does this process of distillation come to an end? Answer: The vapor pressure p_0 of water corresponding to the temperature of the observations applies to the vapor *directly* above the *water surface*. As the column becomes higher, the pressure decreases, according to the barometric pressure

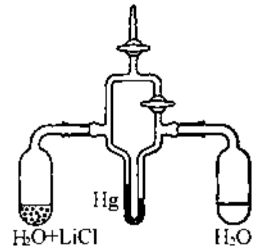


Figure 16.15 The vapor pressure of a solution is less than that of the pure solvent. The upper stopcock is used to pump out the air from the apparatus, while the lower one can be opened to equalize the pressures on both sides.

formula. At a height h above the water surface, it is reduced to

$$p_\text{h} = p_0 e^{-\frac{\varrho_0 g h}{p_0}} \tag{9.20}$$

(ϱ_0 = density of the water vapor above the water surface).
As the height h increases, the pressure p_h at some point becomes equal to the pressure above the surface at α', the vapor pressure p_S of the solution; then the same number of molecules fall onto the surface at α' as are evaporating from it in a given time. The volume increase then comes to an end.
Making use of this result, we wish to derive the relation between the osmotic pressure and the vapor pressure. We set $p_\text{h} = p_\text{S}$, and furthermore, from Eq. (16.17), we take $h = p_\text{os}/\varrho_\text{S} g$. Then we obtain

$$\frac{p_\text{S}}{p_0} = e^{-\frac{\varrho_0 p_\text{os}}{p_0 \varrho_\text{S}}} . \tag{16.18}$$

Furthermore, we treat the water vapor to a good approximation as an ideal gas and set

$$p_0 = \frac{\varrho_0 RT}{M_\text{m}} , \tag{14.21}$$

thus obtaining

$$\ln \frac{p_\text{S}}{p_0} = -\frac{\varrho_0 p_\text{os} M_\text{m}}{\varrho_0 RT \varrho_\text{S}}$$

or

$$p_\text{os} = \frac{\varrho_\text{S} RT}{M_\text{m}} \ln \frac{p_0}{p_\text{S}} \approx \frac{\varrho_\text{S} RT}{M_\text{m}} \cdot \frac{p_0 - p_\text{S}}{p_\text{S}} . \tag{16.19}$$

(The approximation is valid, since p_0/p_S is only slightly greater than 1, so that in the series expansion $\ln x = x - 1 + \ldots$, the higher terms can be neglected.)
Direct measurements of the osmotic pressure are tedious and time-consuming. Equation (16.19) as just derived provides a convenient indirect method. We compare the vapor pressure p_S of the solution with the vapor pressure p_0 of the pure solvent, and compute the osmotic pressure using Eq. (16.19). In fact, one does not usually measure the vapor pressures p_S and p_0 of the solution and the solvent, but rather the corresponding *boiling points* T_S and T_0. Then for the osmotic pressure, we find the simple relation[C16.13]

$$p_\text{os} = \varrho_\text{S} l_v \frac{T_\text{S} - T_0}{T_0} \tag{16.20}$$

(ϱ_S is the density of the solution; for rather dilute solutions, it is ≈ the density of the solvent. T_0 is the boiling point and l_v the specific heat (latent heat) of vaporization of the solvent; cf. Sect. 13.4, especially Eq. (13.7)).

C16.13. The transition from Eq. (16.19) to Eq. (16.20) is intuitively clear. The mathematical derivation of (16.20) proceeds via the CLAUSIUS-CLAPEYRON relation for the vapor pressure curve (cf. Eq. (19.19)) and involves some approximations which are however equally valid as those used to arrive at Eq. (16.19).

All the direct and indirect measurements of osmotic pressure lead in the case of *dilute* solutions (of the order of one-tenth mole/liter) to a surprisingly simple result: For the molecules of the solute, we can apply the equation of state of an ideal gas[4] This formulation is often called "VAN'T HOFF's law":

$$p_{os} V = nRT \quad \text{or} \quad p_{os} = cRT \tag{16.21}$$

(n = amount of substance dissolved in the volume V of solution, that is n/V = concentration c[C16.14]. For $c = 0.1\,\text{kmol/m}^3$ and $T = 273\,\text{K}$ (0 °C), the osmotic pressure is $p_{os} = 2.24 \cdot 10^5\,\text{Pa} \approx 2\,\text{bar}$.)

C16.14. Here, a new quantity is defined, the (molar) *concentration* $c = n/V$, where n is the amount of substance of the dissolved material (solute) and V is the volume of the resulting *solution* (unit: kmol/m^3 or mole/liter). Alternatively, the *mass concentration* can be used: $c\,(\text{mass}) = M/V$, where M is the mass of the solute (unit: kg/m^3 or gram/liter). The conversion factor is the molar mass M_m of the solute: $c\,(\text{mass}) = c\,(\text{molar}) \cdot M_m$. In chemistry and medicine, a practical unit is also used: the 'percent solution'. It is the mass (in gram) of solute dissolved in 100 ml of solution, quoted as a dimensionless number, '% solution'; e.g. 5 g of salt (NaCl) dissolved in 100 ml of aqueous solution is a '5 % saline (salt) solution'.

This occurrence of the same equation of state under quite different conditions is very instructive. We can see that the equation of state is in the end based on the statistical laws governing the thermal motions of a large number of particles, in particular on the fundamental relation

$$E_{kin} = \frac{1}{2} kT\,.$$

This holds even for macroscopic objects and for systems which are not chemically uniform, such as for example dust-like suspended particles in liquids and gases.

16.6 The Experimental Determination of BOLTZMANN's Constant k from the Barometric Equation

Everyone knows what happens in a liquid which is clouded by suspended particles: In the course of time, it *clears* up; the particles "drift down to the bottom" of the container. In this process, the larger particles form a clearly-defined layer, while the finer ones are in a diffuse cloud which gradually becomes thinner. – Interpretation: The particles are pulled downwards by their weight (reduced by their static buoyancy!). But the random thermal motions hinder their sinking downwards[5] The result of this competition is a distribution of the particles over the whole depth of the liquid, just as the molecules of the air are distributed over the depth of the atmosphere.

[4] Therefore, one often uses the osmotic pressure to determine the molar mass M_m of a dissolved substance. One obtains from Eq. (16.21)

$$M_m = R \cdot \frac{M}{V} \cdot \frac{T}{p_{os}}$$

(M/V = density ϱ_S of the solution).

[5] The BROWNian motion (cf. Sect. 9.1). Every suspended particle undergoes a rapid, random series of collisions with the invisible molecules of the liquid. Within the time Δt between two such collisions, each suspended particle has on

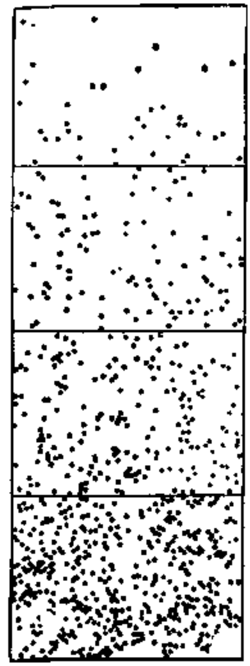


Figure 16.16 The density distribution of suspended particles in water. This drawing is based on photographs taken by J. PERRIN. It represents four horizontal layers at height intervals h of $10\,\mu\text{m}$ each. The particles are grains of gummi gutta with a diameter of $0.6\,\mu\text{m}$ and a density of $\varrho = 1210\,\text{kg/m}^3$. The mass of a single particle is $1.25 \cdot 10^{-16}$ kg; and its "effective" mass after subtracting its buoyancy is $m = 2.17 \cdot 10^{-17}$ kg.

Figure 16.16 shows an example of suspended gummi gutta particles[C16.15] in water, with a diameter of $0.6\,\mu\text{m}$. The instantaneous images show the distribution of the particles in four horizontal layers which are each separated by $10\,\mu\text{m}$ in height. The series of these images is in relatively good agreement with a longitudinal section through the model gas atmosphere, i.e. with Fig. 9.33.

C16.15. Gummi gutta is a natural resin, which is however poisonous. J. PERRIN used a different resin, called mastix.

This qualitative agreement is rather convincing; however, the quantitative evaluation will be decisive.

In Chap. 9, we have already discussed the distribution of the air molecules in the gravitational field of the earth, described by the barometric pressure formula. It was found to be given by:

$$\frac{p_\text{h}}{p_0} = e^{-\frac{\varrho_0 g h}{p_0}} \qquad (9.20)$$

(p_h is the pressure at the altitude h, p_0 the pressure and ϱ_0 the density of the gas at the altitude zero (sea level)).

---

average the energy

$$\tfrac{1}{2}mu^2 = \tfrac{3}{2}kT \qquad (16.2)$$

(m is the mass of the suspended particle).

Unfortunately, the time interval Δt is far too short to permit a measurement of their velocities. Otherwise, one could measure u directly, then insert it into Eq. (16.2) and thus determine k.

Now we replace the ratio of the pressures by the ratio of the number densities of the molecules. We write

$$\frac{p_\mathrm{h}}{p_0} = \frac{N_\mathrm{V,h}}{N_\mathrm{V,0}} = \frac{\text{Number/Volume at the height } h}{\text{Number/Volume at the height zero}} \,. \qquad (16.22)$$

Furthermore, we relate the pressure and the density of the gas via the equation of state for ideal gases. We write this in the form

$$p_0 = \varrho_0 \frac{RT}{M_\mathrm{m}} = \varrho_0 \frac{kT}{m} \qquad (14.21 \text{ and } 14.23)$$

(k = BOLTZMANN constant, M_m = molar mass of the gas, m = mass of one molecule)

and obtain the barometric pressure formula in the form

$$\frac{N_\mathrm{V,h}}{N_\mathrm{V,0}} = e^{-\frac{mgh}{kT}} \,. \qquad (16.23)$$

This equation contains *two* unknowns, namely m and k. If however the mass m is known (it can be found e.g. from the particle diameter and density), then one can determine k by counting the particle numbers N as a function of the height h. This was first carried out by J. PERRIN in 1909)[C16.16]. The "effective" mass of the individual suspended particles was $m = 2.17 \cdot 10^{-17}$ kg (cf. the legend of Fig. 16.16). In air (the mass of a nitrogen molecule is $4.65 \cdot 10^{-26}$ kg), the pressure p and the density ϱ decrease by half for each 5.4 km of altitude increase. In the case of the suspended particles, this same decrease occurs over a height increase of only $\frac{4.65 \cdot 10^{-26}}{2.17 \cdot 10^{-17}} \cdot 5.4 \cdot 10^6\,\mathrm{mm} \approx 0.01\,\mathrm{mm} = 10\,\mu\mathrm{m}$. Compare Fig. 16.16.

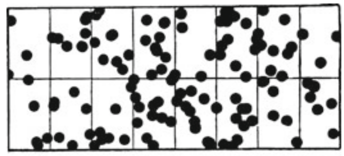
C16.16. See J. Perrin, *"Atoms"*, 2nd ed., Van Nostrand & Co., New York (1923), Chap. 4. – JEAN PERRIN (1870–1942), Nobel Prize 1926.
Another method of determining k is described at the end of Sect. 17.5.

A concentration gradient of the suspended particles can be produced in several other ways. Frequently, the weight is replaced by centrifugal force (Sect. 9.12, Point 1). With today's materials, centrifugal accelerations of up to the 10^6-fold of the earth's acceleration of gravity can be reached (rotational frequency $\approx 1800\,\mathrm{s}^{-1}$ with a circumference velocity of around 900 m/s at a radius of 8 cm). Then instead of g in Eq. (16.23), we can insert $10^6\,g$. In this way, even large *molecules* can be given a density profile like that of suspended macroscopic particles ("ultracentrifuge").

16.7 Statistical Fluctuations and the Particle Number

The upper part of Fig. 16.17 shows a snapshot of our steel-ball model gas (Sect. 9.7, exposure time $\approx 10^{-5}$ s). The whole volume is divided into 16 partial volumes by lines drawn through the container. In the lower part of the figure, the same partial volumes are shown, each one with a number N giving the number of molecules that it contains

Part III



N	3	7	9	10	5	13	9	3
ΔN	−5	−1	+1	+2	−3	+5	+1	−5
$(\Delta N)^2$	25	1	1	4	9	25	1	25
N	5	7	8	11	12	9	5	6
ΔN	−3	−1	0	+3	+4	+1	−3	−2
$(\Delta N)^2$	9	1	0	9	16	1	9	4

Average: $\overline{N} \approx 8$; $\overline{(\Delta N)^2} \approx 8.8$

$$\overline{\varepsilon^2} = \frac{\overline{(\Delta N)^2}}{(\overline{N})^2} = \frac{8.8}{8^2} \approx 14\%; \quad \frac{1}{\mathrm{N}} = \frac{1}{8} \approx 12\%$$

Figure 16.17 The experimental derivation of Eq. (16.25)

at that moment. The average value of N is $\overline{N} \approx 8$, but the individual values exhibit considerable fluctuations around this mean, defined by the equation

$$\varepsilon = \frac{\text{Deviation } \Delta N \text{ of the individual value from the mean}}{\text{Mean value } \overline{N}} . \tag{16.24}$$

The values of ΔN are also shown in Fig. 16.17, along with those of $(\Delta N)^2$. We take the average value of the squared deviations, i.e. $\overline{\varepsilon^2}$, and find after a sufficiently large number of such experiments the result:

$$\overline{\varepsilon^2} = \frac{1}{\overline{N}} . \tag{16.25}$$

In words: *The mean value of the squared deviations is equal to the inverse of the average number of individual objects involved.*

This relation, which we have found here empirically, holds quite generally, e.g. for the volume occupied by the N gas molecules, for the density of a gas, for the fluctuations over time of radioactive decay rates, etc.

We present here a proof for the density fluctuations in an ideal gas. In a large volume of such a gas, we imagine a partial volume V to be confined in a cylinder which is closed at one end by a freely movable piston. This piston can be considered to act as a very large molecule. As such, it participates in the statistical fluctuations due to the thermal motions within the gas.

If the piston during its random back-and-forth movements decreases the volume V by ΔV, then its potential energy increases by

$$E_{\text{pot}} = -\tfrac{1}{2}\Delta p \Delta V .$$

Averaged over time, this must be equal to $\frac{1}{2}kT$, that is

$$\tfrac{1}{2}kT = -\tfrac{1}{2}\Delta p \Delta V . \tag{16.26}$$

Now from differential calculus, we have

$$\Delta p = \frac{\mathrm{d}p}{\mathrm{d}V}\Delta V$$

or, after insertion into Eq. (16.26),

$$(\Delta V)^2 = -\frac{kT}{\mathrm{d}p/\mathrm{d}V} \,. \tag{16.27}$$

The equation of state for the ideal gas (Eq. (14.23)) gives us the denominator:

$$\frac{\mathrm{d}p}{\mathrm{d}V} = -\frac{NkT}{V^2} \,,$$

and thus

$$\left(\frac{\Delta V}{V}\right)^2 = \frac{1}{N} \,. \tag{16.28}$$

The left-hand side of this equation is the relative fluctuation of the volume containing the N molecules. We can now introduce the number density of the molecules, that is $N_\mathrm{V} = N/V$. We have $N_\mathrm{V}(t) \cdot V(t) = N = \text{const}$, so that (from the product rule for differentiation):

$$\Delta N_\mathrm{V} V + N_\mathrm{V} \Delta V = 0 \tag{16.29}$$

or

$$\frac{\Delta N_\mathrm{V}}{N_\mathrm{V}} = -\frac{\Delta V}{V} \,, \tag{16.30}$$

and finally

$$\left(\frac{\Delta N_\mathrm{V}}{N_\mathrm{V}}\right)^2 = \left(\frac{\Delta \varrho}{\varrho}\right)^2 = \frac{1}{N} \tag{16.31}$$

$(\varrho = \text{mass density})$.

16.8 The BOLTZMANN Distribution

We apply the barometric pressure formula:

$$\frac{N_{\mathrm{V,h}}}{N_{\mathrm{V,0}}} = \mathrm{e}^{-\frac{mgh}{kT}} \,. \tag{16.23}$$

The product mgh has a simple physical meaning: it is the difference ΔE of the potential energies of a molecules in the gravitational field at two altitudes separated by the height difference h. We thus obtain

$$\frac{N_{\mathrm{V,h}}}{N_{\mathrm{V,0}}} = \mathrm{e}^{-\frac{\Delta E}{kT}} \tag{16.32}$$

($N_{\mathrm{V,h}}$ = number density of the molecules at the height h, $N_{\mathrm{V,0}}$ = number density of the molecules at the base height. Molecules with the index h have more energy than those at the base height by an amount ΔE.)

This BOLTZMANN *distribution*, derived here for a special case, holds quite generally. *For all processes which take place in thermal equilibrium, it gives the ratio of the numbers of molecules whose energies differ by the energy ΔE in some arbitrary force field.*

Within the scope of this introductory text, a few examples of the possible applications of this very general equation (16.32) will have to suffice. It can be used for example to describe:

The dependence of the vapor pressure of a material on its temperature. Then ΔE refers to the heat of vaporization per molecule.

The MAXWELL-BOLTZMANN velocity distribution (Sect. 16.3). Then ΔE refers to the kinetic energy of the molecules.

The dependence of the equilibrium of a chemical reaction on the concentrations of the reactants (the law of mass action). Here, ΔE refers to the *reaction energy* per molecule for the reaction.

The dependence of the electrical conductivity of a non-metallic electron conductor on the temperature. Then ΔE is the separation energy of an electron from its binding site.

The electron emission of a glowing body. Then ΔE refers to the work function of an electron in the material.

The spectral energy distribution of the radiation from a 'black body'. Then ΔE is the energy $h\nu$ of a light quantum of frequency ν.

Owing to the great general importance of Eq. (16.32), we present below another straightforward derivation:
We assume that two molecules with the energies E_1 and E_2 collide elastically during their thermal motions; after their collision, they have the energies E_1' and E_2'. Then we have from energy conservation

$$E_1 + E_2 = E_1' + E_2' \,. \tag{16.33}$$

In statistical equilibrium (quasi-stationary state), the number of transitions $\overrightarrow{N}$ from left to right in this equation must be equal to the number of transitions $\overleftarrow{N}$ from right to left. We denote the number of molecules with the energy E by $N(E)$. Then we have

$$\begin{aligned} \overrightarrow{N} &= \text{const}\, N(E_1)N(E_2)\,, \\ \overleftarrow{N} &= \text{const}\, N(E_1')N(E_2')\,. \end{aligned} \tag{16.34}$$

We take the two constants to be equal; that is a plausible assumption, which is justified later by the success of our results. With this condition, if follows from Eq. (16.34) that

$$N(E_1)N(E_2) = N(E_1')N(E_2')\,. \tag{16.35}$$

Now we have to search for a function $N(E)$ which fulfills Eqns. (16.33) and (16.35) *simultaneously*. This is the case for the trial function

$$N(E) = N_0 e^{\beta E}\,. \tag{16.36}$$

It converts Eq. (16.35) into

$$N_0^2 e^{\beta(E_1+E_2)} = N_0^2 e^{\beta(E_1'+E_2')} ,$$

and this, when Eq. (16.33) holds, is an identity. Furthermore, it follows from Eq. (16.36) that

$$\frac{N(E_1)}{N(E_2)} = e^{\beta(E_1-E_2)} . \tag{16.37}$$

Finally, the comparison with Eq. (16.23), our special case for the barometric pressure formula, gives $\beta = -1/kT$. We thus obtain quite generally

$$\frac{N(E_2)}{N(E_1)} = e^{-\frac{E_2-E_1}{kT}} . \tag{16.38}$$

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_16) contains supplementary material, which is available to authorized users.

Transport Processes: Diffusion and Heat Conduction

17

17.1 Preliminary Remarks

We have already dealt twice with diffusion processes, both times in connection with the molecular picture for thermal motions (Sect. 16.1). In the present chapter, we want to discuss some aspects of the quantitative treatment of diffusion, and then to make a few remarks about the related topics of heat conduction and heat transport. – Beginners may want to skim over some of this material. These are indeed problems of practical importance, but their quantitative treatment is still not very satisfying.

17.2 Diffusion and Mixing

To begin, we need to demarcate the concept of *diffusion* clearly from other phenomena involving mixing. – First, let us imagine that two different but miscible liquids are arranged in two layers (cf. Fig. 9.1), with the liquid of higher density below. The initially sharp boundary surface between the two layers will gradually become washed out, and in the course of many weeks, the two liquids will become completely mixed into a homogeneous solution. This case represents true diffusion; the mutual intermixing of the two types of molecules is simply a result of their molecular *thermal motions*.

In the second case, we imagine that local density variations are present within the two liquids, produced for example by local temperature differences. Then there will be upwards and downwards-directed *currents* within the liquids, initially forming clearly-recognizable patterns. This type of *free convection* makes a considerable contribution to the intermixing of the two liquids; in comparison, genuine diffusion may become practically insignificant (being orders of magnitude *slower*).

This latter effect is exemplified in a third case. Now, we imagine that convection is *forced*: Using moving solid bodies (e.g. the blades of a *stirrer*), we produce *turbulent flow*.

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_17

In order to observe pure, genuine diffusion by itself, we must suppress convection in its two forms, free and forced, by making use of a suitable experimental arrangement. We can for example allow fluids with a lower density to "float" on other fluids with a higher density, carefully avoiding the occurrence of local temperature variations. Most simply, we can introduce one of the molecular species in the solid phase.

17.3 FICK's First Law and the Diffusion Constant

We refer to Fig. 16.2 and reproduce it schematically in Fig. 17.1: A gas, e.g. H_2, is allowed to diffuse through a porous partition of thickness l. On both sides of the partition and within its channels, air serves as "solvent". *The partition serves only to avoid disturbing convection currents.*

As usual, we define the number density of the molecules by the quotient

$$N_V = \frac{\text{Number } N \text{ of dissolved molecules}}{\text{Volume } V \text{ of the solution}} \,. \qquad (13.1)$$

On the front (left) side of the partition, we maintain the number density constant at $N_{V,a}$; behind the partition, all the molecules that diffuse through are removed by some arbitrary mechanism, e.g. they are blown away by an air stream. Then within the partition, there is a constant gradient in the number density:

$$\frac{\partial N_V}{\partial x} = -\frac{N_{V,a}}{l} \,.$$

We measure the number ∂N of molecules that diffuse through the area A within the time ∂t, and thus experimentally determine the "molec-

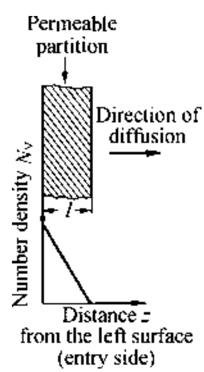


Figure 17.1 The derivation of Eq. (17.1)

ular current":

$$\frac{\partial N}{\partial t} = -DA\frac{\partial N_V}{\partial x}\,. \tag{17.1}$$

In words: The current of diffusing molecules is proportional to the gradient of their number density (FICK's first law). The constant of proportionality D is called the *diffusion constant* (here for H_2 diffusing in the air of the channels). This law holds in general, not just for the particular geometry described above.

So much for the empirical facts. The molecular picture leads us to an interpretation and permits the calculation of the diffusion constant D in simple cases.

The diffusing molecules are continually undergoing collisions with the molecules from their surroundings (the "solvent"). Each of them is subject to a time-averaged force F in the direction of the diffusion, which moves it against the frictional resistance of its surroundings at an average velocity u. The corresponding frictional work is performed with the power

$$\dot{W} = uF \tag{5.33}$$

and is returned to the surroundings as kinetic energy. – For later considerations, we define the quotient

$$\mu = \frac{u}{F} \tag{17.2}$$

as the mechanical *mobility*.

If the mean free path is short compared to the diameter, then it follows for example for spherical molecules from STOKES's formula for frictional motion (Eq. (10.7)) that:

$$\mu = (6\pi R\eta)^{-1}$$

(R is the radius of the molecules, and η is the coefficient of viscosity of the surroundings, i.e. of the solvent).

Figure 17.2 represents a thin layer of the solvent perpendicular to the direction of the diffusive motion. The cross-sectional area of the layer is A, and its thickness is Δx. It contains $N = N_V A\Delta x$ dissolved molecules (black dots). Each one of them is subject to the force F. This force can be replaced by an osmotic pressure $\Delta p = (p_1 - p_2)$ which presses against the surface element A/N. Then we find the relation:

$$F = \Delta p\frac{A}{N} = -\frac{1}{N_V}\frac{\Delta p}{\Delta x}\,. \tag{17.3}$$

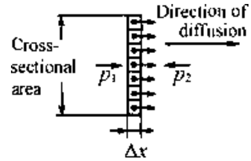


Figure 17.2 The mechanism of FICK's law

For the osmotic pressure, the ideal-gas equation holds:

$$p = \frac{N}{V}kT = N_V kT\,. \tag{14.23}$$

It leads to

$$\Delta p = \Delta N_V kT\,. \tag{17.4}$$

Inserting equations (17.3) and (17.4) into Eq. (17.2) yields

$$F = \frac{u}{\mu} = -\frac{kT}{N_V}\frac{\Delta N_V}{\Delta x}$$

or, with the abbreviation

$$\text{Diffusion constant } D = \mu kT\,, \tag{17.5}$$

we obtain

$$u = -\frac{D}{N_V}\frac{\Delta N_V}{\Delta x}\,. \tag{17.6}$$

This "diffusion velocity" u implies that within the time Δt, ΔN molecules diffuse through an area A of the partition. Then we have

$$\Delta N = \Delta t A u N_V\,, \tag{17.7}$$

or, for the diffusion velocity,

$$u = \frac{1}{A}\frac{\Delta N}{\Delta t}\frac{1}{N_V}\,. \tag{17.8}$$

Finally, we combine Eqns. (17.6) and (17.8) to yield FICK's first law, which we had already obtained above from purely empirical considerations:

$$\frac{\Delta N}{\Delta t} = -DA\frac{\Delta N_V}{\Delta x}\,. \tag{17.1}$$

The diffusion constant D has the dimensions m^2/s. Table 17.1 lists some measured values.

Often, the diffusing particles are electrically-charged molecules. These "charge carriers" acquire a preferred direction when an electric field is applied during their diffusion, and they thus carry an

Table 17.1 Some examples of diffusion constants and diffusion paths

(Molecule)	diffuses	at a temperature of (°C)	with the diffusion constant D $\left(\frac{m^2}{s}\right)$	and a single molecule moves from its original position in one day (according to Eq. (17.16)) by a distance of
H_2 O_2	in air at atmospheric pressure	0 0	$6.4 \cdot 10^{-5}$ $1.8 \cdot 10^{-5}$	3.3 m 1.8 m
Urea Salt Cane sugar	in water	15 10 18.5	$1.0 \cdot 10^{-9}$ $9.3 \cdot 10^{-10}$ $3.7 \cdot 10^{-10}$	13 mm 13 mm 8 mm
Gold	in molten lead	490	$3.5 \cdot 10^{-9}$	25 mm
Gold	in solid lead	165	$4.6 \cdot 10^{-12}$	0.9 mm
H_2 Potassium as color centers[a]	in a KBr crystal	680 650	$2.3 \cdot 10^{-8}$ $5.2 \cdot 10^{-8}$	6 cm 9.5 cm

[a] Color centers: See Vol. 2, Sect. 27.14.

electric current. In this way, for example ionic currents and electron currents can arise in liquids, in gases and in solids. In favorable cases, these directed diffusion processes can be followed directly by eye (see R.W. Pohl, *Elektrizitätslehre*, 21st ed. (1975), Sects. 16, §6 and 25, §22; or cf. the image on the Wiki site for R.W. POHL, https://en.wikipedia.org/wiki/Robert_Pohl).

The mobility μ_e of the charge carriers is not referred to the force, but rather to the electric field E = force F/charge e. We thus obtain the electrical mobility

$$\mu_e = \frac{u}{E} = \frac{ue}{F} = e\mu \tag{17.9}$$

(e = charge of the carriers, e.g. in ampere second (A s); μ as defined in Eq. (17.2)).

17.4 Quasi-Stationary Diffusion

The application of FICK's first law presupposes that we know the gradient of the number density, i.e. $\Delta N_V/\Delta x$. It can be readily determined for a stationary state (Fig. 17.1). This may also be the case to a good approximation for many processes which are only approximately stationary (quasi-stationary states). An example of this type is sketched in Fig. 17.3. A solid body Y contains N molecules of type I in the volume V; their number density is thus $N_V = N/V$. Think of a solid-state solution, e.g. of thallium atoms in a KBr crystal. Into this solid body, we suppose that N^* molecules of a gas diffuse inwards from the left, for example Br_2 molecules. They are supposed to unite at the diffusion front with N molecules of type I and therefore to drop out of further diffusion processes. *By what distance x does the diffusion front move forward within a time t?*

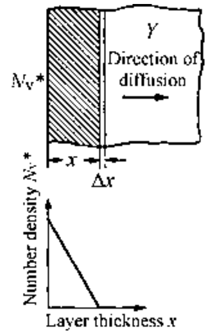


Figure 17.3 Linear concentration gradient for diffusion accompanied by chemical reaction

Within the volume $A\,\mathrm{d}x$, there are $\mathrm{d}N = N_V A\,\mathrm{d}x$ molecules of type I; thus, we have

$$\frac{\mathrm{d}N^*}{\mathrm{d}t} = N_V A \frac{\mathrm{d}x}{\mathrm{d}t}\,. \tag{17.10}$$

This process can still be treated to a very good approximation as stationary. That is, we can neglect $\mathrm{d}x$ relative to x and treat the process as if it occurs practically at a fixed location. As a result, the layer of thickness x which has already been chemically converted takes over the function of the partition in Fig. 17.1. The concentration of diffusing molecules is N_V^* at the left of this layer, and at the right, after the layer, i.e. at the diffusion and reaction front, it is zero. Thus we again find as an approximate expression for the diffusion concentration gradient

$$\frac{\Delta N_V^*}{\Delta x} = -\frac{N_V^*}{x}\,. \tag{17.11}$$

We apply Eq. (17.1) to the N^* molecules and obtain together with Eqns. (17.10) and (17.11) the result

$$N_V \frac{\mathrm{d}x}{\mathrm{d}t} = D\frac{N_V^*}{x}\,. \tag{17.12}$$

The solution of this differential equation is

$$x^2 = 2\frac{N_V^*}{N_V}Dt\,, \tag{17.13}$$

and thus, as the answer to the question in italics above,

$$\frac{x^2}{t} = D \cdot \mathrm{const} \tag{17.14}$$

('const' is a pure number).

This result has been demonstrated with the example mentioned above, the diffusion of Br_2 into a Tl-doped KBr crystal[C17.1]. The previously brown layer x becomes clear, since the TlBr molecules formed are colorless. – Equation. (17.14) plays an important role for surface reactions on metals, i.e. when they "tarnish".

C17.1. The crystal takes on a brown color on being heated in K vapor (E. Mollwo, Ann. Physik, 5th Series, Vol. 29, p. 394 (1937)).

17.5 Non-Stationary Diffusion

In our two examples of applications of FICK's first law, the number density of the diffusing molecules at the front of the diffusion path (at a distance l from the left-hand surface in Fig. 17.1 or the coordinate $x + \Delta x$ in Fig. 17.3, respectively) was held constant at the value zero. In general, the number density N_V of the molecules on both sides of the diffusion region being considered is variable over time. The process is then no longer stationary; the spatial distribution of the diffusing molecules *varies* in the course of time. The increase in the number density N_V in a region between x_1 and x_2 can be obtained from the difference in the numbers of molecules flowing in at x_1 and out at x_2.

If we again assign a positive sign to a particle current moving in the positive x-direction, then for the rate of change of the number density N_V in the volume V between x_1 and x_2 we find

$$\frac{\partial N_V}{\partial t} = \frac{1}{V}\left\{ \left.\frac{\partial N}{\partial t}\right|_{x_1} - \left.\frac{\partial N}{\partial t}\right|_{x_2} \right\}.$$

Setting $V = A \cdot (x_2 - x_1)$ and making use of Eq. (17.1), we obtain from this

$$\frac{\partial N_V}{\partial t} = D\frac{\partial^2 N_V}{\partial x^2}. \tag{17.15}$$

This differential equation is called FICK's second law.

We also give an example of a non-stationary diffusion process here; however, we will not derive it. In our example, at the time $t = 0$, the number density in a whole region is taken to be zero. At the front side of this region, it has the value $N_{V,a}$, and this value is held *constant* during the entire process of diffusion. How does the distance x between the location of a certain concentration $N_{V,x}$ and the entry point $x = 0$ change? Answer: Once again, we have

$$\frac{x^2}{t} = D \cdot \text{const}. \tag{17.14}$$

The conclusions implied by this equation are represented graphically in Fig. 17.4: The distributions of the number densities at various times remain similar to each other. They can be made identical by a suitable choice of the time-axis scales.

Part III

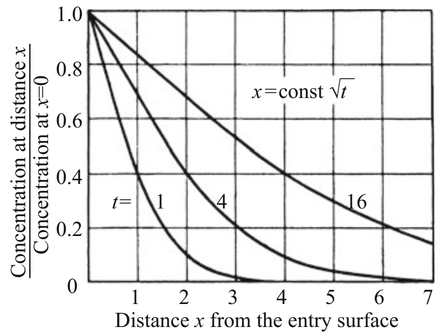


Figure 17.4 The diffusion gradient as a function of time. The distances x traversed by a particular given value of the number density, e.g. 40 % of its initial value $N_{V,a}$, behave as the square roots of the diffusion times, e.g. as $1:2:4$.

In the case of BROWNian motion, one does not observe diffusion as a mass phenomenon, but instead as an individual process. One cannot follow the progress of a certain concentration, but only of single particles. One measures how the distance x of a given particle from an arbitrary starting point gradually increases with time t. In this case, for the constant in Eq. (17.14), we find the number 2. We thus have

$$\frac{x^2}{t} = 2D\,. \tag{17.16}$$

If the particles are spherical, we can compute D by using Eq. (10.7) (Stokes' law) and Eqns. (17.2) and (17.5). We then obtain

$$\frac{x^2}{t} = \frac{kT}{3\pi\eta R} \tag{17.17}$$

(R is the particles' radius, η the viscosity coefficient of the fluid).

From this equation, we can also make an experimental determination of the value of the BOLTZMANN constant ($k = 1.38 \cdot 10^{-23}$ W s/K).

17.6 General Considerations on Heat Conduction and Heat Transport

As we have already mentioned in Sect. 13.3, heat can be transported either through "conduction" or through "radiation". The necessary precondition is a temperature gradient. Heat conduction takes place in the interior of materials via different mechanisms, such as molecular motions in gases and liquids, or elastic waves and electronic motion in solids. In gases and liquids, as for diffusion, "genuine" heat conduction due to processes on the molecular scale must be distinguished from the usually predominant heat transport by free and forced *convection*.

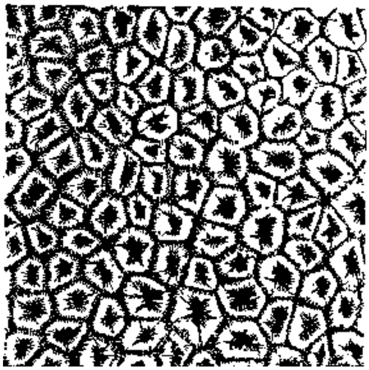


Figure 17.5 Convective heat transport via convection cells, each of which is formed by a flow of hot liquid from the hotplate upwards and of cooled liquid back downwards (BÉNARD cells (1900)[C17.2]) within a liquid layer (e.g. in liquid paraffine). The cells are deformed by mutual pressure into mainly hexagonal shapes, sometimes with a regularity which approaches that of a honeycomb. In each cell, the liquid is rising in the interior and sinking on the outside of the cell. The flow is completely stationary. If it is disturbed by stirring, a new pattern is formed within a fraction of a minute (actual size).

C17.2. E.L. KOSCHMIEDER: "BÉNARD Cells and Taylor Vortices", Cambridge University Press (1993).

Figure 17.5 shows an example for heat transport by convection. A hot metal plate is covered with a layer of liquid about 3 mm thick; above it is the cool room air. The liquid contains suspended fine aluminum particles which make its free convection visible. They exhibit a complicated, honeycomb-like pattern. Heat transport by forced convection can be found for example in the cooling systems of automobiles.

An important and instructive application of heat transport mainly by convection is exhibited by the *countercurrent heat exchanger*. – In the laboratory, one sometimes needs to change the temperature of a flowing substance *temporarily*, e.g. in order to accelerate a chemical reaction within a fluid or to purify a liquid by distillation. Then the scheme which is sketched in Fig. 17.6 can be used: At the left, heat is applied to the flowing substance by a Bunsen burner flame below the flask, while at the right, heat is removed by a water-cooled condenser. Such a setup is convenient but wasteful. All of the heat power added at point *a* is lost again at *b* and carried off by the cooling water. Such an arrangement would be unacceptable for large-scale technical applications; but it can be avoided, in the ideal limiting case completely. This is the function of the countercurrent heat exchanger, developed in 1857 by WILHELM SIEMENS[C17.3]. Its principle is indicated in Fig. 17.7. At the left, the liquid whose temperature is to be varied flows in, for example water at room temperature T_1; below, in the round boiling flask, it has a temperature T_2, say 353 K (80 °C). At the upper right, water at room temperature T_1 again flows out. In principle, one needs to add heat to the flowing liquid only at the beginning of operation; in this example, by heating the water in the boiling flask to 80 °C. From then on, further heat input is theoretically

C17.3. W. SIEMENS: see Comment C15.6 in Chap. 15. – An important technical application of the countercurrent exchanger is in the liquefaction of air by the LINDE process (see Sect. 15.5).

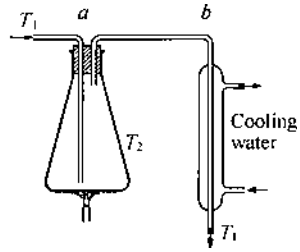


Figure 17.6 Producing temperature variations in a flowing substance *without* energy conservation

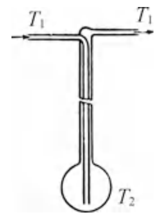


Figure 17.7 Schematic of a countercurrent heat exchanger; the temperature change of the flowing substance is accomplished with minimal energy losses

unnecessary. The upwards-flowing water in the outer tube gives its heat up to the downwards-flowing water in the inner tube. When the tubes are sufficiently long, this "temperature exchange" takes place with only tiny temperature gradients. At every level along the tube, the upwards-flowing (outgoing) water is only slightly warmer than the downwards-flowing (incoming) water adjacent to it.

In reality, no countercurrent exchanger can function without adding any heat energy at all. Firstly, losses due to heat conduction along the tubing are unavoidable[1]. As a result, continuous addition of heat is necessary, but at a much lower power level than in the setup of Fig. 17.6. Secondly, the flow within the tubes must be turbulent[C17.4], in order to ensure good heat exchange. Maintaining a turbulent flow however requires a power input to a pump or some similar arrangement.

C17.4. Turbulent flows were already described in Sect. 10.4.

Summarizing briefly: In the ideal limiting case, the countercurrent exchanger fulfils an important technical function: It makes it possible to temporarily change the temperature of a flowing substance without a continual high power input.

For this reason, this trick is used by many warm-blooded animal species to reduce heat losses, for example when their feet remain in contact with cold water or ice for long periods. – Example: The male Emperor penguin who is brooding an egg. He has to stand on the ice in the south-polar weather for around two months without any nourishment!

[1] They can, however, be minimized by using very long, helically coiled tubing.

17.7 Stationary Heat Conduction

Genuine heat conduction can be readily observed in *solid* bodies. We suppose that in a time Δt, a (thermal) energy ΔQ passes through an area A, driven by a temperature gradient $\Delta T/\Delta x$. Then we obtain for the *heat current*:

$$\frac{\Delta Q}{\Delta t} = -\lambda A \frac{\Delta T}{\Delta x} \,. \tag{17.18}$$

In words: The heat current is proportional to the temperature gradient. The proportionality constant λ is called the *thermal conductivity* coefficient. It depends strongly on the temperature. This is illustrated in Fig. 17.8 with three examples. – In copper, a typical metal, heat transport is accomplished almost entirely by the (conduction) electrons; in crystalline quartz, an insulator, it is due completely to high-frequency elastic waves (quantized sound waves or phonons). Both the electrons and the elastic waves are scattered by other elastic waves (electron-phonon and phonon-phonon scattering). The frequency of occurrence of such scattering processes decreases with

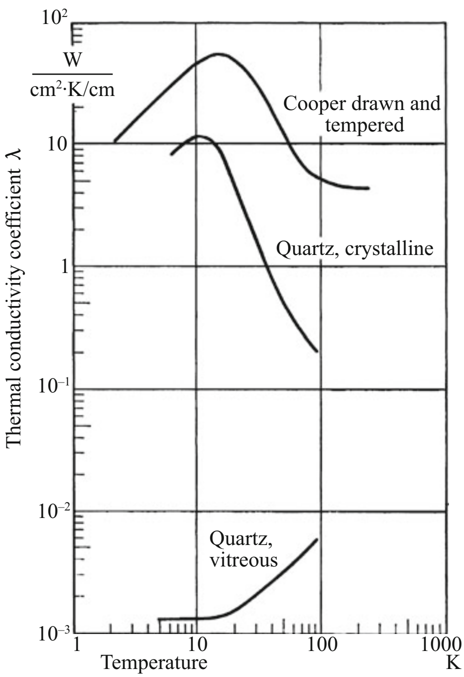


Figure 17.8 The dependence of the thermal conductivity coefficient λ on the temperature

decreasing temperature. At very low temperatures, a different disturbance of heat transport predominates, namely scattering by *localized* lattice defects and by the surfaces of crystallites. Given the complete disorder in quartz glass, for example, its heat conductivity is very small[C17.5]. All these topics belong among the problems of solid-state physics.

C17.5. See e.g. R.O. Pohl, Xiao Liu, and E. Thompson, *Rev. Mod. Phys.* **74**, 991 (2002).

17.8 Non-Stationary Heat Conduction

Non-stationary heat conduction can be described by a differential equation, analogous to FICK's second law for diffusion. For heat conduction which is limited to the x direction (i.e. one-dimensional), it is given by

$$\frac{\partial T}{\partial t} = -\frac{\lambda}{\varrho \cdot c} \cdot \frac{\partial^2 T}{\partial x^2} \,. \tag{17.19}$$

Here, λ is the thermal conductivity coefficient defined by Eq. (17.18), ϱ is the density of the conducting material, and c is its specific heat. The ratio $\frac{\lambda}{\varrho \cdot c}$ is called the *temperature conductivity* (units e.g. m^2/s).

We will give only one example of non-stationary heat conduction. It is analogous to the diffusion process illustrated in Fig. 17.4. In Fig. 17.9, at the time $t = 0$, we suppose that the temperature T is the same at all points within a metal rod. Then it is suddenly increased to the value T_1 at one end of the rod; thereafter, we observe the temperature distribution along the rod as a function of time. Figure 17.9

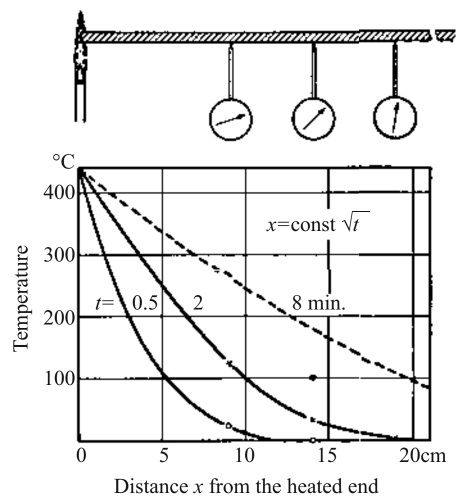


Figure 17.9 Raw data from a demonstration experiment showing the time evolution of a temperature gradient (in an iron rod of 8 mm diameter and 1 m length, without any thermal insulation. The setup is shown in the *upper image*). The distance travelled by a point at a particular temperature is proportional to the square root of the time.

Part III

shows the results: The distance x from the entry point $x = 0$ to the point at which a certain temperature T_x is observed increases proportionally to the square root of the elapsed time. The temperature distributions observed at different times remain similar to each other. They can be made to appear identical by a suitable choice of the time scales.

17.9 Transport Processes in Gases and Their Lack of Pressure Dependence

In gases and liquids (i.e. fluids), the relation between diffusion and heat conduction is intuitively clear. In the case of diffusion, we are dealing with the statistically-ordered forward motion of the molecules. *Heat conduction can be described concisely as the diffusion of an excess of kinetic energy of the molecules*. In Fig. 17.10, we suppose that the wall at the left is at a higher temperature than the gas in the container; convection is excluded. Then the adjacent gas layer will be the first to be heated, i.e. its molecules will acquire an increased kinetic energy. This distinguishes them from the molecules in other layers. No distinguishing feature of any kind can be maintained in a statistical process with a large number of individuals (molecules). The distinguished molecules must therefore give up a part of their excess kinetic energy in collisions with the other molecules. Thus, the excess kinetic energy gradually diffuses into the gas layers further to the right.

In a completely analogous manner, we can understand another phenomenon which depends on molecular motions but which we have thus far not explained: *Internal friction* (Sect. 10.2, see in particular Fig. 10.2). In Fig. 17.11, we suppose that the left wall is moving upwards with a velocity u. The molecules in the adjacent layer acquire

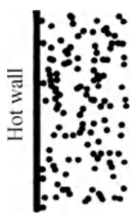


Figure 17.10 The mechanism of heat conduction in gases

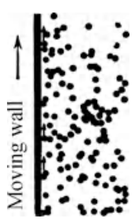


Figure 17.11 The mechanism of internal friction in gases

Part III

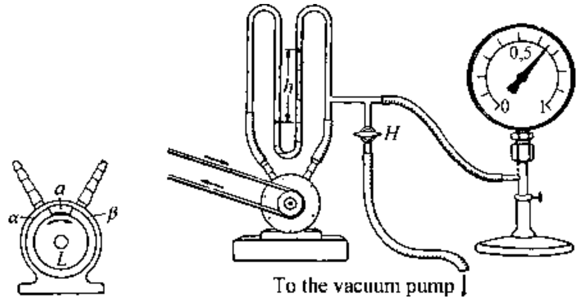


Figure 17.12 The lack of dependence of the internal friction (viscosity) on pressure. The spacing between the rotating cylinder L and its surrounding housing is exaggerated in the drawing for clarity.

a preferred direction of motion (upwards) through collisions with this wall, and therefore have an additional momentum component *mu* directed upwards. This is indicated in the figure by small arrows. This one-sided excess momentum distinguishes the molecules of the layer nearest the wall from all the others in the container. This distinguishing feature cannot be maintained in a statistical ensemble of molecules; thus, the upwards-directed excess momentum is gradually transported into the layers further to the right and causes them to move, albeit more slowly, in the same direction as the wall. We see that *internal friction* or viscosity can be concisely described as the *diffusion of an additional momentum component* of the molecules.

Just like diffusion and heat conduction, internal friction (viscosity) can be strongly enhanced by convection, especially by turbulent convection. We have already seen this in Sect. 10.4.

The relation between diffusion, heat conduction and internal friction in gases becomes very clear through a common feature: All these phenomena are independent of the gas pressure over a wide range. This surprising fact can be most simply illustrated for the case of internal friction.

In Fig. 17.12, an inner cylinder is rotating within an outer cylinder. Their spacing is about 1 mm, apart from a segment a. There, the spacing is reduced to roughly 0.2 mm. During rotation, the air between the two cylinders is set in motion in the direction of the rotation, due to its viscosity. Then, between the two regions α and β, a pressure difference is built up, here $\approx$ 20 hPa. We then pump a large fraction of the air, 80 % or more, out of the chamber containing the cylinders. Nevertheless, the manometer continues to show the same pressure difference of $\approx$ 20 hPa.

The following experiment is even more striking: We put a steel ball into the upper end of a precision glass tube which is standing vertically (diameter $\approx$ 15 mm). The difference in the outer diameter of the ball and the inner diameter of the tube is about 0.01 mm. The tube contains air at atmospheric pressure ($\approx 10^5$ Pa). During the downward motion of the ball as it falls

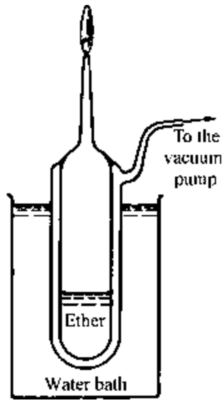


Figure 17.13 A simple demonstration experiment to show the lack of dependence of the heat conductivity of a gas on its pressure. The heat current flows out of the hot water bath through the gas layer in the double-walled vessel into the diethyl ether ($(CH_3CH_2)_2O$) and produces a corresponding current of ether vapor. The strength of this current is shown by the height of an ether flame which is burning at the top of the vessel. Over a large range, it is independent of the pressure of the gas layer between the vessel walls.

through the tube, the gas within the tube must flow around the ball through a very narrow, circular slit. Its viscosity produces a large resistance: The ball does not "fall" with an *accelerated* motion, but instead it "sinks" (after a brief initial period) with a *constant* velocity (Sect. 5.11). At this velocity, it covers a distance s (e.g. 60 cm) in a time t (e.g. 30 s). If we reduce the gas pressure p, the time required for the ball to sink initially remains the same. Only at a pressure of $p \approx 0.16 \cdot 10^5$ Pa does it begin to sink noticeably faster; at $p \approx 1.3$ Pa, it finally approaches free fall.

The lack of dependence of the heat conductivity of a gas on its pressure is likewise relatively easy to demonstrate. Details are shown in Fig. 17.13.

So much for the facts. Now we consider their explanation on a molecular basis: The number of molecules diffusing in a given time through a given area, the amount of their additional momentum or of their excess kinetic energy are all proportional to their number density N_V. They are in addition proportional to the *mean free path* l of the molecules, i.e. to the distance they travel on the average between collisions (Sect. 16.3). N_V increases in direct proportion to the gas pressure, but l decreases in indirect proportion to the pressure; their product is thus independent of pressure. Therefore, in gases, every type of diffusion (or "transport process") does not depend on the pressure.

Hydrogen has a very long mean free path; under standard conditions (0 °C, 1013 hPa), it is $l = 1.4 \cdot 10^{-7}$ m. As a result, hydrogen has a very high heat conductivity (cf. Fig. 17.14).

At very low pressures, the concept of the mean free path l begins to become meaningless: The mean free paths of the molecules become larger than the dimensions of their container. The molecules can then bounce back and forth unimpeded between opposite container walls. The momentum or energy transferred to and from the walls becomes smaller as the density of the gas decreases. This is the principle of the Thermos bottle, a container which has double walls separated by

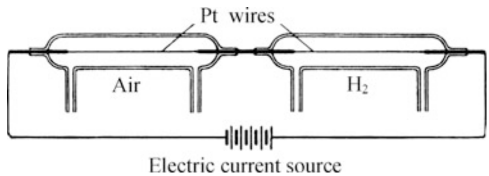


Figure 17.14 A simple demonstration experiment for comparing the heat conductivities of H_2 and air. In addition to genuine heat conduction, free convection also occurs. Two identical platinum wires (connected in series) are heated by the same electric current. The wire in air glows bright yellow, and is thus very hot, while the wire in H_2 remains dark; it is cooled by the high thermal conductivity of the hydrogen. In mixtures of gases, the heat conductivity depends on the composition of the mixture. This is the reason why heat conductivity is often used in technical applications to monitor the composition of a gas mixture. – The basic aspects of the different processes can be readily demonstrated with the above setup.

Part III

an evacuated space. Heat transport to and from its interior can occur only through radiation.

But even these radiation losses can be reduced! To accomplish this, the vacuum space is filled with a large number of thin layers of aluminized plastic, e.g. Mylar, typically 30 layers/cm, which are separated from each other by an electrically-insulating material, for example Nylon netting. Since the temperature difference between neighboring Mylar layers is small, the radiative heat transport between them is also very small. The overall heat transfer through this stack is inversely proportional to its thickness, and thus can be represented as a mean apparent thermal conductivity between the outer, warm surface of the container and its inner, cold surface. Even with a residual gas pressure of some 10^{-2} Pa, the thermal conductivity of these "superinsulated" Thermos or cryogenic containers is of the order of 10^{-7} W/(cm K) between the outer wall at 300 K and the inner wall at 20 K[C17.6], which is about four orders of magnitude lower than the thermal conductivity of vitreous quartz at 10 K (Fig. 17.8).

C17.6. Jack W. Ekin, *Experimental Techniques for Low-Temperature Measurements*, Oxford University Press (2006), p. 57 and Fig. 2.4, p. 58.

17.10 Determination of the Mean Free Path

The relationship between the three diffusion phenomena (of molecules, of momentum and of energy) and the mean free path l makes it possible to determine this important quantity experimentally by three different methods. The necessary formulas (Eqns. (17.22), (17.24) and (17.26)) can be obtained from fairly simple considerations. We proceed in a similar manner as in Sect. 9.8, where we dealt with the pressure of a gas. Instead of Fig. 9.25 as used there, we refer here to Fig. 17.15. We consider the molecules which pass through a cross-sectional area A at the position x, coming from the left and the right. Two other areas are drawn in to the left and the right of the area A at the positions $(x - l)$ and $(x + l)$; l here denotes the mean free path.

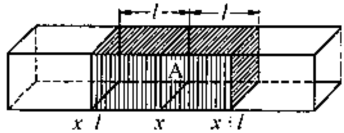


Figure 17.15 The derivation of Eq. (17.20)

On average, the molecules which will pass through A from the left and from the right suffer their last collisions before reaching A at these two adjacent areas. The number densities N_V of the molecules and their velocities u thus remain constant within the two shaded volumes. From the left, in a time dt, a number

$$\mathrm{d}N_1 = A\mathrm{d}t\tfrac{1}{6}(N_V u)_{(x-l)}\,,$$

of molecules arrives at A, and from the right, a number

$$\mathrm{d}N_2 = A\mathrm{d}t\tfrac{1}{6}(N_V u)_{(x+l)}\,.$$

We have encountered the factor $\frac{1}{6}$ previously in Sect. 9.8. The resulting molecular current in the x-direction is then

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{A}{6}[(N_V u)_{(x-l)} - (N_V u)_{(x+l)}] = -\frac{A}{6}\frac{\mathrm{d}(N_V u)}{\mathrm{d}x}2l\,,$$

or

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -A\frac{l}{3}\frac{\mathrm{d}(N_V u)}{\mathrm{d}x}\,. \qquad (17.20)$$

Note that this equation is quite general and could be applied to any ensemble of moving particles, macroscopic or microscopic. We will apply this general equation to particular situations involving molecular motions in the following:

1. *Diffusion of molecules*. The temperature is the same everywhere, and thus $\overline{u}$, the average molecular velocity, is constant[C17.7]. For the diffusing molecular currents, we obtain

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -A\frac{l\overline{u}}{3}\frac{\mathrm{d}N_V}{\mathrm{d}x} = -DA\frac{\mathrm{d}N_V}{\mathrm{d}x}\,, \qquad (17.21)$$

which is FICK's first law with the diffusion constant

$$D = \frac{l\overline{u}}{3}\,. \qquad (17.22)$$

C17.7. We now apply the general equation (17.20) to the particular case of molecules with their random thermal motion. We cannot specify a microscopic velocity here; only a statistical average over the six spatial directions and the wide range of magnitudes can be given. It was introduced in Sect. 9.8 as the "rms (root-mean-square) velocity" u_{rms}, and more precisely defined in Eq. (16.3). The *average molecular velocity* referred to here and used in the next sections is this statistical velocity. It is an average over the microscopic, random velocities of the molecules, and we denote it here for short by $\overline{u}$, to keep it separate from macroscopic velocities such as $u_\perp$, referred to in the discussion of viscosity, or the phonon velocity u mentioned in Comment C17.8.

2. *Diffusion of additional momentum, i.e. internal friction or viscosity* as in Fig. 17.11. Perpendicular to the direction of their diffusion, the molecules have an additional velocity component $u_\perp$ (indicated by small arrows in Fig. 17.11. This is typically a *macroscopic flow velocity*). They thus have an excess momentum $p_\perp$ in the direction of this velocity. For the *momentum current*, we have

$$\frac{\mathrm{d}p_\perp}{\mathrm{d}t} = -\frac{l}{3}A\frac{\mathrm{d}(N_\mathrm{V}\overline{u}\,mu_\perp)}{\mathrm{d}x} = -\frac{l\overline{u}}{3}AN_\mathrm{V}\frac{\mathrm{d}u_\perp}{\mathrm{d}x}m$$

or, using Eq. (5.41),

$$\frac{F}{A} = -\eta\frac{\mathrm{d}u_\perp}{\mathrm{d}x}\,, \tag{17.23}$$

and, assuming a homogeneous velocity gradient and generalizing the velocity to u,

$$F = \eta A\frac{u}{x}\,, \tag{10.2}$$

we find the viscosity coefficient η to be:

$$\eta = \frac{l\overline{u}}{3}N_\mathrm{V}m\,. \tag{17.24}$$

3. *Diffusion of energy: heat conduction* as in Fig. 17.10. Each molecule transports the additional energy $\frac{1}{2}fkT$, and all the molecules together transport in this way a thermal energy Q (f = number of degrees of freedom, k = BOLTZMANN constant). For the *energy current*, we find:

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\frac{l\overline{u}}{3}A\frac{1}{2}N_\mathrm{V}fk\frac{\mathrm{d}T}{\mathrm{d}x} \tag{17.25}$$

or

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\lambda A\frac{\mathrm{d}T}{\mathrm{d}x} \tag{17.18}$$

with the "thermal conductivity coefficient" λ[C17.8]

$$\lambda = \frac{l\overline{u}}{6}N_\mathrm{V}fk = \frac{1}{3}l\overline{u}\varrho c \tag{17.26}$$

(ϱ = density, c = specific heat).

C17.8. This expression (the right-hand term in Eq. (17.26)) holds also for the heat conductivity in electrically-insulating solids (see Fig. 17.8), where c is the specific heat contribution due to lattice vibrations (see Fig. 14.5) that propagate through the crystal lattice at the velocity u as elastic waves (then u replaces the average molecular velocity $\overline{u}$).

17.11 The Mutual Relations of Transport Processes in Gases

Up to now, we have treated the various transport processes as independent of one another. This is quite correct in a first approximation. However, when we look more closely, we find experimentally that the different transport phenomena are mutually interrelated. We offer four examples of these interrelations:

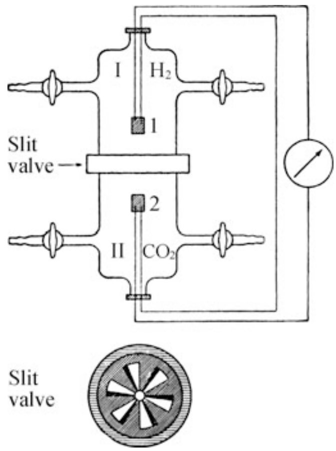


Figure 17.16 The production of a temperature difference by diffusion. The smaller image below shows the brass slit valve which separates the two chambers *I* and *II*. The thermocouples 1 and 2 are made from silver foils with welded-on wires of steel and Constantan.

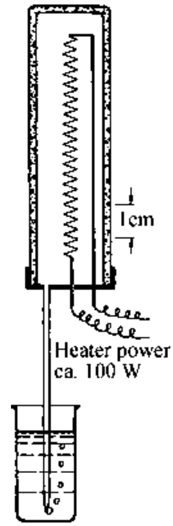
1. *Diffusion in gases produces temperature differences* and they in turn lead to heat conduction. In Fig. 17.16, chamber *I* contains hydrogen, and thus a gas with a small molar mass, $M_\mathrm{m} = 2\,\mathrm{g/mol}$. Chamber *II* contains carbon dioxide, with $M_\mathrm{m} = 44\,\mathrm{g/mol}$. Both of these gases are at the same pressure and temperature; *1* and *2* are thermocouples connected in opposition. – By a slight rotation around the long axis of the chambers, we can open a slit valve which separates the two chambers (cf. small image below!). Then the two gases can mutually diffuse into one another. After we open the valve, for around 30 seconds, a temperature difference of about 0.6 °C is established, with chamber *II* at a lower temperature.

Explanation: The small H_2 molecules quickly penetrate into the CO_2 via isothermal diffusion, so that the number density N_V and the pressure p temporarily drop in chamber *I*. In order to restore their original values, the gas in chamber *II* expands adiabatically, thereby compressing the contents of chamber *I* and performing external work. As a result of this work, the gas in chamber *II* cools; the temperature thus exhibits a gradient, increasing in the direction $II \to I$, in the same direction as the diffusion of the heavier molecules (CO_2).

2. *Temperature differences produce a pressure difference in gases (*KNUDSEN *effect).* In Fig. 17.17, a portion of room air is confined within a porous clay cell. Inside the cell, there is an electric heater. As a result, the temperature in the tiny channels of the clay walls is higher near the inside of the walls than on the outside. A glass tube which extends below the cell into a beaker of water allows the air in the cell to expand and escape to the outside (bubbles!). We observe a continual stream of escaping air: Room air is constantly pulled through the porous clay walls into the heated cell, and as a result, the pressure within the cell is higher than outside it.

The explanation starts with the condition for a stationary state. The number of molecules passing through a cross-sectional area A in the time $\mathrm{d}t$ is $N_\mathrm{V}\overline{u}A\,\mathrm{d}t$. Assuming that the same number enter and leave the channels

Figure 17.17 Porous clay cylinder for demonstrating the KNUDSEN effect



C17.9. In fact, Eq. (17.27) is a *continuity condition*. The air molecules outside the cell at (2) are in equilibrium at the (lower) temperature T_2. They enter the channels with the net average (one-dimensional) velocity $\overline{u_{ch,2}} = \overline{u_2}$. The mean free path l plays no role within the channels. There, the molecules are rapidly thermalized to the (higher) inner temperature T_1 and exit at (1) with the somewhat higher velocity $\overline{u_{ch,1}} = \overline{u_1}$.
Their density N_V has decreased slightly.
Since no molecules are lost or gained within the channels, their entering molecular current dN_2/dt is equal to the exiting current dN_1/dt. These currents are given by $(N_V\overline{u}A)_{1,2}$ (A is the net cross-sectional area of the channels). This gives Eq.(17.27), taking $A_1 = A_2$.
The rest of the calculation leading to Eq. (17.28) is simple manipulations of Eq. (16.1), making use of the ideal gas equation (14.23).

in the cell walls in a given time, and that the channels on average have a uniform cross-section, this leads to:

$$(N_V\overline{u})_1 = (N_V\overline{u})_2\,, \qquad (17.27)$$

where the indices 1 and 2 refer to the quantities on the hot and the cool sides of the porous wall, respectively.
This equation (17.27)[C17.9] is combined with the ideal gas law,

$$p = N_V kT \qquad (14.23)$$

and (from the kinetic theory of gases, Chap. 9) with

$$\frac{1}{2}m\overline{u^2} = \frac{3}{2}kT\,, \qquad (16.1)$$

yielding after some simple manipulations:

$$\frac{p_1}{\sqrt{(T)_1}} = \frac{p_2}{\sqrt{(T)_2}}\,. \qquad (17.28)$$

That is, when the temperature is different on the two sides of the porous wall, the pressures also become different.

3. *In gas mixtures, temperature differences produce concentration gradients (thermodiffusion).* The KNUDSEN effect which we discussed above under Point 2 requires only one species of gas molecules. It was simply convenient to describe it there using a mixture of gases, namely room air.

We could leave out the porous wall, use a mixture of gases instead of a pure gaseous substance, and maintain a temperature gradient within the gas mixture. Then the molecules of higher mass will be enriched in the cooler region; the heavier molecules thus move in the

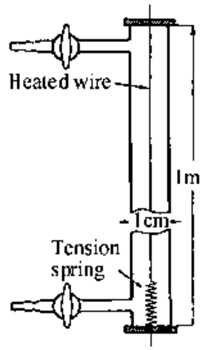


Figure 17.18 Separation of a gas mixture by thermodiffusion in a "separation tube". A wire, tightly stretched in the middle of the tube, is heated until it glows in a mixture of CO_2 and H_2 (partial pressures $\approx$ 0.37 bar and 0.13 bar; the CO_2 should be filled first!). In about 5 minutes, the hydrogen becomes enriched in the upper part of the tube, and prevents the wire from glowing there due to its good heat conductivity. (For the demonstration setup, the tube length is 1 m and its inner diameter is 1 cm. The wire should be well centered and the long axis of the tube exactly vertical). One can also use a mixture of argon and bromine vapor. Then, the enrichment of bromine at the lower end of the tube causes it to liquefy there.

direction of decreasing temperature. This phenomenon is called *thermodiffusion*. It was applied by K. CLUSIUS very successfully for the separation of molecular mixtures, in particular mixtures of different isotopes. His "separation tube" consists of a long, vertically-mounted glass tube with an electrically-heated wire along its central axis. The warm mixture of gases rises near the center of the tube, while cooler gas sinks near the outer wall. The molecules of greater mass diffuse preferentially radially outwards and are carried down by the gas flow at the perimeter of the tube so that they are enriched near the bottom. Figure 17.18 shows a demonstration setup. Thermodiffusion also occurs when one of the "molecular species" consists of larger particles. For example: Warm air rises above a radiator, and between the radiator and the cooler wall, there is a temperature gradient. Dust accumulates in front of the wall, so that it becomes soiled with streaks of dust. – When a pot is used for cooking over an open fire, small particles of soot from the hot flame gases drift towards the bottom of the pot and cover it with a layer of carbon black.

The explanation of thermodiffusion also follows from Eq. (17.20). We must simply take into account that N_V and u in Fig. 17.15 are, in a precise description, somewhat different on each side of the surface A when a temperature gradient is present along the x-direction. – In Eq. (17.20), we replace the number density N_V by $3p/m\overline{u^2}$ (this expression follows from Eq. (9.14) with the density $\varrho = N_V m$). Then we obtain

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -A\frac{pl}{m}\frac{\mathrm{d}(1/\overline{u})}{\mathrm{d}x}\,. \tag{17.29}$$

Now using

$$\frac{1}{2}m\overline{u^2} = \frac{3}{2}kT \tag{16.1}$$

we find after some straightforward manipulations

$$\frac{\mathrm{d}N}{\mathrm{d}t} = +A\frac{lp}{2T\sqrt{3mkT}}\frac{\mathrm{d}T}{\mathrm{d}x}\,. \tag{17.30}$$

Thus, a molecular flow results, and it moves in the direction of increasing temperature. This flow is stronger for the lighter molecules in a gas mixture than for heavier ones. In the stationary state, therefore, the lighter molecules are enriched towards the hotter side and the heavier molecules are enriched towards the cooler side.

4. *Pressure differences in gases produce temperature differences.* Figure 17.19 shows a "vortex tube", above as a section along its length, and below in cross-section at the location b. At this point, air enters tangentially into the tube at a high pressure p. Centrifugal force ensures that the pressure is higher just inside the walls of the tube than at its center axis. To the right of position b, there is an orifice of about 2 mm diameter. A valve H provides a means to regulate the relative strength of the airflows which are moving to the right and to the left. The airstream which exits the tube at the right is cool, while the stream exiting at the left is warm. With $p = 6\,\text{bar}$, we can readily obtain a temperature difference of 40 °C. Within a short time, the right-hand end of the tube is covered by a thick layer of frost.

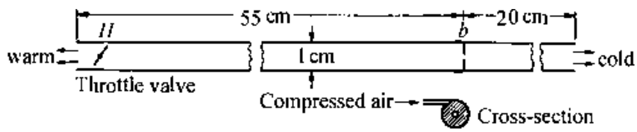


Figure 17.19 Vortex tube (Ranque-Hilsch vortex tube: R. Hilsch, *Z. Naturforschung* **1**, 208 (1946))

The State Function Entropy, *S*

18

18.1 Reversible Processes

All mechanical, electrical and magnetic processes in which – in the ideal limit – no temperature differences occur, are *reversible*. This means that such processes could in principle be made to 'run backwards' along the path that they have taken previously. Their initial state can be restored, *without* causing a *permanent* change of state in any of the components involved. Some examples:

A mechanical or an electrical *oscillation* takes place reversibly; it reproduces its initial state over and over in a periodic sequence.

The *free fall* of a steel ball is likewise reversible, but restoring its initial state requires some *auxiliary apparatus*, e.g. a hard steel plate as in Fig. 5.10. With its help, the accelerated downwards motion can be converted into an upwards motion. The steel plate is not changed in any lasting way in this process; it serves only as a temporary storage medium for potential energy.

A third example of a reversible process can serve to explain the concept of "quasi-static". It is illustrated in Fig. 18.1. The force F of a stretched helical spring and the weight F_G of a mass are kept always very close to equilibrium; this is accomplished by a lever system whose leverage is continuously variable. Then an arbitrarily small difference between F and F_G can precipitate motion in the one or the other direction. The initial state can thus be restored at any time. This process must proceed very slowly, and thus practically without

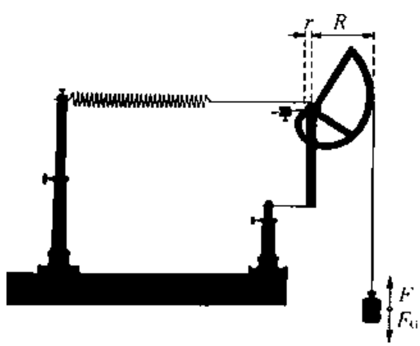


Figure 18.1 The quasi-static expansion of a stretched spring

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_18

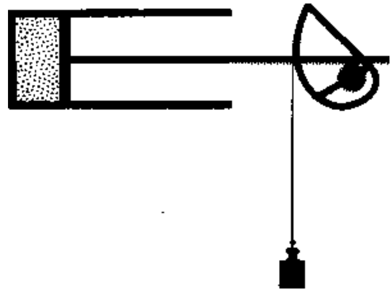


Figure 18.2 Quasi-static expansion of a working substance, for example compressed air

acceleration. Such a process is called *quasi-static*. *Briefly, we define a quasi-static process as a sequence of equilibrium states.*

Processes in which temperature differences occur *may* also take place quasi-statically. They *must* take place quasi-statically if they are to be considered reversible. Examples:

Figure 18.2 shows the reversible, quasi-static expansion of a gas. The variable leverage has to be adjusted to fit the particular gas.

As a second example, we consider the reversible, quasi-static conversion of a liquid into a gas or *vice versa*. In Fig. 18.3, we see a cylinder with a piston. Below the piston is a liquid, with its gaseous phase (vapor) between the piston and the liquid surface. The piston is pressed down by a weight; above the piston, the cylinder has been evacuated. The pressure of the piston can be adjusted by choosing the weight so that the pressure of the piston is practically equal to the saturation vapor pressure of the liquid substance, Case *B*. Then the piston can either rise very slowly and convert all the liquid into vapor, Case *A*; in this case, heat energy must be supplied from the surroundings. Or else it can descend very slowly and condense all the gas to liquid, Case *C*; in this case, heat must be given up to the surroundings. Both processes, *A* and *C*, take place extremely slowly, i.e. *quasi-statically*, and therefore *reversibly*. *Arbitrarily small* temperature differences suffice to push the process in one or the other direction.

Summary: All reversible processes are characterized by three features: Reversible processes can be brought back to their initial states

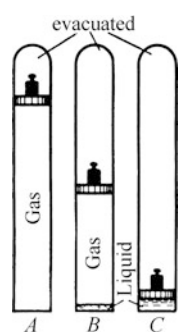


Figure 18.3 Reversible evaporation

(if necessary using auxiliary apparatus) by simply following their path backwards. Restoring their initial states requires *no* net expenditure of energy, and they do not cause a *permanent* change of state in *any* of the bodies involved.

18.2 Irreversible Processes

The converse of reversible processes are the *irreversible* processes. Among them are in particular diffusion, the expansion of gases through a throttle or a nozzle, or into a vacuum, external and internal friction, plastic deformation of objects, heat conduction (when the temperature differences are not vanishingly small), energy input by radiation, as well as all those chemical reactions which do not run their course with infinite slowness.

Irreversible processes are characterized by *three features*:

1. *All irreversible processes proceed spontaneously in only one direction.* This corresponds to our daily experience. The molecules of a perfume which are diffusing into the air in a room never return voluntarily to the open bottle from which they came. An object whose motion has been slowed by air friction will never be accelerated again by the air molecules so that it regains its original velocity. A portion of the air outside a house will never give up its internal energy to heat our apartment or even to fire up the boiler of a steam locomotive. When a stone falls on the floor, it suffers an inelastic collision there and comes to rest. We never experience the converse of that process: No one has ever observed a stone that one day rose back up spontaneously. These possibilities could readily be reconciled with the First Law of thermodynamics, but the molecules never realize them in practice. They can be depended on to divide up a large fortune (of energy, for example); but they will never agree to transfer that fortune voluntarily to a single, distinguished individual (the perfume bottle, the stone etc.).

2. *In all irreversible processes, work is wasted*, i.e. the possibility of obtaining useful work which exists is always defaulted upon. Instead of useful work, the processes give us only warmer objects. Examples:

A certain amount of air is confined within a cylinder equipped with a movable piston. The air is heated while the piston is held fixed. Thereafter, it gives up energy by conduction and radiation until it has again cooled to room temperature. During this cooling process, work is wasted, and the chance of obtaining useful work is lost: We could have given the heated air the opportunity to push out the piston and to perform work until its expansion had cooled it back to room temperature. – In the first case (cooling by heat conduction), the additional heat energy delivered from the fuel used to heat the air is divided up among the enormous number of individual air molecules in the room and is no longer available for performing macroscopic

Part III

work; in the second case (cooling by expansion while performing work), it can be made use of.

Figure 14.9 shows a gas that is being expanded irreversibly without performing work. We could have placed a turbine in the tube connecting the two pressure cylinders and obtained work from it during the pressure equalization. Instead, the air in the right-hand cylinder, after being cooled by the expansion through the valve between the cylinders (work of acceleration!), is simply warmed back to ambient temperature through internal friction in the right-hand cylinder.

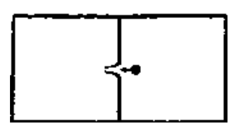
We let a stone which has been lifted up fall back to earth and waste its kinetic energy on impact, warming the ground through work of deformation and friction. If it had been connected to a suitable apparatus, it could have been allowed to sink slowly to the ground and to perform useful work in the process; think of the weights of a pendulum clock and of its chimes.

3. *In* closed *systems, irreversible processes lead to permanent changes of state*. We could indeed restore the original state after an irreversible process has run to completion[1], by replenishing the work which was wasted during the process – *however, there is an essential limitation*: We must not be dealing with a "closed system", i.e. the work must be input to the objects which make up the system from *outside* it, and excess heat must be given up to the outside. For example, the turbine mentioned above would have to be driven backwards as a pump, with the necessary work provided from outside; or the stone would have to be lifted back up using muscle power. In such cases, fuel is burned or food is consumed *outside* the system, and thus the state of some body *outside* the system is *permanently* changed.

"The existence of irreversible processes is a fact confirmed by experience. It has been completely verified by the efforts of many unhappy inventors."

The existence of irreversible processes is a fact confirmed by experience. It has been completely verified by the efforts of many unhappy inventors. Such an inventor might for example try to trick the molecules; we could imagine the setup sketched in Fig. 18.4. It is supposed to disturb the homogeneous temperature distribution within a gas without consuming work. The gas in the left-hand part of the container is to be made warmer, while the gas at the right is to become cooler. Overall, thermal energy will be conserved. The gas at the left could then be used to heat the boiler of a steam engine, while the cool gas at the right could be used to condense its exhaust steam.
How does our inventor proceed? He bores a hole in the partition separating the two halves of the container and closes it on one side with a weir made of fine hairs. His plan is the following: the velocities of the molecules are

Figure 18.4 The irreversibility of temperature equalization



[1] That is why it is preferable to use the word "irreversible", which is derived from Latin, rather than the literal translation "impossible to turn around".

statistically distributed. Only the fastest molecules coming from the right should be able to slip through the weir; the slower ones would rebound from it. Then the fastest molecules, with their higher kinetic energies, would pass over into the left-hand part of the container. There, they will have to share their energy with all the other molecules present; but at least the average kinetic energy would increase on that side. Its temperature would rise, while the temperature at the right would fall. – What is the problem with this "invention"? Answer: The BROWNian motion of the hairs of the weir. They would have to be so fine that they could be moved aside by individual, fast molecules. However, if they were that fine, they themselves would act as large "molecules" in the statistical distribution of the thermal motions. The weir would open and close *statistically*. Often enough, it would be open just when an undesirable molecule with only a small kinetic energy arrived at the partition. Thus, on the average, nothing would be gained and both halves of the chamber would maintain the same temperature.

18.3 Measurement of the Irreversibility Using the State Function *entropy*

Completely irreversible processes occur frequently[2]. Figure 18.5 shows an irreversible expansion as discussed in Sect. 14.8 in connection with an experiment using the steel-ball molecular model. Completely reversible processes, in contrast, are ideal limiting cases; all real processes are only *partially* reversible, they always contain irreversible contributions.

These facts made it necessary to characterize reversible processes and to measure the degree of their irreversibility. This goal was met by discovering a new state function, called the *entropy*. One can arrive at this new function by the following route:

In Fig. 18.6, a gas consisting of an amount of substance n at a constant temperature T_1 can be allowed to expand in a reversible, quasi-static manner. Its volume is allowed to increase from V_1 to V_2, while the pressure decreases from p_1 to p_2. The work performed by the gas

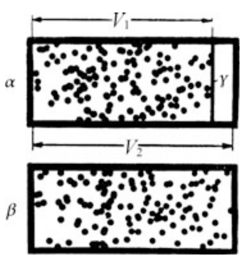


Figure 18.5 A steel-ball model of a gas; at the top *before*, and at the bottom *after* a small expansion from a net volume V_1 to V_2 (cf. Sect. 18.4)

[2] For example, all the motions treated in Sect. 5.11 are irreversible. In such motions, temperature differences occur due to internal and external friction.

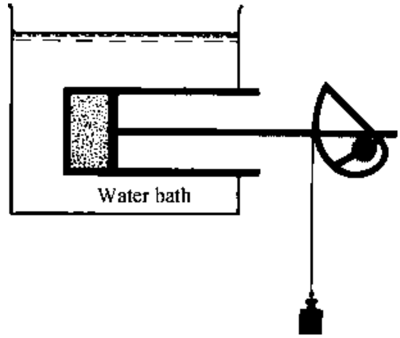


Figure 18.6 The definition of a reversible process

during this isothermal process is

$$W_1 = -T_1 nR \ln \frac{V_2}{V_1} = -T_1 nR \ln \frac{p_1}{p_2} \qquad (14.35)$$

(n = amount of substance (the gas), R = universal gas constant).

At the same time, the gas takes up an amount of heat energy $Q_{\text{rev}(1)} = -W_1$ reversibly from the large water bath, thus keeping its temperature constant.

This heat which is taken up is however *not* uniquely connected with the expansion; it is *not* determined only by the initial and final states; it it thus not a *state variable*.

We can show this by carrying out the expansion of the gas, maintaining *the same initial and final states*, via a different *path*. For this purpose, we extract a quantity of heat Q_{rev} from the whole system in Fig. 18.6 reversibly before the expansion starts by using a suitable auxiliary apparatus, and thereby reduce its temperature to T_2. Then we carry out a slow isothermal expansion at this lower temperature, and for the heat that is taken up, we now find only

$$Q_{\text{rev}(2)} = T_2 nR \ln \frac{V_2}{V_1} \,.$$

At the end of the expansion, we put back reversibly the heat Q_{rev} which we had previously removed, and thus restore the initial temperature T_1.

Therefore, we have the same initial state, namely V_1 and T_1, and the same final state, namely V_2 and T_1; but nevertheless, $Q_{\text{rev}(1)}$ and $Q_{\text{rev}(2)}$ are different, because the "path" was different! – In contrast, however, the quotient

$$\frac{\text{Reversibly absorbed heat}}{\text{Temperature during the heat uptake}}$$

is the same in both cases, namely

$$\frac{Q_{\text{rev}(1)}}{T_1} = \frac{Q_{\text{rev}(2)}}{T_2} = nR \ln \frac{V_2}{V_1} \,. \qquad (18.1)$$

This quotient is independent of the path taken by the process; it is thus a *state function*. We can use this state function as a measure of the irreversibility. It has been given its own name, *entropy*[3].

For the potential energy of a body or its internal energy, the zero point is always arbitrary; we can measure only *changes* or *differences* in these quantities. The same applies to the state function *entropy*: Its zero point is also arbitrary. The special case that we have just considered, a reversible expansion at constant temperature, simply makes a contribution to the entropy already present, since of course the ideal gas in our example had already taken up thermal energy or released it at some temperature or other. Therefore, we finally define the *entropy increase* by

$$\Delta S = S_2 - S_1 = \frac{Q_{\text{rev}}}{T}\,. \tag{18.2}$$

In Fig. 18.6, the expansion was performed reversibly. The "system" consisted of a large water bath and a cylinder containing the confined gas. The gas thus took up heat quasi-statically ($Q_{\text{rev}(1)}$), while the water bath gave up heat quasi-statically ($-Q_{\text{rev}(1)}$). According to the defining equation (18.2), in this reversible process, the entropy of the gas increased by $\frac{Q_{\text{rev}(1)}}{T_1}$ and that of the water bath decreased by $-\frac{Q_{\text{rev}(1)}}{T_1}$. For a reversible process, we thus have

$$\sum \frac{Q_{\text{rev}}}{T} = 0\,. \tag{18.3}$$

Thus, a reversible process in a closed system produces no change in its entropy. We can henceforth use Eq. (18.3) for such a system as the defining feature of a reversible process.

A completely reversible process is an ideal limiting case which cannot be attained in practice. All real processes are more or less irreversible. Then for a closed system, we find[C18.1]

$$\sum \frac{Q_{\text{rev}}}{T} > 0\,. \tag{18.4}$$

As an example, we choose the irreversible process of heat conduction. It is likewise presumed to occur in a system composed of two parts. The thermal energy Q_{rev} is given up by the system at the higher temperature T_1 and is taken on at the lower temperature T_2. In this process, the entropy of the hotter body *decreases* by $-Q_{\text{rev}}/T_1$, while

[3] The name was coined by R. CLAUSIUS in 1865 from the ancient Greek for "turning towards", referring to what he had previously called the "transformation contents".

C18.1. Here, we are dealing with entropy *production*. Corresponding to the observations of irreversible processes described in Sect. 18.2, their entropy can only increase. The magnitude of this entropy production is the measure of the irreversibility of the process. – In the limiting case of reversible processes, the entropy production is zero, so that the entropy of the system remains constant, i.e. it obeys a conservation law (compare Sects. 18.5 and 18.6).

that of the cooler body increases by an amount Q_{rev}/T_2. The sum $Q_{\text{rev}}/T_2 - Q_{\text{rev}}/T_1 = \Delta S$ is thus positive. This *entropy increase* ΔS of the system is a unique measure of the *irreversibility* of the observed heat conduction phenomenon.

18.4 Entropy in a Molecular Picture

We have so far derived the entropy only for a special case. Nevertheless, we will apply the defining equation

$$\Delta S = \frac{Q_{\text{rev}}}{T} \tag{18.2}$$

quite generally. In order to justify that, we wish to clarify the role of the ratio Q_{rev}/T in the molecular picture. In this picture, we will be able to "visualize" the state function *entropy* just as clearly as other state variables and functions, i.e. the temperature, pressure, internal energy and enthalpy. This level of intuitive understanding is possible as a rule only for the simple case of ideal gases.

We refer to Fig. 18.5 and imagine how the corresponding model experiment would function. The small volume V_1 is the x-th fraction of the large volume V_2, i.e. $V_1 = V_2/x$. In the volume V_2, we initially assume that only a single molecule is present. It can be located with absolute certainty, and thus with the probability $w_2 = \frac{1}{1}$, *somewhere* within the volume V_2, but only with the probability $w_1 = 1/x$ within the x-th partial volume, i.e. within the volume V_1. This means that with x observations, on statistical average we find it *once* in the volume V_1. For two molecules, the probabilities of finding both molecules simultaneously in V_2 or in V_1 are:

$$w_2 = \frac{1}{1}, \quad w_1 = \left(\frac{1}{x}\right)^2;$$

and for three molecules,

$$w_2 = \frac{1}{1}, \quad w_1 = \left(\frac{1}{x}\right)^3;$$

and for N molecules,

$$w_2 = \frac{1}{1}, \quad w_1 = \left(\frac{1}{x}\right)^N. \tag{18.5}$$

The *ratio* $W = w_2/w_1$ denotes how much more probable it is to find all the molecules in V_2 instead of in V_1. We obtain

$$W = x^N$$

Part III

or

$$\ln W = N \cdot \ln x \,. \tag{18.6}$$

With $N = nR/k$ (Sect. 14.6) and $x = V_2/V_1$, we obtain

$$k \ln W = nR \ln \frac{V_2}{V_1}$$

or, together with Eqns. (18.1) and (18.2),

$$\Delta S = k \cdot \ln W \,. \tag{18.7}$$

The increase of the entropy which accompanies the expansion of an ideal gas can thus be reduced to the ratio of two probabilities. We also require the value of the universal constant $k = 1.38 \cdot 10^{-23}$ W s/K. *An increase in the entropy means a transition to a state of higher probability*. In Fig. 18.5, finding *all* of the gas molecules in the partial volume V_1 is not impossible, but rather only extremely improbable. That already holds for the relatively few molecules of the model gas, and *a fortiori* for the enormous number of molecules in a genuine gas. The relationship between entropy and probability was recognized by LUDWIG BOLTZMANN (1844–1906). For this reason, the constant k carries his name.

Part III

Think of a mixture of ice and water at 0 °C. In the ice, the molecules are arranged very regularly in the form of a crystal lattice; this is a very improbable state. In the water, they are in a more probable state, with a large degree of randomness. As a result, the entropy of the water is considerably greater than that of the same amount of ice. Nevertheless, a thermally insulated block of ice does not spontaneously convert even a part of itself into water. That would lead the whole system to an extremely improbable state; a portion of the ice would have to cool below 0 °C in order to provide the necessary *heat of melting* for the rest. That would *reduce* the entropy of the whole closed system: The entropy of the ice would have to decrease *more* at a temperature *below* 0 °C than the entropy of the water at 0 °C would increase through the addition of heat.
Another example is perhaps intuitively clearer. The letters which make up this text are in an extremely improbable state; they therefore have a much lower entropy than if they were simply jumbled up in a box with no kind of order. Nevertheless, the letters of this text will certainly not spontaneously convert themselves into a disordered heap; such a transition would have to pass through an extremely improbable intermediate state: Many of the letters would have to accumulate a very great thermal energy and – as "molecules" – use it to jump over their neighbors.

18.5 Examples of the Calculation of the Entropy

Working through examples and applications is always the fastest way to become accustomed to a new physical concept and to be able to use it productively. To this end, we will first calculate the state function entropy for several important cases, and then in Sect. 18.6 we will treat the first applications of the calculated values. – In order to measure the state function entropy, we will always have to employ *quasi-static*, reversible processes; that follows directly from its definition in Sect. 18.3.

C18.2. The *increase* in entropy in all the examples treated in this section always refers to the substance considered, which exchanges entropy with its surroundings via reversible processes. For the overall system, the entropy remains constant, in accord with Eq. (18.3). Thus, we are not dealing with true entropy *production*, but rather with entropy *exchange* (accompanied by a transfer of heat).

1. *Entropy increase*[C18.2] *on melting*. Let us assume that an object has a mass M and a specific heat of melting l_f. Its melting point is T. The process of melting takes place in surroundings which are only slightly higher in temperature. The latent heat Ml_f is thus taken up by the object at practically the melting temperature, i.e. reversibly. In that case, the entropy of the melting object increases by the amount

$$\Delta S = \frac{Ml_f}{T} \,. \tag{18.8}$$

Numerical example
for water at standard atmospheric pressure (1013 hPa)

$$T = 273\,\mathrm{K}\,, \quad l_f = 3.34 \cdot 10^5\ \mathrm{W\,s/kg}\ \text{(see Table 13.1)}\,.$$

Then the increase in the specific entropy is

$$\frac{\Delta S}{M} = 1.22 \cdot 10^3\ \frac{\mathrm{W\,s}}{\mathrm{kg\,K}}$$

or, with the molar mass $M_\mathrm{m} = M/n = 18$ g/mol, the molar entropy increase becomes

$$\frac{\Delta S}{n} = 22\ \frac{\mathrm{W\,s}}{\mathrm{mol\,K}} = 2.64\,R$$
$$(R = 8.31\ \mathrm{W\,s/(mol\,K)}).$$

For mercury ($M_\mathrm{m} = 200$ g/mol), the corresponding numbers are

$$T = 234.1\,\mathrm{K}\,, \quad l_f = 11.8 \cdot 10^3\ \frac{\mathrm{W\,s}}{\mathrm{kg}}\,, \quad \frac{\Delta S}{n} = 10\ \frac{\mathrm{W\,s}}{\mathrm{mol\,K}} = 1.20\,R\,.$$

On reversible melting, the entropy of the surroundings decreases by the same amount as it increases within the object that melts. Thus, as in every reversible process in a closed system, the overall amount of entropy remains unchanged. – The corresponding conclusion holds also in all the following examples.

2. *Entropy increase on heating*. A substance of mass M is heated at constant pressure from an initial temperature T_i to a final temperature T_f. In this process, small amounts of heat are transferred at gradually increasing temperatures, and thus reversibly. For the entropy increase of the substance heated, we therefore find

$$\Delta S = \frac{\Delta Q_{\mathrm{rev}(1)}}{T_1} + \frac{\Delta Q_{\mathrm{rev}(2)}}{T_2} + \cdots = \sum_\mathrm{j} \frac{\Delta Q_{\mathrm{rev(j)}}}{T_\mathrm{j}}\,, \tag{18.9}$$

$$\Delta S = M\left(\frac{c_{\mathrm{p}1}\Delta T}{T_1} + \frac{c_{\mathrm{p}2}\Delta T}{T_2} + \cdots\right) = M\sum_\mathrm{j} \frac{c_{\mathrm{pj}}\Delta T}{T_\mathrm{j}}\,, \tag{18.10}$$

Part III

Table 18.1 Specific state variables for water along its vapor-pressure curve. The reference point for the enthalpy and the entropy is chosen to be 0 °C.

		Liquid			Saturated vapor		
Temperature	Vapor pressure	$\frac{\text{Volume } V}{\text{Mass } M}$	$\frac{\text{Enthalpy } H}{\text{Mass } M}$	$\frac{\text{Entropy } S}{\text{Mass } M}$	$\frac{\text{Volume } V}{\text{Mass } M}$	$\frac{\text{Enthalpy } H}{\text{Mass } M}$	$\frac{\text{Entropy } S}{\text{Mass } M}$
(°C)	$\left(10^5\,\frac{\mathrm{N}}{\mathrm{m}^2}\right)$	$\left(\frac{\mathrm{m}^3}{\mathrm{kg}}\right)$	$\left(10^5\,\frac{\mathrm{W\,s}}{\mathrm{kg}}\right)$	$\left(10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}}\right)$	$\left(\frac{\mathrm{m}^3}{\mathrm{kg}}\right)$	$\left(10^5\,\frac{\mathrm{W\,s}}{\mathrm{kg}}\right)$	$\left(10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}}\right)$
17.2	0.02	0.001	0.724	0.255	68.3	25.33	8.71
59.7	0.20	0.001	2.495	0.829	7.79	26.09	7.91
99.1	0.98	0.001	4.149	1.298	1.73	26.71	7.37
151	4.9	0.0011	6.364	1.851	0.382	27.47	6.83
211	19.6	0.0012	9.043	2.437	0.101	27.97	6.36
310	98	0.0014	13.984	3.345	0.0185	27.26	5.61
374	221	0.0037	20.625	4.313	0.0037	22.07	4.61

or, in the limit of infinitely small increments and with constant specific heat,

$$\Delta S = Mc_\mathrm{p}\int_\mathrm{i}^\mathrm{f}\frac{\mathrm{d}T}{T} = Mc_\mathrm{p}\,\ln\frac{T_\mathrm{f}}{T_\mathrm{i}}\,, \tag{18.11}$$

and, with $Mc_\mathrm{p} = nC_\mathrm{p}$ (see Eqns. (13.5) and (13.6)),

$$\Delta S = nC_\mathrm{p}\,\ln\frac{T_\mathrm{f}}{T_\mathrm{i}}\,. \tag{18.12}$$

Numerical example

for water on heating from its melting point up to its boiling point at standard pressure (1013 hPa). Its heat capacities are nearly constant over this temperature range:

$$T_\mathrm{i} = 273\,\mathrm{K}\,,\quad T_\mathrm{f} = 373\,\mathrm{K}\,,$$

$$c_\mathrm{p} = 4.19\cdot 10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}}\quad\text{or}\quad C_\mathrm{p} = 75.5\,\frac{\mathrm{W\,s}}{\mathrm{mol\,K}}\,,$$

$$\frac{\Delta S}{M} = 1.31\cdot 10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}}\quad\text{or}\quad \frac{\Delta S}{n} = 23.6\,\frac{\mathrm{W\,s}}{\mathrm{mol\,K}} = 2.84\,R\,.$$

Table 18.1 gives some additional values of $\Delta S/M$ at various temperatures. These values play an important role in technical applications. To simplify the calculations, the specific entropy of liquid water at 0 °C (273 K) and standard atmospheric pressure (1013 hPa) is arbitrarily set equal to zero. We make use of this convention in quoting measured values, and denote the entropy thus defined by S.

3. *Entropy increase on evaporation:* PICTET-TROUTON *rule.* Let a liquid have the mass M and the specific heat of vaporization l_v. It is evaporated at constant pressure (its saturation vapor pressure) and the corresponding temperature T. Then for the increase in entropy, we have

$$\Delta S = \frac{Ml_v}{T}\,. \tag{18.13}$$

Numerical example

For water at standard pressure (1013 hPa), $T = 373\,\mathrm{K}$ and $l_v = 2.26 \cdot 10^6\,\mathrm{W\,s/kg}$ (see Fig. 14.3); then we find

$$\frac{\Delta S}{M} = 6.06 \cdot 10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}} \quad \text{or} \quad \frac{\Delta S}{n} = 109\,\frac{\mathrm{W\,s}}{\mathrm{mol\,K}} = 13.1\,R\,.$$

For the molar entropy increase $\Delta S/n$, quite similar values for the evaporation of many other substances are obtained. This is the content of "TROUTON*'s rule*"[C18.3].

C18.3. See e.g. A. Eucken and E. Wicke, "*Grundriss der physikalischen Chemie*", Akad. Verlagsges., 10th ed. (1959), Chap. 2. Apparently published independently by R. PICTET (1876) and R.T. TROUTON (1883). English: see JAIME WILNIAK, *Chemical Educator* **6** (2001), p. 55. Available online at https://de.scribd.com/document/71366734/Frederick-Thomas-Trouton-The-Man-The-Rule-And-the-Ratio

When liquid water is converted into water vapor, the increase in its entropy is thus about five times greater than when ice is converted into liquid water ($2.64\,R$). When liquid *water* at 0 °C is converted into saturated water *vapor* at 100 °C, the molar entropy of the water increases by

$$\frac{\Delta S}{n} = \underset{\substack{\text{on heating}\\ \text{(Pt. 2)}}}{2.84\,R} + \underset{\substack{\text{on vaporization}\\ \text{(Pt. 3)}}}{13.1\,R} \approx 16\,R$$

or, referred to its mass,

$$\frac{\Delta S}{M} = (1.31 + 6.06) \cdot 10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}} = 7.37 \cdot 10^3\,\frac{\mathrm{W\,s}}{\mathrm{kg\,K}}\,.$$

This latter quantity is called the specific entropy of the saturated vapor. Values for other temperatures can be found in Table 18.1.

4. *Entropy changes associated with changes of state of ideal gases.* We look at the p-V diagram in Fig. 18.7 and proceed in two steps from the state 1 (temperature T_1) to the state 2 (temperature $T_2 > T_1$). First, the gas (amount of substance $= n$) takes up heat at constant pressure along the path $1 \to 3$; and then along the path $3 \to 2$, it receives work at constant temperature. The constancy of the temperature T_2 is possible only if a quantity of heat equivalent to the work is given up by the gas to its surroundings. Then, making use of Eqns. (18.12) and (14.35),

$$\Delta S = n \int_1^3 \frac{C_\mathrm{p}\mathrm{d}T}{T} - \frac{nRT_2 \ln \frac{p_2}{p_3}}{T_2}\,. \qquad (18.14)$$

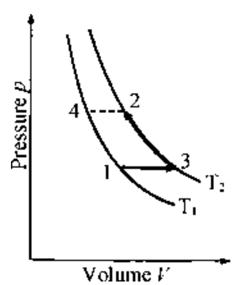


Figure 18.7 Calculating the entropy of ideal gases

Now, $T_3 = T_2$, $p_3 = p_1$. Then we find for the increase of the molar entropy

$$\frac{\Delta S}{n} = C_p \ln \frac{T_2}{T_1} - R \ln \frac{p_2}{p_1}\,. \tag{18.15}$$

The entropy of an ideal gas thus increases with increasing temperature and decreases with increasing pressure. The entropy of an ideal gas accordingly consists of two parts: The first part depends on the temperature, and the second part depends on a geometrical constraint, namely the available volume, which determines the pressure and the number density. On isothermal compression, the entropy of an ideal gas becomes smaller. – For the derivation of this equation, we could have carried out the transition from state 1 to state 2 along some other arbitrary path, e.g. in the two steps $1 \to 4$ and $4 \to 2$. The entropy is a state function; it is independent of the *path* along which a change of state is accomplished.

18.6 Application of Entropy to Reversible Changes of State in Closed Systems

When reversible processes occur adiabatically, i.e. without any exchange of heat with their surroundings, then the sum of the entropies of all the bodies involved remains constant; the processes are *isentropic*. *This constancy of the entropy in reversible, adiabatic processes is often employed.*

We begin by showing in Fig. 18.8 some adiabatic curves in the p-V_m diagram ($V_m = V/n$) of an ideal gas for reversible expansion: At the lower end of each adiabatic curve, the constant value of the associated molar entropy S/n is noted.

Formation of fog or clouds on adiabatic expansion. Water vapor with a vapor pressure of p_1 is expanded adiabatically, and its pressure decreases to p_2. What fraction y of the water will precipitate as fog droplets? This case plays an important role in meteorology – think of an upwardly-directed current of warm air.

Before the expansion and cooling, the vapor pressure p_1 is associated with the temperature T_1. At this temperature, the water *vapor* of mass M has the entropy S_1. During the expansion and cooling process, a fraction y of the vapor is condensed to liquid water (fog droplets). The mass of the vapor is reduced to $M(1 - y)$, and at its temperature T_2, it retains the entropy $(1 - y)S_2$. Furthermore, water droplets of mass My are formed. At the temperature T_2, this liquid water has the entropy yS'_2. Setting the entropies before and after the condensation equal, we find

$$S_1 = (1 - y)S_2 + yS'_2\,.$$

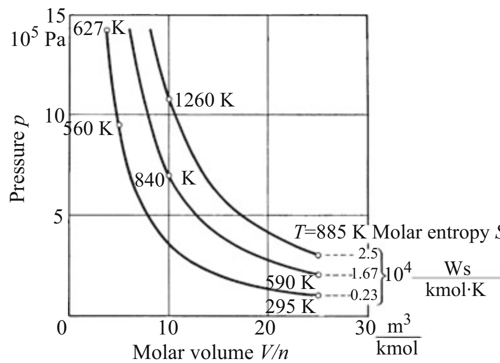


Figure 18.8 Adiabatic curves for a diatomic, ideal gas as curves of constant entropy. The reference point for the entropy has been chosen to be 0 °C and standard atmospheric pressure (1013 hPa) (cf. **Exercise 18.1**).

Part III

Furthermore, we have

$$\underset{\text{vapor}}{S_2} - \underset{\text{liquid}}{S_2'} = \frac{l_v}{T_2} M \,. \tag{18.13}$$

Combining these two equations, we obtain

$$y = \frac{S_2 - S_1}{M} \frac{T_2}{l_v} \,. \tag{18.16}$$

Numerical example for water vapor[C18.4]

$$p_1 = 200\,\text{hPa}\,, \quad T_1 = 333\,\text{K}\,(59.8\,°\text{C})\,, \quad \frac{S_1}{M} = 7.91 \cdot 10^3 \,\frac{\text{W s}}{\text{kg K}}\,.$$

$$p_2 = 20\,\text{hPa}\,, \quad T_2 = 290.3\,\text{K}\,(17.1\,°\text{C})\,, \quad \frac{S_2}{M} = 8.71 \cdot 10^3 \,\frac{\text{W s}}{\text{kg K}}\,.$$

The specific heat of vaporization (latent heat) of water at T_2 is (Fig. 14.3)

$$l_v = 2.45 \cdot 10^6 \,\frac{\text{W s}}{\text{kg}}\,.$$

Result:

$$y = 0.095\,,$$

i.e. 9.5 % of the saturated vapor has been condensed to fog droplets.

C18.4. Determined using Eq. (18.15). Taking $T_0 = 273$ K and $p_0 = 1013$ hPa as starting values, use p_1, T_1 and p_2, T_2 as end points, respectively, to determine S_1 and S_2 (Eq. (18.15)), then convert to specific entropies.

18.7 The *H-S* or Mollier Diagram with Applications. Supersonic Gas Jets

Thus far, we have depicted the states of substances only in p-V or p-T diagrams. Their ordinates represent the pressure, while the abscissa shows the specific volume V/M, or the molar volume V/n, or the temperature T. One could however equally well use other pairs of state variables to construct such diagrams[C18.5].

As one of the many possibilities, in Fig. 18.9 we show an *H*-*S* diagram for the case of air. On the ordinate, the specific enthalpy is plotted, i.e. H/M, and the abscissa represents the specific entropy, i.e. S/M. The values on the ordinate are computed from Eq. (14.27), while those on the abscissa are found with Eq. (18.15). In both cases, the temperature dependence of the specific heats has been taken into account (see, for example, Fig. 16.10).

In an *H*-*S* diagram, the adiabatic curves are *straight* lines which run parallel to the ordinate axis. The isotherms, drawn here for several temperatures between $-130\,°\mathrm{C}$ and $+50\,°\mathrm{C}$, are straight lines only at low pressures, and then they are parallel to the abscissa axis. – In a *p*-*V* diagram, the isobars and the isochores would be *straight* lines, while in the *H*-*S* diagram, these lines of constant pressure and of constant volume are *curved*. In Fig. 18.9, only a few isobars are shown for pressures in the range $(0.01 - 200) \cdot 10^5$ Pa.

The *H*-*S* diagram plays an important role in the adiabatic changes of state of flowing substances. One can determine the technical work which can be obtained from a change of state without any computations; only the values on the ordinate need be read off. In the following, we treat an example of an application which is equally important in basic physics and in technology. It concerns the adia-

C18.5. For some examples of MOLLIER diagrams, see http://www.engineersedge.com/thermodynamics/enthalpy_entropy_mollier.htm
Supersonic jets are described for example in https://www.fas.org/sgp/othergov/doe/lanl/pubs/00326958.pdf

Part III

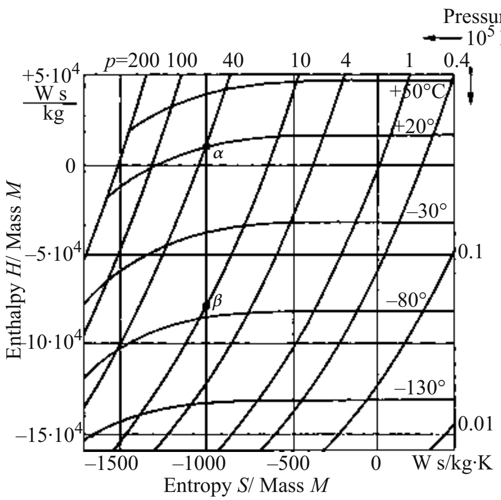


Figure 18.9 A portion of the *H*-*S* or MOLLIER diagram for air (from *Zeitschrift des VDI* **48**, 271 (1904)). R. MOLLIER (Professor in Dresden) introduced the enthalpy as the quantity on the ordinate of state diagrams in 1904. The curves are isotherms and isobars, respectively. The values of the enthalpy and the entropy are referred to standard conditions, 0 °C and an air pressure of 1013 hPa.

batic flow of a gas which is escaping from a pressurized container: a "supersonic jet".

For this example, we choose the gas to be air. The air is compressed to a high pressure p_1 within a pressure cylinder, where the pressure is held constant. The air is allowed to flow out through a nozzle, forming a jet, and enters a space where the ambient pressure p_2 is lower. During the expansion, the air performs work of acceleration and imparts kinetic energy (in one direction!) to itself. How does the resulting jet velocity u depend on the initial and final pressures?

In an adiabatic process, no thermal energy is exchanged with the surroundings. Therefore, in the equation for the First Law, Q can be set to zero. For the work of acceleration of the flowing air, we then find, using Eq. (14.11) and setting the work of acceleration equal to the kinetic energy of the air in the jet, $\frac{1}{2}Mu^2$:

$$H_2 - H_1 = W_{\text{tech}} = \tfrac{1}{2}Mu^2 \,. \tag{18.17}$$

The enthalpy difference H_2-H_1 can be read directly off the H-S diagram for air (Fig. 18.9). We can for example assume that the air in the pressure cylinder is at a pressure of $p_1 = 40\cdot10^5$ Pa and a temperature of 20 °C. Its state is marked in Fig. 18.9 by the point α. The adiabatic expansion proceeds to a final pressure of $p_2 = 10 \cdot 10^5$ Pa. Then the final state of the air is marked by the point β in Fig. 18.9. The difference between α and β along the vertical axis gives the decrease in the specific enthalpy during the expansion. It is

$$-\frac{H_2 - H_1}{M} = 9.6 \cdot 10^4 \, \frac{\text{W s}}{\text{kg}} \,.$$

Inserting this value into Eq. (18.17) shows the final or muzzle velocity of the air jet to be $u = 438$ m/s.

In a similar fashion, the flow velocities for other final pressures p_2 can be obtained, and they are plotted in Fig. 18.10a. In all cases, a constant initial pressure of $p_1 = 40 \cdot 10^5$ Pa has been assumed. – Result: *The jet velocity can be considerably greater than the speed of sound* c_{sound} (= 340 m/s at 20 °C). *Nevertheless, one cannot achieve velocities greater than a certain limiting value* $u_{\max}$. In our example, with an initial pressure of $p_1 = 40 \cdot 10^5$ Pa, the highest obtainable muzzle velocity is $u_{\max} = 760$ m/s. This maximum value is obtained when the air flows out into a *vacuum*.

On expansion, the density of the air (i.e. the quotient $\varrho = M/V$) decreases. This is shown for our example in Fig. 18.10b. The values shown here were calculated with the aid of Eq. (14.42). See also Footnote 1 there.

The mass M of the air that flows out is proportional to the time of flow t, to the density ϱ of the air, to the cross-sectional area A of the nozzle and to the flow velocity u. It is given by the product of these four quantities, that is

$$M = t\varrho A u \,. \tag{18.18}$$

Part III

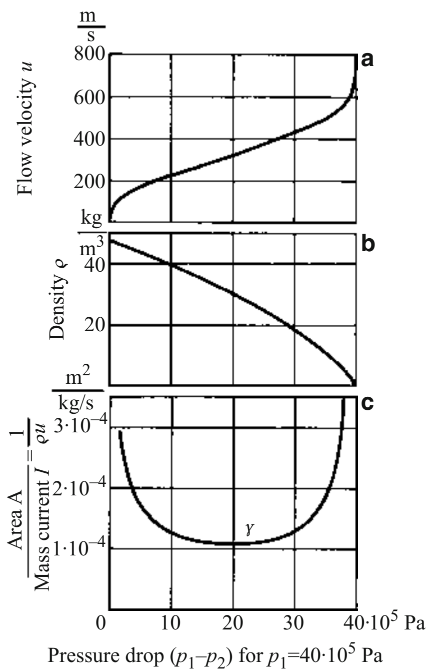


Figure 18.10 The supersonic flow of a gas jet from a nozzle. All three curves apply to an initial pressure of $p_1 = 40 \cdot 10^5$ Pa. The values of u and ϱ are those found *after* the expansion.

With the mass current

$$I = \frac{\text{Mass } M \text{ of the flowing air}}{\text{Time } t}, \tag{18.19}$$

we obtain

$$\frac{A}{I} = \frac{1}{\varrho u}. \tag{18.20}$$

This quantity is plotted in Fig. 18.10c. We consider its content in detail, making use of the other two parts of the figure. We find the following:

In Fig. 18.10a, the curve of the flow velocity hardly deviates from the ordinate axis up to about 70 m/s. Thus, the density curve for low velocities in Fig. 18.10b corresponds to a single fixed point, namely the point on the ordinate axis: The density ϱ is thus constant up to about $1/5$ of the velocity of sound (Sect. 10.1). At "low" velocities, gases behave like non-compressible fluids: The quotient A/I decreases in Fig. 18.10c with increasing values of u. – However, at high velocities, the situation is quite different: Here, the density ϱ decreases rapidly with increasing velocity. As a result, in Eq. (18.20), the increase in u is compensated by a decrease in ϱ, so that the quotient A/I becomes constant over a certain range (in Fig. 18.10c). Later, the decrease in ϱ begins to predominate over the increase of u, and the quotient A/I once again rises. At its minimum (Point γ), the flow velocity is equal to the *velocity of sound*

$$c_{\text{sound}} = \sqrt{\kappa \cdot \frac{R}{M_{\text{m}}} \cdot T} \tag{14.56}$$

(T = temperature of the adiabatically expanded gas at the point of the smallest cross-section).
The minimum of A/I, and thus the maximum of the mass current, is reached when

$$\frac{p_2}{p_1} = \left(\frac{2}{\kappa + 1}\right)^{\frac{\kappa}{\kappa - 1}} .$$

For air, this occurs at an external pressure of $p_2 = 0.53\, p_1$.

This can be derived in the general case, but it can also be understood qualitatively: If the velocity of sound has been attained at the location of the smallest cross-section in the nozzle by a sufficient reduction in the external pressure p_2, then additional "downstream" reductions in the pressure have no further effect. They can propagate at most with the velocity of sound, and therefore cannot overcome the flow velocity and penetrate upstream to the point of smallest cross-section.

If a simple nozzle is used (as shown in Fig. 18.11), the point of smallest cross-section falls together with the muzzle of the nozzle. *Then, in the muzzle of a simple nozzle, the flow velocity can be at most equal to the velocity of sound*. If we want to obtain higher muzzle velocities, then the nozzle must be enlarged conically after its narrowest point (Fig. 18.12). The cross-sectional area A of the nozzle must be adapted to the area required by the mass current I at every point. Then the gas can exit the muzzle with the full velocity predicted by the H-S diagram. *At the narrowest part of the nozzle, the flow velocity is still equal to the velocity of sound*; therefore, the mass current I remains the same as before, without the conical extension of the nozzle.

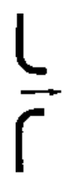


Figure 18.11 Example of a simple nozzle, not suitable for producing supersonic velocities

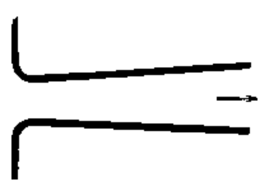


Figure 18.12 A de LAVAL nozzle for producing supersonic gas jets (C. G. P. DE LAVAL, 1845–1913, Sweden)

Exercise

18.1 Using the adiabatic curves shown in Fig. 18.8, determine the corresponding molar entropies and compare them with the values given in the figure. (Sects. 18.4 and 18.6. *Hint*: Refer to Eq. (18.15).)

Converting Heat into Work. The Second Law

19

19.1 Heat Engines and the Second Law[1]

Heat engines were developed as a technological innovation in order to make use of temperature differences for performing mechanical work. The most important types of heat engines, the steam engine and the internal combustion engine, are well known to everyone today. All heat engines *mediate* the transition from a hot region to a cool region by means of a flowing working substance and *repeat* this process *periodically*. The initial state of the machine is repeatedly and periodically reproduced; the only quantity which is permanently consumed is the supply of fuel.

Without the mediation of a machine, temperature differences are equalized by *thermal processes* alone, i.e. through heat conduction and radiation. Both of these processes are irreversible; in both, work is thus wasted, or in other words, an opportunity to perform useful work is not utilized (Sect. 18.2).

In contrast, we could obtain the ideal maximum amount of work that could be performed by eliminating all irreversible effects such as friction, heat conduction and radiation, and instead permit the temperature difference to be equalized via a "heat engine" in a *reversible* manner; that is, all the steps in the process must occur quasistatically. This will be shown in the following.

A working substance takes on the heat energy $Q_{\mathrm{rev}(1)}$ (> 0) from a hot reservoir at the high temperature T_1 in an isothermal and reversible step. When the working substance has flowed to the cool reservoir at the lower temperature T_2, it gives up the smaller amount of heat energy $Q_{\mathrm{rev}(2)}$ (< 0), again in an isothermal and reversible step, to the cool reservoir. Then the excess heat energy $Q_{\mathrm{rev}(1)} + Q_{\mathrm{rev}(2)} = -W$ can be completely converted to mechanical work, assuming that all the other intermediate steps in the machine's operation take place reversibly. In that case, from Eq. (18.3), we see that the sum of all the entropy changes is zero, i.e.

$$\frac{Q_{\mathrm{rev}(1)}}{T_1} + \frac{Q_{\mathrm{rev}(2)}}{T_2} = 0\,. \qquad (18.3)$$

[1] It is suggested that the reader familiarize him- or herself with the properties of irreversible processes (in Sect. 18.2, in particular the examples given under Point 2) before reading this chapter.

K. Lüders, R.O. Pohl (Eds.), *Pohl's Introduction to Physics*, DOI 10.1007/978-3-319-40046-4_19

Part III

The ratio of work W (< 0) performed to heat $Q_{\text{rev}(1)}$[C19.1] consumed,

$$\frac{-W}{Q_{\text{rev}(1)}} = \frac{Q_{\text{rev}(1)} + Q_{\text{rev}(2)}}{Q_{\text{rev}(1)}} = \eta_{\text{ideal}} \tag{19.1}$$

defines the *thermal efficiency* η_{ideal} of an ideal heat engine. The combination of Eqns. (18.3) and (19.1) yields

$$\eta_{\text{ideal}} = \frac{-W}{Q_{\text{rev}(1)}} = \frac{T_1 - T_2}{T_1}\,. \tag{19.2}$$

C19.1. The negative sign of the output quantities is due to the formulation of the First Law (Eq. (14.6)), here $Q + W = 0$.
For a heat engine, $Q > 0$ and $W < 0$. The energy balance is then
$Q = Q_1 + Q_2 = -W$
or
$Q_1 = -W - Q_2$.
(Q_1 is positive, while W and Q_2 are negative.)

The highest theoretically-possible efficiency η_{ideal} of a *heat engine* is thus independent of all the details of its construction and operation. The essential point is only that all possible irreversible processes be eliminated; then the determining factors are simply the higher temperature at which the heat $Q_{\text{rev}(1)}$ is input quasi-statically, and the lower temperature at which the heat $Q_{\text{rev}(2)}$ is output quasi-statically.

Equation (19.2) is a quantitative statement of the Second Law of thermodynamics. Its essential content was formulated in 1824 by SADI CARNOT[C19.2]. CARNOT's considerations were still based on the assumption of a "heat substance" ("phlogiston"). The modern interpretation of Eq. (19.2) and our knowledge of its general applicability are due in the main to RUDOLF CLAUSIUS (1822–1888).

C19.2. S. CARNOT (1796–1832): "*Réflexions sur la puissance motrice du feu et les machines propres a développer cette puissance*", Paris (1832).

The First Law states that the sum of all the energies which contribute to a change of state remains constant. This can be demonstrated experimentally by converting work completely into heat, for example using friction. – The reverse process is not possible: *The Second Law limits the conversion of heat into work.*

For $T_1 = T_2$, from Eq. (19.2), both η_{ideal} and also the work W performed are equal to zero. This fact is the basis of a formulation of the Second Law which is due to PLANCK. It states that, "It is not possible to construct a machine which *does nothing more* than lifting a weight and cooling a heat bath by withdrawing the equivalent amount of heat"[C19.3].

C19.3. This is the reason why the Second Law is often summarized by the statement: "It is impossible to construct a perpetual motion machine of the second kind". That would be a machine which does nothing other than convert heat to mechanical work with 100 % efficiency.

In the isothermal expansion of a gas, the entire amount of heat taken up by the gas is indeed converted to work (see Sect. 14.11). Nevertheless, this does not contradict the Second Law, because in addition to performing work, for example the work of lifting a weight, something else is happening: The *density* of the gas is *reduced* or the amount of pressurized, stored air is decreased.

The Second Law of thermodynamics is also a purely *empirical result*. The discussion above clearly shows this. The 'law' is based on the successes of technology and its experience in constructing heat engines.

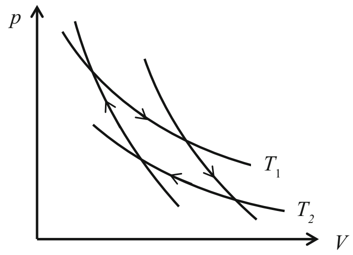
19.2 The CARNOT Cycle

The considerations in Sect. 19.1 stand and fall with the possibility that a working substance can take up an amount $Q_{\mathrm{rev}(1)}$ of heat energy reversibly and can release a smaller amount $Q_{\mathrm{rev}(2)}$ reversibly.

The fundamental possibility of such processes can be demonstrated with a machine which carries out a "CARNOT cycle" (Fig. 19.1). It makes use of a gas which is confined in a cylinder, as shown in Fig. 18.2. This cylinder is first brought into thermal contact with a hot reservoir or heat bath (T_1 in Fig. 18.6). The gas expands isothermally and quasi-statically by taking up the heat $Q_{\mathrm{rev}(1)}$ reversibly from the heat bath and performing mechanical work with it. The cylinder is then thermally isolated from its surroundings and the gas is further expanded adiabatically until it reaches the temperature T_2 of a cooler heat reservoir. During these two expansion steps, the gas performs all together the work W_1.
In the third step of the cycle, thermal contact is made with the cooler heat bath. The gas is compressed isothermally and quasi-statically, thereby giving up the heat $Q_{\mathrm{rev}(2)}$ reversibly to the cooler heat bath. In the fourth step, the cylinder is again thermally isolated and the gas is compressed adiabatically until it reaches its initial temperature T_1. Then the initial state of the system has been restored. During the two compression steps, the work W_2 must be performed on the gas. The net output of work is $Q_{\mathrm{rev}(1)} + Q_{\mathrm{rev}(2)} = -(W_1 + W_2) = -W$. At each changeover between isothermal and adiabatic volume variations, the adjustable lever arm shown in Fig. 18.2 must be set to the appropriate length.

The decisive point in the CARNOT *cycle is that the temperature of the working substance (the gas in the cylinder) is equilibrated to the temperature of the respective heat bath with which it is brought into contact*. This requirement can be fulfilled by a slightly different cyclic process in the STIRLING engine, which can be more readily constructed. It is based, as we shall soon see, on the use of a *regenerator*.

Figure 19.1 CARNOT cycle



19.3 The STIRLING Engine

C19.4. It was invented by ROBERT STIRLING, a Scottish clergyman, in 1816.

This engine[C19.4] was previously used for small-scale industrial processes and as a toy. It demonstrates the essentials of a heat engine in an especially clear way: that is, the mediation of heat transfer between a hot and a cool reservoir by a working substance in a *periodic* process. We will explain its construction and operation by making use of the semi-schematic illustrations in Fig. 19.2, and will employ it several times in order to demonstrate the content of Eq. (19.2) experimentally.

The two heat baths *I* and *II* in Fig. 19.2 at the temperatures T_1 and T_2 ($T_1 > T_2$), respectively, are attached to the left and the right ends of a cylinder. In the cylinder, besides the piston there is a drum *D*. It contains channels along its length. This drum is pushed back and forth within the cylinder by means of a connecting rod to the crankshaft (not shown), with a phase shift of around 90° relative to the piston. The drum fulfills a double function: First, it acts as a *displacer*, moving the working substance (usually air) alternately towards the hot and the cool heat bath. Second, it acts as a *heat storage medium* (*regenerator*). This means that during the displacement of the gas, the latter flows through the channels in the drum. The drum takes up heat from the gas flowing to the right (Part *b*) and gives it up to the gas which is flowing to the left (Part *d*).

The four-step operation of this heat engine is illustrated in the four parts of the figure. At *a*, the air has expanded isothermally by taking up heat at the higher temperature T_1 and is moving the piston to the right. At *b*, the displacer shoves the air towards the cooler heat bath.

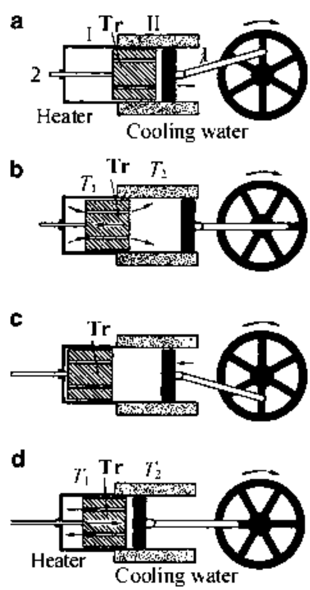


Figure 19.2 The operation of a STIRLING engine. At **a** and **c**, the displacer *D* is at its reversal point, while at **b** and **d**, the motion of the piston is reversed (rest points) **(Video 19.1)**.

Video 19.1:
"Operation of a STIRLING engine"
http://tiny.cc/gigvjy
The operation is explained using a lecture-demonstration model. There, the function of the regenerator is performed not only by the displacer drum, but also by a wire net within the cylinder.

On the way, it is cooled to the lower temperature T_2 while flowing through the channels in the displacer[C19.5]. At c, the piston is moved to the left by the flywheel (stored mechanical energy!) and the air is compressed isothermally at the lower temperature T_2, giving up heat. At d, the displacer shoves the compressed air back towards the hot reservoir. On the way, it is warmed up to T_1 by flowing through the channels, taking up the stored heat from the regenerator. Then the cycle can begin anew: The compressed air can again take on more heat at the higher temperature, expand isothermally and push the piston to the right, performing work. After 1/4 of a rotation, the initial state a has again been reached.

C19.5. This step is an *isochoric* change of state (Sect. 14.9, Point 3). The cyclic process here is composed of two isotherms and two isochores. The important point however (just as in the CARNOT cycle) is that the expansion occurs at the higher temperature and the compression at the lower temperature.

In the ideal limiting case, the work performed by the engine with this cyclic process would be

$$-W = Q_{\text{rev}(1)} \frac{T_1 - T_2}{T_1} . \tag{19.2}$$

Here, $Q_{\text{rev}(1)}$ is the heat energy which is taken up at the higher temperature, causing the gas to expand. This heat uptake is isothermal in the ideal case. Then we have[C19.6]

$$Q_{\text{rev}(1)} = nRT_1 \ln \frac{V_2}{V_1} = nRT_1 \ln \frac{p_1}{p_2} \tag{14.35}$$

(n = amount of substance (air), R = universal gas constant, p_1 (V_1) and p_2 (V_2) are the pressures (volumes) before and after the isothermal expansion).

C19.6. These equations hold under the assumption that the working substance is an ideal gas, which is the case for air near room temperature. The statement of the Second Law (Eq. (19.2)) is however independent of this assumption; it holds quite generally.

Combining Eq. (14.35) and Eq. (19.2) yields

$$-W = nR(T_1 - T_2) \ln \frac{p_1}{p_2}$$

or

$$-W = \text{const}\,(T_1 - T_2) . \tag{19.3}$$

In words: The work performed by the STIRLING engine is determined only by the temperature difference $T_1 - T_2$. This assertion can readily be confirmed in a demonstration experiment. Figure 19.3 shows a STIRLING engine in silhouette. The upper half of the cylinder is held at the constant temperature of +20 °C (293 K) by circulating water, and the lower half is alternately heated by a glycerine bath at +220 °C (493 K) and cooled in liquid air at −180 °C (93 K). In both cases, the temperature difference is the same, namely 200 K. Indeed, the machine runs in both cases at the same speed of rotation and thus performs the same amount of work in equal times (here the work serves only to overcome the friction of its bearings).

A more modern engine would use a fixed regenerator. It connects two cylinders in which the pistons move with an appropriate phase shift.

Part III

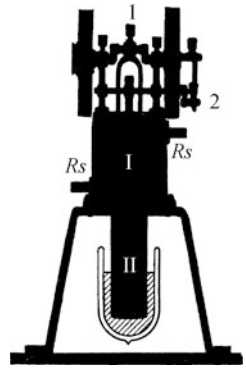


Figure 19.3 Verification of Eq. (19.3) with a STIRLING engine. The crank 2 moves the displacer drum (D) sketched in Fig. 19.2. The tube ends *Rs* serve as input and output openings for the room-temperature water which maintains the temperature of the upper half of the cylinder. Its lower half is in this case cooled by liquid air, and is thus the cooler end, denoted by *II* (**Video 19.2**).

Video 19.2: "Technical Version of a STIRLING engine"
http://tiny.cc/kigvjy
In the video, it is shown that only the temperature *difference* between the two halves of the cylinder determines the operation of the STIRLING engine. The engine shown here is about 100 years old. For a demonstration of its operation, see also **Video 19.1** (http://tiny.cc/gigvjy).

19.4 Technical Heat Engines

Technical heat engines do not operate reversibly. The most important steam engines today are steam turbines. In their construction, the dependence of the gas density on its pressure must be taken into account. The points treated in Sect. 18.7 must be considered. The *head* (height of the water level above the turbine) for water turbines corresponds to the decrease of the *specific enthalpy* of the steam in a steam turbine. In modern turbines, it can be up to 0.33 kWh/kg, corresponding to a water head of 122 km(!). Therefore, if the steam were expanded in a *single* step, the corresponding velocity would be around 1.5 km/s. For this reason, steam turbines are subdivided into several stages connected in series.

The working substance for turbines remains today in most cases water vapor; in special cases, the steam stages are preceded by a stage using a circuit of Na or biphenyl vapor. The water vapor, after being evaporated, is passed through a "superheater", i.e. it is converted to unsaturated vapor, a pure gas. Temperatures of up to about 500 °C are employed.

Apart from steam engines, *internal combustion* engines are of great technological importance. In their case, the heat energy is produced *within* the cylinder, at its upper end. The working substance is *air* with a small component (less than 21 mole percent) of gaseous combustion products. – Assume the volume of the combustion chamber above the piston to be V_1 (Fig. 19.4). After combustion, the temperature rises up to T_1. While pushing out the piston, the working substance expands adiabatically up to the cylinder volume V_2. In the process, it cools to a temperature of

$$T_2 = T_1 \left(\frac{V_1}{V_2}\right)^{\kappa - 1} \qquad (14.48)$$

κ = adiabatic exponent; for air, $(\kappa - 1) \approx 0.4$ (Table 14.1).

The remainder of the heat produced, which is not converted to work, is given up to the surrounding air along with the exhaust gases. Its

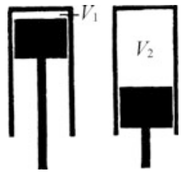


Figure 19.4 Regarding the efficiency of an internal combustion engine

temperature decreases in this process from T_2 down to the outside temperature. In order to avoid taking average values for the temperatures, we insert T_1 and T_2 into Eq. (19.2) and obtain as the largest theoretically-possible efficiency

$$\eta_{\text{ideal}} = \frac{T_1 - T_2}{T_1} = 1 - \left(\frac{V_1}{V_2}\right)^{\kappa-1} . \tag{19.4}$$

The smaller the ratio V_1/V_2 (its inverse is called the *compression ratio*), the cooler the exhaust gases and the higher the efficiency.

The required quantity of air and fuel can be introduced into a small combustion chamber only by compressing them strongly. If the piston compresses a fuel-air *mixture* (NIKOLAUS OTTO, 1876), the ratio $V_2/V_1 \approx 8$ cannot be exceeded if we wish to avoid premature ignition. This corresponds to an efficiency of $\eta_{\text{ideal}} = 57\,\%$. If only the air is compressed by the piston and the fuel is then injected into the combustion chamber (RUDOLF DIESEL, from 1893), then we can increase the compression ratio to $V_2/V_1 \approx 16$. This corresponds to $\eta_{\text{ideal}} = 67\,\%$.

OTTO and DIESEL engines have roughly the same temperature in their combustion chambers, $T_1 \approx 1900\,\text{K}$. But the DIESEL engine, with $V_2/V_1 \approx 16$, can eject its exhaust gases at a temperature T_2 which is lower than in the OTTO engine, with $V_2/V_1 \approx 8$. The practically achievable efficiencies are for the OTTO engine $\approx 30\,\%$, and for the DIESEL engine $\approx 35\,\%$.

19.5 Heat Pumps (Refrigeration Devices)

In Fig. 19.3, we showed a STIRLING engine as a readily understandable heat engine. At its upper end, heat was applied, while the cooler reservoir was below. Based on this particular experiment, we can establish a general scheme which is applicable to every heat engine (Fig. 19.5, left). It represents the ideal limiting case of complete reversibility. A working substance moves periodically between two containers *I* and *II* at different temperatures. It mediates the transfer of heat from the warmer container *I* to the cooler container *II*. The working substance takes up the heat $Q_{\text{rev}(1)}$ at the higher temperature T_1. At the lower temperature T_2, it gives up the smaller quantity of

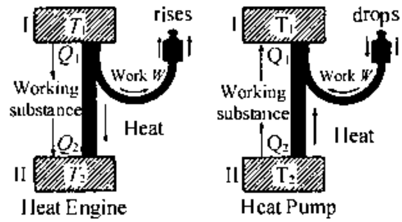


Figure 19.5 The *heat pump* (refrigeration device) as the reverse of a heat engine (we have left off the index *rev* on the heats Q.)

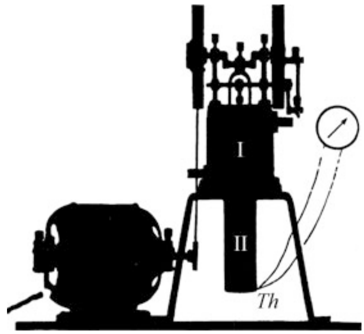


Figure 19.6 A STIRLING engine used as a heat pump (refrigeration device) *Th* is a thermocouple (**Video 19.3**).

Video 19.3:
"Heat pump/refrigerator"
http://tiny.cc/1hgvjy
The experiment is carried out with the same lecture-demonstration model as in **Video 19.1** (http://tiny.cc/gigvjy).

Part III

heat $Q_{\text{rev}(2)}$ (<0). The difference between these two heat energies is employed to perform useful work W (<0) (cf. Comment C19.1). In the scheme shown, the work is stored as the potential energy of a weight lifted to a certain height. The process comes to an end when the temperature difference has been equalized by the energy transport, i.e. when $T_1 = T_2$ is reached.

Can the temperature equalization between the heat baths *I* and *II* be reversed again, can we warm *I* at the cost of *II*? We surely can! We need only supply the quantity of work W previously performed by the machine and allow it to run backwards. Then it no longer acts as a *heat engine*, but instead as a *heat pump*. The cyclic process of the STIRLING engine then proceeds in the reverse direction, i.e. compression at a higher temperature and expansion at a lower temperature.

We first show this experimentally. In Fig. 19.6, the STIRLING engine is driven by an electric motor; in the process, the lower half of the cylinder *II* is cooled, and the upper half is heated correspondingly. After a short time, a temperature difference of $T_1 - T_2 = 10\,\text{K}$ has already been produced. By applying mechanical work, heat has been "pumped up" from *II* to *I*.

This experiment leads directly to the idealized scheme for all heat *pumps* (Fig. 19.5, right). Compare it to the scheme of all *heat engines* to the left; this needs no further explanation.

Often, heat pumps are used under the name *refrigeration devices*. As 'refrigerators', they have the task of cooling an insulated space *II*, for example the interior of a household refrigerator, relative to its surroundings *I*, for example the room air. As heat pumps in the narrower sense, they have the task of heating an insulated space *I*, for example a living room, relative to its surroundings *II*, for example the free atmosphere outside; or, as "air conditioners", of cooling a similar space relative to the outside air. Depending on their mode of use, their efficiency has to be appropriately defined. We do this again for the *the ideal limiting case* of complete reversibility; the required work is then minimal. For a *refrigeration device*[C19.7], we have

$$\eta_{\text{ideal}} = \frac{Q_{\text{rev}(2)} \text{ (Heat taken up at the low temperature) } (T_2)}{W \text{ (Work expended)}}$$
$$= \frac{Q_{\text{rev}(2)}}{-Q_{\text{rev}(1)} - Q_{\text{rev}(2)}} \,,$$

or, with

$$\frac{Q_{\text{rev}(1)}}{Q_{\text{rev}(2)}} = -\frac{T_1}{T_2} \,, \qquad (18.3)$$

$$\eta_{\text{ideal}} = \frac{T_2}{T_1 - T_2} \,. \qquad (19.5)$$

C19.7. the energy balance for the refrigeration device or heat pump is given by $Q_2 + W = -Q_1$. (Q_2 and W are input to the device, Q_1 is the heat output. Compare Comment C19.1).

For the *heat pump*, we find

$$\eta_{\text{ideal}} = \frac{-Q_{\text{rev}(1)} \text{ (Heat taken up at the high temperature) } (T_1)}{W \text{ (Work expended)}}$$
$$= \frac{-Q_{\text{rev}(1)}}{-Q_{\text{rev}(1)} - Q_{\text{rev}(2)}}$$

or, with Eq. (18.3),

$$\eta_{\text{ideal}} = \frac{T_1}{T_1 - T_2} \,. \qquad (19.6)$$

Technical details would lead us too far afield. We will have to be content with a few remarks:

1. From Eq. (19.5), we obtain the *basic rule* of refrigeration technology: In order to cool an object to a low temperature T_2, the working substance should never take up heat at a temperature *lower* than T_2. The lower T_2, the less efficiency, according to Eq. (19.5). In short: You shouldn't cool champagne with liquid air.

"In short: You shouldn't cool champagne with liquid air".

2. Gases are not suitable as working substances for refrigeration devices and heat pumps. The volume of the gas cannot be changed rapidly and isothermally, i.e. practically without temperature changes, because the heat exchange with the surroundings is too *slow*; furthermore, high pressures are required for effective cooling. For this reason, one preferentially uses substances for which under the given conditions the liquid and the gas phases exist in equilibrium, e.g. CO_2, ammonia, freon. Their volumes can then be readily changed isothermally by a *phase change*, i.e. evaporation and liquefaction.

3. A numerical example of Eq. (19.6): A residence is to be heated by using a heat pump. The heat input to the machine is to be taken from

the air outside. At an outside temperature of 0 °C, an inside temperature of 20 °C is to be maintained. Then $T_1 = 293\,\text{K}$, $T_2 = 273\,\text{K}$. From Eq. (19.6) *for the ideal limiting case* of complete reversibility, we find

$$\eta_{\text{ideal}} = \frac{T_1}{T_1 - T_2} = \frac{293}{293 - 273} = \frac{293}{20} = 14.7\,.$$

Today, our living space is often heated by electric radiators. That is very convenient, but inefficient. A more physically acceptable method would be the following: We should use the electrical energy to *pump in* the heat from outside the house. In our example, about 7 % of the otherwise necessary electric power would suffice! This means that with an energy consumption of one kilowatt hour, we could bring around 14 kilowatt hours of heat into our rooms! The increased use of heat pumps would be highly desirable to conserve our energy resources.

19.6 The Thermodynamic Definition of Temperature

Equation (19.2) contains no material constants. We can therefore define a measurement procedure for the temperature with its aid which is *independent of all such materials properties*. We need only fix one or the other of the two temperatures T_1 or T_2 to an arbitrarily-chosen numerical value, e.g. T_2. Then the other temperature is uniquely determined by the thermal efficiency of a completely reversible machine. In principle, we have only to measure the efficiency of such a machine in order to obtain the unknown temperature[2] T_1. This was first recognized by WILLIAM THOMSON, later Lord KELVIN (1824–1907). For that reason, the absolute temperature scale (that is, the scale which contains no negative values) is named after KELVIN (Sect. 14.6). Practically, it is identical to the temperature scale as defined by good-quality gas thermometers.

19.7 Pneumatic Motors. Free and Bound Energy

Thus far, we have treated the conversion of heat into work based on making use of a *temperature difference* and carried out by heat engines. It is however also possible to convert heat into work *without*

---

[2] Instead of fixing *one* of the two temperatures at an arbitrarily chosen value, we could choose such a value for the *difference* of two temperatures, i.e. $(T_1 - T_2)$ (cf. Sect. 14.6).

a temperature difference, that is in an *isothermal* process. A clear example of this is a pneumatic motor which is operated isothermally.

We repeat from Sect. 14.11: The isothermally-operating pneumatic motor is a machine which converts the heat that it takes up from its surroundings into work. Its efficiency is ideally 100 %. The conversion is accomplished by the expansion of compressed air.

Now we add a new point: The expansion increases the entropy of the compressed air. Its entropy increase is:

$$\Delta S = \frac{Q_{\text{rev}}}{T} \quad \text{and thus}^{\text{C19.8}} \quad Q_{\text{rev}} = T \cdot \Delta S\,. \tag{18.2}$$

C19.8. This equation in its general form is: $Q_{\text{rev}} = \int T\,\mathrm{d}S$.

This entropy increase characterizes the permanent change which the working substance (compressed air) undergoes when it performs work isothermally. We will show the effects of this permanent change in the general case, i.e. not limited to the expansion of compressed air.

The First Law:

$$\left.\begin{array}{ccccc} Q & = & -W & + & \Delta U \\ \text{heat input} & & \begin{array}{c}\text{energy output}\\ \text{as external work}\end{array} & & \begin{array}{c}\text{increase of}\\ \text{the internal}\\ \text{energy}\end{array}\end{array}\right\} \tag{14.6}$$

leaves it completely undecided as to how the thermal energy input is to be divided between the two terms on the right. This is determined only by the Second Law.

Inserting Eq. (18.2) into Eq. (14.6), with

$$Q = Q_{\text{rev}}\,,$$

we obtain

$$W = -T \cdot \Delta S + \Delta U \tag{19.7}$$

or, for an isothermal process, i.e. a process occurring at constant temperature,

$$W_{\text{isoth}} = \Delta(U - T \cdot S)\,. \tag{19.8}$$

The brackets contain only quantities of state. Therefore, their content is also a quantity of state (a state function)[3]. It is called the *free energy F*; thus,

$$F = U - T \cdot S\,. \tag{19.9}$$

The free energy is smaller than the internal energy; their difference

$$U - F = T \cdot S \tag{19.10}$$

[3] Only *differences*, such as ΔU, ΔS, ΔF, etc. are measurable. If one determines numerical values for U, S, F, etc., they always hold only for a particular reference temperature, which must be quoted, as for example in Table 18.1.

is also referred to as the *bound energy*.

Then

$$W_{\text{isoth}} = \Delta F \tag{19.11}$$

which thus represents the maximum work that can be obtained from an isothermal and reversible process[4].

The bound energy $T \cdot S$ is not wasted, but rather it is determined in such a way that it is not available for the performance of additional work[5].

For many applications, one requires the influence of the temperature on the free energy F or on the maximum work W_{isoth} which can be performed by an isothermal process. We can obtain this influence from Eq. (19.2) by applying it to a very small temperature difference $T_1 - T_2 = \mathrm{d}T$. Then we obtain

$$\frac{-\mathrm{d}W_{\text{isoth}}}{Q_{\text{rev}}} = \frac{\mathrm{d}T}{T}\,. \tag{19.12}$$

From this, with Eq. (18.2), it follows that

$$\frac{-\mathrm{d}W_{\text{isoth}}}{\mathrm{d}T} = \frac{Q_{\text{rev}}}{T} = \Delta S\,. \tag{19.13}$$

Inserting this into Eq. (19.8) or (19.7), we obtain the equation (named for HELMHOLTZ) for the maximum work which can be obtained from an isothermal process:

$$W_{\text{isoth}} = T\,\frac{\mathrm{d}W_{\text{isoth}}}{\mathrm{d}T} + \Delta U\,. \tag{19.14}$$

It plays an important role in physical chemistry.

19.8 Examples of Applications of the Free Energy

1. A *compressed-air cylinder as an energy storage medium*. In the particular case of compressed air, the situation is especially simple, since the internal energy U of the ideal gas remains constant on

[4] Maximum work, because in Eq. (18.2), only reversible processes were presupposed.
[5] The internal energy of a material can be compared with the assets of a company; the free energy with its "liquid" assets, and the bound energy with its "non-liquid" assets (e.g. plants and property).

Part III

isothermal expansion and therefore, $\Delta U = 0$. We thus obtain from Eqns. (19.11) and (19.8)

$$W_{\text{isoth}} = \Delta F = -T \cdot \Delta S\,. \tag{19.15}$$

Here, we have from Eq. (18.15)

$$\Delta S = nR \ln \frac{p_1}{p_2}\,. \tag{19.16}$$

Example

A steel cylinder of mass 64 kg and volume 42 liters at a pressure of $p_1 = 190 \cdot 10^5$ Pa ($= 190$ bar) contains 330 moles of compressed air at room temperature. On isothermal expansion to $p_2 = 1$ bar, its free energy is reduced by

$$\Delta F = -293\,\text{K} \cdot 330\,\text{mol} \cdot 8.31\,\frac{\text{W s}}{\text{mol K}} \cdot \ln \frac{190}{1}$$

(see also **Exercise 15.1**). The logarithmic factor has the value $\ln 190 = 5.25$. Then

$$\Delta F = -4.2 \cdot 10^6\,\text{W s} \approx -1.2\,\text{kWh}\,.$$

An electrical storage battery with about the same mass (70 kg) reduces its free energy on complete discharge by around 2 kWh[C19.9].

C19.9. For comparison, we mention that the free energy of fuels (butter, coal) amounts to about 10 kWh/kg.

2. *Entropy elasticity or rubber elasticity*. For the majority of solid materials, e.g. the metals, the elastic forces which accompany deformation arise from the change in their internal energies. – The situation is quite different for ideal gases. The setup sketched schematically in the upper part of Fig. 19.7 permits the measurement of the elastic force F_{el} which is produced by a column of air confined in a container. If the sliding carriage is moved isothermally to the left, the length of the air column will be changed by $\Delta l < 0$. The force F_{el} remains practically constant. The air is compressed and, from Eq. (19.15), work is performed on it:

$$-F_{\text{el}}\Delta l = \Delta F = -T \cdot \Delta S$$

For the elastic force, we thus obtain

$$F_{\text{el}} = \frac{\Delta S}{\Delta l} \cdot T\,. \tag{19.17}$$

The elastic force F_{el} is therefore proportional to the temperature. It arises simply because the entropy of the air is reduced on isothermal *compression*. As a result, we speak of its *entropy elasticity*.

The entropy of the air, according to Sect. 18.5, Point 4., is given by

$$\frac{\Delta S}{n} = C_{\text{p}} \ln \frac{T_2}{T_1} - R \ln \frac{p_2}{p_1}\,. \tag{18.15}$$

If the air column at the top of Fig. 19.7 is shortened *adiabatically* by Δl, the entropy of the gas remains constant, but its two contributions

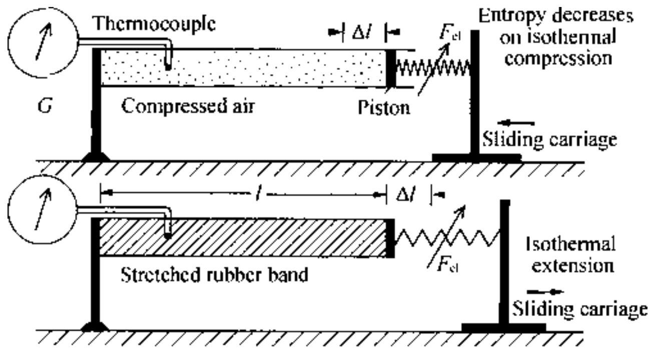


Figure 19.7 The demonstration of entropy elasticity with a compressed column of air and a stretched rubber band (G is a galvanometer for temperature measurements; the force meter is shown only schematically. At the *top*, F_{el} acts to the right, against the motion of the carriage; at the *bottom*, the elastic force acts to the left).

in Eq. (18.15) both change: On adiabatic compression from p_1 to p_2, the first term increases at the cost of the second. The temperature of the gas is increased by adiabatic compression.

Among *solid* materials, one can also observe entropy elasticity, e.g. in rubber and the rubber-like polymers (Fig. 19.7, bottom). Equation (19.17) is again applicable. We find that F_{el} at a constant length l of a rubber band is usually proportional to T to a good approximation. When the band is adiabatically stretched by Δl, it will become warmer. From these two facts, it follows that the internal energy of rubber-like materials is nearly independent of their extension and volume. Therefore, here also the elastic force is determined only by the *entropy* of the band. But now the entropy *decreases* on isothermal *extension*. This can be readily understood:

Rubber-like materials belong to the group of high-polymer substances. In these materials, identical and relatively small molecules (monomers) are joined like the links of a chain into long threadlike molecules. In a disordered pile, they form a matted wad as the most probable arrangement (high entropy). When the band is stretched, the threadlike molecules are pulled apart and become parallel to some extent; their arrangement becomes more ordered and thus less probable (lower entropy). As a model, we can simulate the polymer chains by macroscopic chains of ca. 10 cm length, whose links are magnetic. If we let them lie on a glass plate and vibrate the plate to simulate the thermal motions, they will form a disordered, matted wad (with high entropy), like the polymeric molecular chains.

When a rubber band is adiabatically relaxed, so that its entropy remains constant, it will cool.

3. The *dependence of the melting temperature on the pressure*. Suppose that a substance at constant temperature and pressure changes

its volume by ΔV on melting. Differentiation of the molar work of expansion $W = -p\Delta V_m$ with respect to T yields

$$\frac{dW}{dT} = -\frac{dp}{dT}\Delta V_m. \tag{19.18}$$

Inserting this equation into Eq. (19.13), we obtain from the Second Law the so-called CLAUSIUS-CLAPEYRON equation:

$$\frac{dT}{dp} = \frac{T}{L_f}\Delta V_m. \tag{19.19}$$

It shows the pressure dependence of the melting temperature. (L_f is the molar latent heat of melting[C19.10]; cf. Comment C13.10). – For some substances, the melting temperature increases with increasing pressure. Examples: Wax and CO_2 (Fig. 15.10). In their case, $dT/dp > 0$. Then, according to Eq. (19.19), $\Delta V > 0$ must also hold, i.e. these substances must *expand* on melting.

C19.10. A similar equation applies to the vapor pressure. The derivative dT/dp then holds along the vapor pressure curve, ΔV_m refers to the difference in molar volumes of the gas (vapor) and the liquid, and the constant L_v is the molar latent heat of vaporization.

For other substances, the melting temperature decreases with increasing pressure. Example: Water (Fig. 15.11). In this case, $dT/dp < 0$. Then, from Eq. (19.19), $\Delta V < 0$ must hold, i.e. these substances must *contract* on melting. Both kinds of behavior are observed experimentally.

Part III

19.9 The Human Body as an Isothermal Engine

Energy is input to the body by the oxidation of our foodstuffs. For example, we find for

$$\left.\begin{array}{ll} \text{Butter}\ldots & 9.1 \\ \text{Oatmeal}\ldots & 4.2 \\ \text{Rice}\ldots & 3.9 \\ \text{Bread}\ldots & 2.3 \\ \text{Potatoes}\ldots & 0.9 \end{array}\right\} \frac{\text{kWh}}{\text{kg}}$$

At rest, maintaining the life functions of an adult human requires a power of about 80 W. That is, a human body requires an energy input of around 2 kWh per day. When the person is performing mechanical work, the energy input must be increased up to 3 to 4 kWh per day; for heavy workers, this can be as much as 6 kWh per day. On the average, a human requires an energy input of only about 1300 kWh per year (commercial value around 3 $!)[C19.11].

C19.11. For electrical energy, this corresponds today in Germany to a consumer cost of around 400 Euros or 440 $. In the U.S., the average consumer cost would be around 130 Euros or 145 $.

The thermal efficiency of our muscles is in general around 20 %; through training, it can be increased up to ca. 37 %. Therefore, the muscles cannot possibly be working as heat engines. At an ambient temperature of $T_2 = 293\,\mathrm{K}$ (20 °C), from Eq. (19.2), the internal body temperature would have to be $T_1 = 465\,\mathrm{K}$ (192 °C)! So only an *isothermal* production of muscular work is a possible explanation. In that case, around 60 to 80 % of the chemical energy of our food is converted to heat! Work, for example climbing a mountain, generates heat. (In these figures, the basal metabolism of the body, i.e. its requirement of 2 kWh per day at complete rest, is not included.)

In a careful observation, one has to consider two different processes in the work performed by muscles. During one of them, force is generated; this process is similar to discharging a storage battery: Stored chemical energy is converted to mechanical work. The efficiency of this process can reach 90 %. Then a second process follows; figuratively speaking, it is like recharging the battery. This second process, in contrast to the first, can proceed only in the presence of O_2. It requires an oxidation step, has a low efficiency and produces quite a lot of heat.
Continuous athletic activity, isotonically or in motion, requires a power input of about 1.4 kW (corresponding to an oxygen consumption of 4 liters per minute). Around 1/5 of this power, i.e. about 300 W, is available for carrying out mechanical work (in contrast to isotonic work). For brief high-level activity, the 'muscular storage battery' has an energy reserve of the order of 100 kilowatt-*seconds*. It can be replaced in about a half hour after complete exhaustion, accompanied by an uptake of 15 liters of O_2. A small fraction, which decreases strongly with increasing demand, can be converted to mechanical work. By using this energy reserve, a human can expend a mechanical power of several kilowatts for a few seconds (Sect. 5.2).

The work performed by our muscles is by no means reversible. Their work is just as irreversible as that of technical heat engines. Work performed reversibly is much too ponderous and slow. Work performed reversibly is an ideal, but even this ideal is, like some others, not necessarily worth striving for.

"Work performed reversibly is an ideal, but even this ideal is, like some others, not necessarily worth striving for".

Exercise

19.1 At an outside temperature of −5 °C, a heating boiler in a house is to be maintained at a temperature of 40 °C by a heat pump. How much electrical energy W_{el} will be consumed in the ideal case of completely reversible operation of the pump, if the required heat input to the boiler amounts to 1 kJ? (Sect. 19.5)

Electronic supplementary material The online version of this chapter (doi: 10.1007/978-3-319-40046-4_19) contains supplementary material, which is available to authorized users.

Table of Physical Constants

Some important physical constants		(CODATA values from Dec. 2014)
Gravitational constant	γ	$= 6.667_4 \cdot 10^{-11}\ \mathrm{N\,m^2/kg^2}$
Electric field constant	ε_0	$= 8.854\ 188 \cdot 10^{-12}\ \mathrm{A\,s/V\,m}$
Magnetic field constant	μ_0	$= 12.566\ 370 \cdot 10^{-7}\ \mathrm{V\,s/A\,m}$
Velocity of light in vacuum	c	$= (\varepsilon_0\mu_0)^{-1/2} = 2.997\,925 \cdot 10^8\ \mathrm{m/s}$
Wave resistance of vacuum	Z	$= (\mu_0/\varepsilon_0)^{1/2} = 376.73\ \mathrm{Ohm}$
Relative atomic mass of the proton	$(A)_p$	$= 1.007\,277\ \mathrm{u}$
Relative atomic mass of the neutron	$(A)_n$	$= 1.008\,665\ \mathrm{u}$
Proton mass	m_p	$= 1.672\,621 \cdot 10^{-27}\ \mathrm{kg}$
Rest energy of the proton	$(W_p)_0$	$= 9.382\,720 \cdot 10^8\ \mathrm{eV}$
Rest mass of the electron	m_0	$= 9.101\,383 \cdot 10^{-31}\ \mathrm{kg}$
Rest energy of the electron	$(W_e)_0$	$= 5.109\,99 \cdot 10^5\ \mathrm{eV}$
Ratio of proton mass/electron mass	m_p/m_0	$= 1836.152$
Elementary electric charge	e	$= 1.602\,177 \cdot 10^{-19}\ \mathrm{A\,s}$
Specific charge of the electron	e/m_0	$= 1.760\,366 \cdot 10^{11}\ \mathrm{A\,s/kg}$
Boltzmann's constant	k	$= 1.380\,648 \cdot 10^{-23}\ \mathrm{W\,s/K}$ $= 8.617\ 325 \cdot 10^{-5}\ \mathrm{eV/K}$
Planck's constant	h	$= 6.626\,070 \cdot 10^{-34}\ \mathrm{W\,s^2}$ $= 4.135\,667 \cdot 10^{-15}\ \mathrm{eV\,s}$
Smallest orbital radius of the H atom	a_H	$= \varepsilon_0 h^2/\pi m_0 e^2 = 0.529\,177 \cdot 10^{-10}\ \mathrm{m}$
Bohr magneton	μ_B	$= \mu_0 he/4\pi m_0$ $= 1.165\,407 \cdot 10^{-29}\ \mathrm{V\,s\,m}$
Classical electron radius	r_{el}	$= \mu_0 e^2/4\pi m_0 = 2.820\,419 \cdot 10^{-15}\ \mathrm{m}$
Rydberg frequency	R_y	$= e^4 m_0/8\varepsilon_0^2 h^3 = 3.289\,842 \cdot 10^{15}\ \mathrm{s^{-1}}$
Rydberg constant	R_y^*	$= e^4 m_0/8\varepsilon_0^2 h^3 c = 10\,973\,731.6\ \mathrm{m^{-1}}$
Compton wavelength	λ_C	$= h/m_0 c = 2.426\,310 \cdot 10^{-12}\ \mathrm{m}$
Sommerfeld's fine structure constant	α	$= e^2/2\varepsilon_0 hc = 1/137.036$
$= \dfrac{\text{Electron's velocity } u \text{ in the smallest H orbit}}{\text{Velocity of light } c}$		

Solutions to the Exercises

I. Mechanics

1.1. After 100 years, each day will be $1.5 \cdot 10^{-3}$ s longer than today; then after one year, it will be $1.5 \cdot 10^{-5}$ s longer, and tomorrow, it will be $1.5 \cdot 10^{-5}/365.25$ s, i.e. $4.106 \cdot 10^{-8}$ s longer. Then in 100 years, the clocks will be ahead by $4.106 \cdot 10^{-8}\,\mathrm{s} \cdot (1 + 2 + 3 + \ldots + 36\,525) = 27.4$ s. (The sum can be solved analytically and its solution can be found in many mathematics textbooks).

1.2. $T = 0.05\,\mathrm{s}$

2.1. Downriver, at an angle of 53.13° to the bank.

2.2. $s = 34.3\,\mathrm{m}$

2.3. $t = 17.35\,\mathrm{s}$; $h = 293\,\mathrm{m}$

2.4. $a = 2.475\,\mathrm{m/s^2}$; $F = 49.5\,\mathrm{N}$

2.5. a) $t = 1\,\mathrm{s}$; b) $t = (\pi/4)\,\mathrm{s}$

2.6. $a_\mathrm{r} = 2.10 \cdot 10^{-2}\,\mathrm{m/s^2}$

3.1. $F = 63.1\,\mathrm{N}$; $F_\mathrm{tot} = 42.75\,\mathrm{N}$ to the east ($\alpha = 90°$)

3.2. a) $F = 0.431\, F_\mathrm{G}/(0.959 \sin\Theta + 0.413 \cos\Theta)$; b) $\Theta = 66.7°$

3.3. $v = 86.9\,\mathrm{cm/s}$; $E_\mathrm{kin} = 0.196\,\mathrm{Nm}$; $F_\mathrm{F} = 0.189\,\mathrm{N}$

4.1. $u_0 = 0.628\,\mathrm{m/s}$; $a_0 = 3.95\,\mathrm{m/s^2}$

4.2. $T_\mathrm{M} = 3.16\,\mathrm{s}$; $m_\mathrm{M} = g r_\mathrm{M}^2/(10G)$ (G = gravitational constant)

4.3. The new angular velocity is $\omega_1 = l^2\omega/(l - l_1)^2$

4.4. $z = 5086\,\mathrm{km}$

4.5. With a tangential velocity of $1.5 \cdot 10^{-5}$ m/s, the model earth could circle the sun on a stable orbit. Its "year" would be exactly as long as a real year!

5.1. $W = A\varrho g(H-h)^2/4$

5.2. $W = 0.136\,\text{kWh}$; Power $P = 8.16\,\text{kW}$

5.3. $h_2 = h_1 + (\sqrt{2g(h-h_1)} - \Delta p/m)^2/2g;\ v = \sqrt{2gh_2}$

5.4. a) $v = 10\,\text{m/s}$; b) $v = 5\,\text{m/s}$

5.5. a) From the conservation of energy and momentum, we find: $v_2 = 2(m/M)v_1/(1+m/M) \approx 2(m/M)v_1$ and $\Delta v = 2v_1/(1+m/M)$. b) After the first impact, the balls swing outwards until their kinetic energy is completely converted to potential energy. Then they swing back again, so that the first impact is played out again in reverse. c) When the large ball collides with the small one, the latter flies away with the velocity $\approx 2v_2$ (this can be seen especially clearly in the frame of reference of the large ball). The large ball loses a velocity of $2v_2(1+M/m)$, that is less than 1 %. This result is verified by comparison of the two amplitudes.

5.6. $h = m^2v^2/(2g(M+m)^2)$

5.7. $v = 2.97\,\text{m/s}$; 13.36 J still remains from the original 1350 J.

5.8. This follows from the conservation laws for momentum and energy, together with the Pythagorean Theorem.

5.9. $M_\text{p} = 1.718M_0$ or 63 % of the total initial mass.

5.10. $\mathrm{d}M/\mathrm{d}t = M_\text{t}(a_\text{o}+g)/u_\text{r}$ (g = acceleration of gravity).

6.1. a) $m_2 = 1.875\,\text{kg}$; b) 7/18 of the length of the beam, as measured from m_2; $F = 441\,\text{N}$.

6.2. One can either use the equation of motion to compute the acceleration of the cylinder axles, $\mathrm{d}v/\mathrm{d}t = (2/3)g\sin\alpha$, or else use the law of energy conservation to find the final velocity, $v = \sqrt{4gh/3}$. Both expressions contain neither the mass (or the density) nor the radii of the cylinders. The two cylinders thus arrive at the bottom simultaneously.

6.3. $W = 976\,\text{J}$

6.4. $l = 2a/3$

6.5. a) At one of the two points at a distance of $a/(2\sqrt{3})$ from the center of gravity; b) at one of the two points at a distance of $a/6$ from the center of gravity.

6.6. $\Theta_0 = 0.05\,\text{kg}\,\text{m}^2$; $\Theta_\text{S} = 0.03\,\text{kg}\,\text{m}^2$

6.7. a) $\Theta = 0.00579\,\text{kg}\,\text{m}^2$ (make use of Eq. (6.12)); b) $d = 0.28\,\text{mm}$ (here, the torsion coefficient changes by $2mgd$).

6.8. $\varphi = 2\pi mr^2/(\Theta + mr^2)$

6.9. Under the assumption that the end of the pencil does not slide immediately off the edge of the table, the pencil will begin to rotate around its point with the angular momentum $\Theta\omega$ (where

$\Theta = (1/3)ml^2$), which is equal to the orbital angular momentum $mvl/2$ at this point. It follows that $\omega = (3/2)\sqrt{2gh}/l$, or $\nu = 5.3$ rotations per second. This result could be checked simply by using a video camera.

6.10. See Eq. (6.16)

6.11. The gyroscope is also rotating in the moving frame of reference, but without precession. For each of its volume elements, the rotation can be decomposed into a vertical velocity component parallel to the vector of the angular velocity ω_P and a horizontal component. The latter leads to a CORIOLIS force on the volume element, which causes a torque to act on the axle of the gyroscope. The sum of all these torques leads to an overall torque which is parallel to $\boldsymbol{M}$ but opposite in direction. Its magnitude would have to be found from a computation; here, we content ourselves with the statement that it would have the same magnitude as $\boldsymbol{M}$. In the rotating frame of reference, the sum of all the torques along the axle of the gyroscope is thus zero.

7.1. a) $F_C = 2mu\omega = 50\,\text{N}$. b) While the bullet is moving within the barrel of the pistol, it moves along a circular orbit with a continually increasing radius. The distance s along this orbit is given by $s = R\omega t$ and $R = ut$. It then follows that $s = u\omega t^2$, and thus the acceleration is $\mathrm{d}^2s/\mathrm{d}t^2 = 2u\omega$ and the force is $F_H = 2mu\omega = 50\,\text{N}$, so that it has the same magnitude as F_C, but the opposite direction.

7.2. For the observer in the rotating frame of reference of the swivel chair, the pendulum is moving on a circular orbit with the velocity $u = \omega R$. This requires that a radial force $m\omega^2 R$ directed towards the center of rotation be acting. This force is composed of the centrifugal force $F_Z = m\omega^2 R$, which is directed outwards, and the CORIOLIS force $F_C = 2mu\omega = 2m\omega^2 R$ which is directed towards the center.

7.3. $a_C = 2\omega v \sin\varphi = 1.9 \cdot 10^{-3}\,\text{m/s}^2 \approx 10^{-4}g$ (g = acceleration of gravity). The CORIOLIS force is thus negligible compared to the inertial forces which are due for example to the imperfect flatness and smoothness of the rails.

7.4. $T = 2\pi\sqrt{R/g}$, i.e. the SCHULER period!

7.5. $s = 1.5 \cdot 10^{-9}\,\text{m} \approx 1\,\text{nm}$

7.6. During seven complete swings, that is after 48 s, the image of the pendulum wire has shifted by seven wire diameters, or by 2.8 mm on a circle of radius $A = 1\,\text{m}$. From this, at the latitude of Göttingen, we find an angular velocity of $\omega_G = 5.83 \cdot 10^{-5}\,\text{s}^{-1}$. Dividing by the factor $\sin 51.5° = 0.78$ then yields $\omega_E = 7.5 \cdot 10^{-5}\,\text{s}^{-1}$ (the precise value is $7.3 \cdot 10^{-5}\,\text{s}^{-1}$).

8.1. $E = 1.25 \cdot 10^5\,\text{N/mm}^2$

8.2. From L and H, with a certain amount of trigonometry, we find for the radius of curvature $r = 2.56\,\text{m}$. s is measured to be 0.26 m. Then we can compute the torque which deforms the rod: $M = 1\,\text{kg} \cdot g \cdot 0.26\,\text{m} = 2.6\,\text{N m}$ (note that the torque which is due to the second kg weight serves only to keep the rod from being accelerated). The geometrical moment follows from Eq. (8.23): $J = 64 \cdot 10^{-12}\,\text{m}^4$. Then, finally, we obtain $E = 10.3 \cdot 10^4\,\text{N/mm}^2$, in good agreement with the value given in Table 8.1.

8.3. The light pointer (broadened by diffraction) is displaced by nearly d_s. From this, we find $\alpha = 7 \cdot 10^{-4}$. With the geometrical moment from Eq. (8.26), we obtain $G = 8.0 \cdot 10^4\,\mathrm{N/mm^2}$, in agreement with the value given in Table 8.1.

8.4. $s = 11.5\,\mathrm{m}$

9.1. $F_2 = 34.5\,\mathrm{N}$

9.2. $A_2 = 5A_1$

9.3. $F = ((4/3)\pi d^3/8)g\varrho_{\text{water}}$ (g = acceleration of gravity)

9.4. $F = 12\,700\,\mathrm{N}$

9.5. $F = 10^4\,\mathrm{N}$

9.6. $m = 7.35\,\mathrm{g}$ (weight 0.072 N)

9.7. The heat of evaporation (vaporization) for one molecule is $7.6 \cdot 10^{-20}\,\mathrm{W\,s}$. In each cm^2 of water surface, there are $1.03 \cdot 10^{15}$ molecules, as can be found from the number density of the water molecules. Then from Table 9.1, we find the surface work for one molecule to be $7.2 \cdot 10^{-21}\,\mathrm{W\,s}$. Around 10 % of the heat of vaporization is thus required to bring the molecules to the surface where they can evaporate.

9.8. $\tau = 4.8\,\mathrm{s}$ (in order to keep the effort required for counting to a minimum, it was begun after 6 s following the end of the pouring).

9.9. The constants of BOYLE's law are, at 0 °C: $0.8\,\mathrm{m^3\,bar/kg}$; and at 200 °C: $1.4\,\mathrm{m^3\,bar/kg}$; $p_o - p_c = 1.6 \cdot 10^2\,\mathrm{Pa}$ (= 1.6 mbar)

9.10. $(p_2 - p_1)A = gAh(\varrho_{\text{air}} - \varrho_{\text{hydrogen}})$; this is the same as for the buoyant force.

9.11. The pressure in the methane is greater than in the surrounding air, by $\Delta p = 0.57\,\mathrm{Pa}$ (this corresponds to the gravity pressure (Sect. 9.4) of a column of water 0.057 mm "high"!).

10.1. $\eta = 2gR_1^2R_2^2(\varrho_2 - \varrho_1)/9(R_1^2u_2 - R_2^2u_1)$; $\quad \varrho = (R_1^2u_2\varrho_1 - R_2^2u_1\varrho_2)/(R_1^2u_2 - R_2^2u_1)$

10.2. From energy conservation, Eq. (10.12), and neglecting the viscosity of the water, we find $r = (1/\sqrt{\pi\sqrt{2gh}})\sqrt{\mathrm{cm^3/s}}$, as long as $R \gg r$, so that we can consider the water in the cylinder to be practically at rest.

II. Vibrations and Waves

11.1. These beats are produced by the harmonics $2\nu_2$ and $3\nu_1$.

11.2. a) $d_1 = 0.326\,\mathrm{m}$; b) $d_2 = 0.163\,\mathrm{m}$ (the reduction of the resonance frequency at a fixed length is used for example in constructing organ pipes; see E. Skudrzyk, "*The Foundations of Acoustics*" Springer (Heidelberg, New York) 1971).

11.3. $\lambda = 2.4\,\text{cm}$; $\nu = 14.17\,\text{kHz}$

11.4. $c_1 = 3840\,\text{m/s}$

11.5. $\delta^{-1} = \tau = 50\,\text{s}$; $\nu_e = 0.42\,\text{Hz}$; $\Lambda = 4.8 \cdot 10^{-2}$; $K = 1.05$; $H = 6.4 \cdot 10^{-3}\,\text{Hz}$

11.6. For the symmetrical normal mode oscillation, we obtain the frequency $\nu_s = 0.900\,\text{Hz}$, and for die antisymmetrical normal mode, $\nu_{as} = 0.997\,\text{Hz}$. For the beat frequency, $\nu_B = 0.100\,\text{Hz}$. The difference $\Delta\nu = \nu_s - \nu_{as} = 0.097\,\text{Hz}$ is found, in agreement with the beat frequency determined experimentally.

11.7. a) $\nu_0 = \sqrt{2D/m}/2\pi$, $\nu_1 = \sqrt{D/m}/2\pi$, $\nu_2 = \sqrt{3D/m}/2\pi$; b) $\nu_0 = \sqrt{(D + D')/m}/2\pi$, $\nu_1 = \sqrt{D/m}/2\pi$, $\nu_2 = \sqrt{(D + 2D')/m}/2\pi$. Note that in both cases, one of the eigenfrequencies is always larger and the other is always smaller than the frequency of a single pendulum as defined in the problem. The difference between the two eigenfrequencies becomes smaller as the coupling becomes weaker.

12.1. $\nu' = 531\,\text{Hz}$

12.2. Maxima appear along the axis of symmetry and at a spacing of $\pm m y \lambda / a$ from it, where $m = 1, 2, 3, \ldots$.

12.3. On the diagonals of the square, since in these directions, the openings are at the greatest distances from each other.

12.4. $D = \lambda a / b$

12.5. $c_1 = 5000\,\text{m/s}$

12.6. $d = 108\,\text{m}$

12.7. $u = 60\,\text{km/h}$

12.8. $\sin\alpha_{max} = 1/n = 0.78, \quad \alpha_{max} = 51°$

12.9. $m = 1$: $\gamma = 9.9°$ (grazing angle), $\beta = 80.1°$; $m = 2$: $\gamma = 20°$, $\beta = 70°$; $m = 3$: $\gamma = 31°$, $\beta = 59°$; $m = 4$: $\gamma = 43.4°$, $\beta = 46.4°$.

12.10. $c^* = c(1 + \lambda(\mathrm{d}n/\mathrm{d}\lambda)/n)$

12.11. $\lambda = 4.25\,\text{m}$ to $0.425\,\text{m}$

III. Thermodynamics

13.1. The clock advances too slowly by 10^{-4}, thus losing about one minute per week.

14.1. Applying Eq. (14.19), we find $p = 166 \cdot 10^5\,\text{Pa}$ ($= 166\,\text{bar}$).

14.2. Addition of the partial pressures yields $p_{tot} = 9.4 \cdot 10^5\,\text{Pa}$.

14.3. With V_1 = initial volume, V_2 = final volume and the temperatures $T_1 = 291\,\mathrm{K}$ (room temperature) and $T_2 = 773\,\mathrm{K}$, we obtain from Eq. (14.48): $V_1/V_2 = 15.1$. The volume must therefore be reduced by a factor of about 1/15.

15.1. a) $V_m = 1.217 \cdot 10^{-4}\,\mathrm{m^3/mol} = 3.5 V_{m,solid}$, i.e. the density of the gas is 28 % of the density in the solid phase. b) From the VAN-DER-WAALS equation, with this value of V_m, we calculate a value of RT which is only around 1 % smaller than the true value. An exact solution of the VAN-DER-WAALS equation yields $V_m = 1.23 \cdot 10^{-4}\,\mathrm{m^3/mol}$, that is a value which is about 1 % larger. The ideal-gas equation thus holds rather precisely for nitrogen, even at a packing density at which the intermolecular distance is only 50 % larger than in the solid state.

18.1. Make use of Eq. (18.15) and take the standard conditions, $T_1 = 273\,\mathrm{K}$ and $p_1 = 1013\,\mathrm{hPa}$. Then you obtain the molar entropies as shown in Fig. 18.8.

19.1. $W_{el} = 0.143\,\mathrm{kJ}$

Index

A

Acceleration, 20, 35
  definition, 21
  measurement, 23
  of a sinusoidal oscillation, 50
Acceleration of gravity = Apparent gravity, 25
Acceleration of gravity = Free-fall acceleration, 35
Acoustic replica method, 323
Acoustic wave resistance, 350
actio = *reactio*, 32, 59, 76, 141
Adiabatic change of state, 400, 489
Adiabatic curve, 401
Adiabatic expansion, 487
Adiabatic exponent, 390, 402
  measurement, 405
Air friction, 83
Air pressure, 199
Air resistance, 26, 83
Aircraft
  lift, 241
  propulsion, 86
  thrust, 244
Airfoil, 238
Amount of substance, 29, 369
Anemometer, 236
Angle measurement, 9
Angle of incidence, 302
Angle, unit of, 10
Angular acceleration, 97
Angular momentum, 97, 104, 115, 117
  as vector, 104
Angular momentum conservation, 105
Angular velocity, 22, 27, 96, 97, 277
  as a vector, 96
Apparent gravity, 25
Apparent viscosity, 217
ARCHIMEDES' principle, 184
Area law, 55
Atmosphere (atm), 178
Atmosphere = "ocean of air", 198, 445
Atmosphere of the earth, 198
Auto control, 283
Auto-controlling
  of a tuning fork, 249
AVOGADRO constant, 370
Avoidance flow, 226, 238
Axis
  free, 108
Axis of incidence, 302
Axis of rotation, 91

B

Balance, 43, 104
Ballistic curve, 65
Ballistic pendulum, 81
Bar exercises, 108
Bar, unit of pressure, 178
Barometric equation, 445
Barometric pressure formula, 203, 204, 446, 449
Base quantities, 29
Base units, 29
Basilar membrane, 363
Bats, 325
Beam balance, 43, 104
Beam of sound, 303
Beat frequency, 253
Beats, 14, 253, 284, 342
BEHN's tube, 206
Bending vibrations, 270
BERNOULLI's equation, 223, 327
BEUERMANN, G., 254
Bicycle, 84
Bicycle riding, 119
Bicycle, riding 'no hands', 119
Bimetallic strip, 371
Blood circulation, 216
Blood reservoir, 216
Boiling, 394, 425
Boiling temperature, 421
BOLTZMANN distribution, 449
BOLTZMANN, LUDWIG, 483
BOLTZMANN's constant, 393, 483
  determination by J. PERRIN, 445
Boom, supersonic, 354
Boomerang, 121
Bound energy, 506
Boundary energy, 188
Boundary layer, 211, 212
  in the RUBENS' flame tube, 267
  thickness, 213
Bow wave, 307
BOYLE's law, 194
BRAGG angle, 321, 335
Breakers, 192
Breathing sphere, 351
BROWN, R., 175
Brownian motion, 175, 445, 479
  in gases, 193, 204
Bulk modulus, 156
Buoyancy, 86, 184
  in gases, 205
  in liquids, 183

C

Calorimeter, 389
Capillary depression, 190
Capillary elevation, 190
Carbon dioxide
  isotherms, 410, 411
  phase diagram, 422
CARNOT cycle, 497
CARNOT, SADI, 496
Cartesian diver, 185
CAVENDISH, H., 60
Cavitation, 190, 349
Celestial mechanics, 62
CELSIUS temperature scale, 373, 392
CELSIUS, A., 371
Center of gravity, 77, 93
Centrifugal force, 134, 177, 206
Centripetal force = radial force, 44
Chain 'syphon', 200
Changes of state
  by irreversible processes, 478
  of ideal gases, 398
  reversible, 487
Chemical potential, 384
Chimney, 206
CHLADNI sound patterns, 272
Circular frequency, 50
Circular motion, 44

Circular orbits, 26
Circulation, 229
CLAUSIUS, RUDOLF, 496
CLAUSIUS-CLAPEYRON equation, 509
Closed system, 478, 487
Cloud chamber, 425
Clouds, sink velocity, 216
Cold working, 167
Collision
  elastic, 79
  inelastic, 80
  non-central, 82
Combination oscillations, 280
Compressibility, 156, 177, 193
  isothermal, 398
  of water, 181
Compression, 153
Conical waves, 307, 354
Conservation of angular momentum, 115
Consonants, 357
Convection, 460
Cooling bottle, 378
Coriolis force, 138, 149, 435
Countercurrent heat exchanger, 420
Countercurrent process, 419
Coupled pendulums, 283
Critical temperature, 386, 414
Curl, 229
Current impulse, 75
Cutting, 171
Cyclic process, CARNOT, 497

D
DALTON's law, 394
Damping ratio, 273
Dancing balls, 75
dB(A), 361
de LAVAL nozzle, 492
Dead water, 340
DEBYE temperature, 389
Decibel (dB), 351
Decompression, 396
Deep drilling, 165
Deformation
  elastic, 153, 165
  of solid bodies, 31, 96
  of solids, 153
  plastic, 167
Deformation ellipsoid, 158
Degree Celsius (°C), 371, 392
Degrees of freedom, 123, 437
Delay in boiling, 425
Density, 38
  of air, 193
Density = Mass density, 38
Density fluctuations, statistical, 414, 448
DEWAR vessel, 376
Diabolo top, 116
DIESEL engine, 501
Difference frequency, 253, 358
Difference oscillation, 255, 280
Diffraction, 299
  by a slit, 311, 331
  FRAUNHOFER, 311
  FRESNEL, 311
Diffraction grating, 319
Diffusion, 176, 413, 432, 453, 469, 471
Diffusion boundary as surface
  for gases, 201
  of liquids, 413
Diffusion constant, 455, 456
Dilation, cubic, 156
Dipole, 228
Directional characteristic, 331
Directional hearing:, 359
Directional receiver, 356
Discus, 117, 121
Dispersion, 301, 341, 342, 347
Distance, 4
Distance measurement
  direct, 4
Doorbell, 249
DOPPLER effect, 293, 325
Drag, 212
Droplets, 191, 192
Dry ice, 423
DULONG-PETIT rule, 439
Dynamic pressure = Stagnation pressure, 223
Dynamics, 31

E
Ear, 362
  detection threshold, 360
  frequency range, 356
  model of the cochlea, 363
  spectral sensitivity distribution, 361
Earth
  as a top, 123
  as an accelerated frame of reference, 146, 148
  atmosphere, 198
Earth's rotation
  detection by HAGEN, 150
  verification by FOUCAULT, 149
Eddy-current damping, 274
Efficiency, 85
  of a heat engine, 496, 501
  of a heat pump, 503, 504
  of a refrigeration device, 503
  of muscles, 510
Elastic, 74
Elastic aftereffect, 165
Elastic aftereffects, 165
Elastic coefficient, 94
Elastic deformation, 153, 165, 264
Elastic oscillator, 52
Elementary oscillator, 259
Elementary waves, 301, 308, 312
Elevator feeling, 43
Elliptical orbit, 56
Elongation, 153, 154
Energy, 71
  bound, 506
  conservation of, 71, 74
  free, 505
    applications, 506
  internal, 379, 384, 396
  kinetic, 73, 98, 429
    of rotation, 97, 98, 100
  potential, 72
Energy conservation, 385
Energy release, stimulated, 278
Enthalpy, 387, 416, 489
Enthalpy of vaporization, 378, 387
Entropy, 479, 481, 489
  and probability, 483
  calculations, 483
  in a molecular picture, 482
Entropy changes of ideal gases, 486
Entropy elasticity, 507
Entropy increase, 481, 482
  by melting, 484
  on evaporation, 485
  on heating, 484
Entropy production, 481
Envelope, 410
Equation of motion
  for rotations, 97–99, 117
Equation of state
  caloric, 395
  for ideal gases, 391, 392
  for real gases, 414
  thermal, 384, 391
  VAN DER WAALS', 414, 415
Equilibrium
  indifferent, 47
  stable, 47
Equipartition principle, 437, 438
Escape velocity, 64
Expansion, 479, 480
Expansion coefficient, 370

F
Feedback, 247, 250, 283
FICK's law
  first, 455, 469
  second, 459
Field-line patterns
  in electrodynamics, 228
Figure axis, 112, 113

Fire pump, 405
First Law, 385, 496, 505
Fixed temperature points, 372
Flame, sensitive, 218
Floatation, 184
Flow
laminar, 214
turbulent, 217
Flow apparatus, 220
Flow around a disk, 225
Flow field, 220, 226
irrotational, 227
Flow limit, 155
Flow resistance, 234
Flow resistance around objects, 235
Flow through a nozzle, 490, 491
Flow velocity in a nozzle, 491
Flue pipe, 268, 269, 323
Fluid, 211
Focal length, 304
Focal point, 303, 311, 326
Fog formation, 411, 412, 425, 487
Force
as vector, 32
conservative, 73
hydrodynamic, 224
measurement of, 35
unit of, 35, 37
Force law
linear, 46, 52
nonlinear, 47
Forced oscillations, 273
amplitude, 275
phase difference between excitation and resonator, 275
Forces
aerodynamic, 266
hydrodynamic, 349
Formant range, 259, 357
FOUCAULT's pendulum experiment, 149
FOURIER analysis, 255, 259
FOURIER integral, 258
Frame of reference, 17, 33
accelerated, 129, 141, 206
with pure path acceleration, 129, 130
with pure radial acceleration, 130
Frame of reference with radial acceleration, 133
FRAUNHOFER diffraction, 311
Free axes, 108
of humans and animals, 111
Free balloon, 205
Free energy, 505
applications, 506
Free fall, 23, 38
Frequency, 13, 19, 27, 28
amplitude-dependent, 74
Frequency meter, 280
Frequency spectrum, 358
FRESNEL diffraction, 311
FRESNEL zones, 315, 331
Friction
external, 32, 169
internal, 211, 466, 470
sliding, 169
Frictional resistance, 212, 214, 216, 236
Frictional work, 219
Frictionless fluid motion, 220
Fundamental equation
of kinetic theory of gases, 196
of motion, 37
Fundamental vibration, 261

G

Gas
flow out of a nozzle, 491
ideal, 194
in an accelerated frame, 206
real, 194
Gas compressors, 406
Gas constant, universal, 393
Gas law, ideal, 391
Gas liquefaction, 420
Gas molecules, velocity, 198, 430, 435
Gas syphon, 201
Gas thermometer, 372
Gases
heat conduction, 465, 470
ideal, 195, 392, 395
internal friction, 466, 470
model experiment, 196, 204, 432
real, 409
transport processes, 465
Gauge block, 5
GAY-LUSSAC's throttle experiment, 395
Geometrical moment of inertia, 162, 163
GIBBS' phase rule, 423
Glass teardrops, 182
Gliding flight, 242
Gravitation, 34, 35
Gravitational constant, 60, 62
Gravity pendulum, 53
in a rotating frame of reference, 134, 138
Gravity pressure, 182, 183
in the air, 203
in water, 202
Group velocity, 341, 342
GUERICKE, OTTO VON, 199
Gyrocompass, 140, 150

H

HAGEN-POISEUILLE law, 216
Hair, diameter of, 6
Hardening, 155, 167
Harmonic oscillator, 50, 247
Harmonics, 261
Heat, 375
as random motion, 429
conversion into work, 495
latent, 377
Heat capacity, 376
molar, 376, 389
of gases, 391, 438
of some solids, 389
temperature dependence, 439
specific, 376, 388
Heat conduction, 460, 481
genuine, 460
in gases, 465, 470
non-stationary, 464
stationary, 463
Heat conductivity
temperature dependence, 463
Heat current, 463
Heat engine, 495
efficiency, 496, 501
technical, 500
Heat exchanger
countercurrent, 419, 420, 461
Heat of condensation, 377
Heat of crystallization, 378
Heat of melting, 378
Heat of transition, 379
Heat of vaporization, 377, 378, 450, 485
of water, 386
Heat pump, 501, 502
efficiency, 503, 504
Heat transport, 375, 460
through radiation, 375, 468
Heating pad, 379
Hectopascal (hPa), 178
HELE-SHAW flow, 215
Helium, superfluid, 191
HELMHOLTZ resonator, 273
HELMHOLTZ's equation, 506
Herpolhode cone, 116
Hertz = 1 s^{-1}, 13, 247
Hold point, 379
Hollow vortex, 231
HOOKE's law, 155, 322
Hour glass, 15
H-S diagram = MOLLIER diagram, 489
HUYGHENS' principle, 308
HUYGHENS, CHR., 26
Hydraulic press, 181

Hydrodynamic attraction, 225
Hydrofoil, 238
Hysteresis, 166
Hysteresis curve, 166
Hysteresis loop, 166

I
Ideal gas, 194, 195, 395
  changes of state, 398
  equation of state, 392
  molar heat capacity, 398
Image formation, 303
Image point, 303, 311
Impact = impulse, 75
Impulse, 75
Index of refraction, 302, 330
Indices, 296, 318
Individual unit of mass, 370
Inertia, 34
Inertial forces, 129
Initial vortex, 238
Instantaneous axis, 94
Interference, 295
Interferometer, 334
Internal combustion engine, 500
Internal energy, 379, 384, 396
Internal friction
  in liquids, 211
  independence of pressure, 466
internal friction, 470
Internal pressure, 416, 418
International temperature scale, 372
Intonation, 357
Inversion temperature, 417
Irradiance, 348, 360
Irreversibility, 479, 482
Irreversible process, 477, 481
Irrotational vortex field, 228, 231
Isochoric change of state, 400
Isotherm, 398
Isothermal change of state, 398
Isotherms
  of CO_2, 410, 411

J
Jet engines, 244
Jet propulsion, 86
JOULE, 68
JOULE-THOMSON effect, 417, 418, 420
JOULE-THOMSON throttle experiment, 416
Jumping techniques, 70

K
KELTING, H., 359
Kelvin
  as unit, 373
  temperature, 392
  temperature scale, 373, 504
KEMPELEN, W. V., 358
KEPLER, JOH., 62
KEPLER's elliptical orbits, 58, 63
KEPLER's laws, 62
Kilogram, 34
Kilopond, 38
Kilowatt (kW), 71
Kilowatt hour, 68
Kinematics, 17
Kinetic theory of gases, 196
Kite, 242
KNUDSEN effect, 471
KUNDT's dust figures, 266, 341, 405

L
Larynx, 357
Latent heat, 377
  of melting, 484
Latent heat of vaporization, 378, 444
Lattice constant, 13, 319, 334
Lattice plane, 329, 334
Law of areas, 63, 108
Law of gravity, 60
Law of reflection, 302
Law of refraction, 302
Law of thermodynamics
  first, 385, 496, 505
  second, 495, 505, 509
Leaf spring, 15, 279
Length, 4
Length measurement
  by interference, 7
  indirect, 7
  microscopic, 6
Lens for water waves, 303
Lens wheel, 263
Lift, 237, 240
  of a wing, 116, 239
Light mill, 433
LINDE, liquefaction of air, 419
Line spectrum, 256
Liquefaction
  of air, 419
  of gases, 418, 420
Liquid
  in an accelerated frame, 206
  supercooled, 424
  tensile strength of, 425
Liquid helium, 419
Liquid manometer, 184
Liquid surface, 412
  as diffusion boundary, 413
  in an accelerated frame of reference, 177
Liquids
  internal friction, 211
  tensile strength, 186
LISSAJOUS orbits, 57
Logarithmic decrement, 273
Longitudinal vibrations, 260, 263, 270
  of linear solid bodies, 263
Longitudinal waves, 321
  in air, 323
Loudness, 360
Loudspeaker, 354
Lubrication, 171
Lunar motion, 27

M
MACH's angle, 309, 354
Magdeburg hemispheres, 199
Magnitude of a quantity, 20
MAGNUS effect, 240
Manometer
  liquid, 184
Mass, 34
  molar, 370, 376, 393
Mass density, 38
Materials, solid, 508
MAX WIEN's experiment, 285
Maxima, 261
MAXWELL-BOLTZMANN velocity distribution, 450
MAXWELL's velocity distribution, 436
MAXWELL's wheel, 43
Mean free path, 435, 436, 467, 468
Mean squared deviation, 448
Melting temperature, 422
  pressure dependence, 508
Memory effect, 165
Metacenter, 185
Meter, 5–7
Metronome, 12
MICHELSON interferometer, 335
Microphones, 355
Microscope, 6
Midpoint of oscillation, 103
Mobility, 455
Model gas, 195, 197, 479
Model gas atmosphere, 446
Model of liquid structure, 414
Modulus of compression, 349, 405
Modulus of elasticity, 155
Molar heat capacity, 376, 389
  in a molecular picture, 437
  of ideal gases, 398
Molar mass, 370, 376, 393
  determination, 393
Molar volume, 370, 393
  of ideal gases, 393
Mole, 370
Molecular beam, 435
Molecular mass

comparison after R. BUNSEN, 431
Molecular velocity
calculation, 197
measurement, 435
Molecular weight, 376
MOLLIER diagram, 488, 489
MOLLWO, E., 287, 344, 346, 359
Moment of inertia, 97, 98
computation, 99
measurement of, 100
of a human body, 102, 108
of cylinders, 101
Momentary axis, 94
Momentum, 76, 470
Momentum conservation, 76, 77, 81
Motion
around a central point, 54
Muscular work, 510
Musical instruments, 262, 353
Muzzle report or bang, 354
Muzzle velocity, 81

N

Natural gas, 206
Neper (Np), 351
Neutral fiber, 161
NEWTON ISAAC, 59, 76, 87
Newton meter, 68
Newton, unit of force, 37
NEWTON's equation
applications, 41
NEWTON's fundamental equation of motion, 34
Nodal lines, 271
Nodes, 261, 297
Normal modes
in liquids and gases, 265
maxima, 261, 264
nodes of, 261, 264
of 2- and 3-dimensional bodies, 271
of linear solid bodies, 263
of stiff linear objects, 269
of vibration, 260
Normal stress, 159
Normalized stress, 154
Nose, blocked, 219
Note, 357
Number density, 369, 454
Nutation, 114, 115, 121
cone of, 116, 118

O

Object point, 303
Observation point, 315
Octave, 357
OHM, GEORG SIMON, 357
Optical lattice, 319
Optics, geometric, 302
Orbital angular momentum, 104
Ordinal number, 296
Oscillation
damped, 273
elliptically polarized, 56
gravity pendulum, 53
linearly polarized, 52
Oscillation frequency
of an elastic pendulum, 52
of torsional oscillations, 97
Oscillation period
of a gravity pendulum, 54
Oscillations, 247
amplification of, 282
damped, 258
excitation of, 247
forced, 147, 273
non-sinusoidal, 251, 262
of strings, 262
sinusoidal, 49, 252
spectral representation of, 256
superposition of, 252
Osmosis, 440
Osmotic pressure, 442, 444
OTTO engine, 501

P

Pan mill, 124
Parabolic trajectory, 64
Parametric excitation, 261, 298
Parametric excitation of oscillations, 108
Partial pressures, 394
PASCAL, BLAISE, 183
Pascal, unit of pressure, 178
Path acceleration, 21, 22, 54, 130
constant, 25
Path difference, 312
Pendulum
coupled, 283
mathematical, 54, 103
period of, 54
physical, 103
simple, 54
Pendulum clock
automatic control, 247
Pendulum length, reduced, 102, 103
Pendulum-Top, 121
Period, 13
of torsional oscillation, 101, 103
Phase, 50, 292
Phase angle = Phase, 50
Phase changes
of real gases, 409
Phase diagram
of CO_2, 422
of water, 422
Phase differences in sound waves, 357
Phase rule of GIBBS, 423
Phase shift, 51
Phase transition
hindrance, 424, 425
liquid-gas, 425
of CO_2, 413
solid-liquid, 412, 424
Phase velocity, 291, 341
Phon, 361
Phonometry, 360
Photophoresis, 434
Physical quantities, expressions, 38
Pipe vibrations, 268
Pirouette, 111
Pitch
of the voice, 357
PITOT tube, 224
PLANCK's radiation law, 373
Plasticity, 167
Pneumatic motors, 406, 504
POCKELS, AGNES, 191
POHL's pendulum, 273
Point mass, 47
POISSON's law, 401
POISSON's number, 155
POISSON's relation, 155
Polhode cone, 116
Polytropic curve, 402
Polytropic exponent, 402
measurement, 403
Potential flow, 226, 228
Power, 71
of a human, 71
of rotational motions, 97
PRANDTL tube, 224
Precession, 117
of the earth, 123
Precession cone, 118, 121
Press, hydraulic, 181
Pressure, 154, 178
in liquids, 178
in the lungs of a person, 394
osmotic, 442
production of, 195
static, in fluids, 223
units of, 178
Pressure distribution
in a gravitational field, 183
in air, 203
in gas lines, 205
in standing waves, 324
in water, 202
Pressure probe, 223
Pressure receiver, 355
Principal ray, 303
Principal stress, 159

Prism
for sound waves, 330
spectral resolving power, 347
Probability, 482, 483
Productivity, 226
Projectile velocity, 18
Propeller, 244
Proper angular momentum, 104
Pulsed excitation, 247, 257
p-V diagram, 382
of carbon dioxide, 410

Q

Quantities referred to the amount of substance (molar), 376
Quantity
derived, 29
of of substance, 369
physical, 28
Quasi-static processes, 476

R

Radar, 325
Radial acceleration, 21, 26, 44, 130, 133
Radial force, 44
Radian, 10, 95
Radiation power, 326, 348
Radiation pressure of sound, 326
Radiometer force, 433
Radius vector, 55, 62
Rapidity, 348
RAYLEIGH disk, 349
Real gases, 194, 409
isotherms, 410
phase changes of, 409
Recording, distortion-free, 281
Rectification, 420
Reference frame, 44
Reflection coefficient, 350
Reflection of sound waves by warm air, 329
Reflection of water waves, 302
Refraction of sound waves, 330
Refraction of water waves, 302
Refrigeration device, 501, 502
Refrigeration technology, basic rule, 503
Regenerator, 498
Relaxation oscillations, 288
Relaxation time, 165, 288
Resolving power, 347
of the ear, 357
Resonance, 276
importance for the detection of sinusoidal oscillations, 279
Resonance chamber, 352
Resonance curves, 276
Resonator, 276
Restoring force, 52
Reversible processes, 475, 480
Reversion pendulum, 103
REYNOLDS number, 217, 219
Ripple tank, 293
Rocket equation, 86
Rockets, 86
Rolling friction, 172
Rosette orbit
of a pendulum, 139, 148
Rotation pendulum
POHL's pendulum, 273
Rotation, sign or sense, 104
Rotations of liquids, 228
Rubber
elasticity, 370, 508
mechanical hysteresis, 166
RUBENS' flame tube, 266
Rule of DULONG-PETIT, 439
Rupture strength
of solids, 167

S

Sailing
close-hauled, 242
Satellite, artificial, 135
Saturation pressure, 377, 386, 422
Scales, 34
Scattering, 330
Schlieren
in air, 331
SCHULER period, 145
Screw caliper, 6
Screw micrometer, 5
Screws, loosening of, 171
Seasickness, 43
Second, 11
Second Law, 505, 509
of thermodynamics, 495
Secondary radiation, 325
Seismograph, 281
Separation fracture, 168
Separation of gases, 420
Shadow, 299, 328
colored, 3
Shaft
supple, 109
Shallow-water lens, 303
Shape memory alloys, 167
Shear modulus, 155, 157
Shear strain, 157
Shear stress, 157, 159
Shockwave, 307
SI = Système International d'Unités, 29
Simple pendulum, 53
Sine curve, 13, 251
Sink, 226
Sinking, 83
Sliding friction, 170
coefficient of, 170
Slit, 299
Solid bodies, 153
deformation, 31, 96, 153
Solids, 153, 508
Somersault, 111
Sonoluminescence, 190
Sound detection threshold, 361
Sound field
energy of, 328, 347
pressure amplitude, 348, 350
velocity amplitude, 349
Sound radiometer, 326
Sound receivers, 355
Sound shadows, 328
Sound source, 351
primary, 353
secondary, 353
Sound velocity, 322
in gases, 349
supersonic, 354, 488, 492
Sound wave resistance, 347
Sound waves
diffraction of, 331
in air, 323
standing, 323
Sound-pressure amplitude, 348
Source, 226
Speaking machine, 358
Specific, 39
Specific enthalpy of vaporization, 378
Specific heat, 376, 388
Specific heat of vaporization, 444
Specific volume, 39
Spectral apparatus, 279
Spectral representation, 256
Spin, 104
Spiral illusion, 4
Spring balance, 36
Spring constant, 46, 52, 94
Spring pendulum, 52, 53
Stabilizing tank, 286
Stagnation point, 223
Stagnation pressure, 223
Standard conditions, 393
Standing waves, 297, 323
State variables
critical, 409
for water, 485
simple, 369
thermal, 384
Statistical fluctuations, 447
Steam
superheated, 411
Steam engine, 500

Steamboats, 236
STEINER's law, 100, 102, 143
Steradian, 10
Stereogrammetry, 8
Sticking friction, 169
Stimulated energy transfer, 278
STIRLING engine, 498, 501
STOKES' law, 216
Strain, 157
Strain ellipsoid, 158
Stream formation, 233
Streamlines, 216, 220, 225, 237, 239
  in a model experiment, 224, 225, 239
  in a water wave, 336
Stretching limit, 155
String
  emission of, 352
  vibrations of, 262, 352
Stroboscope, 14
Sublimation, 423
Sublimation temperature, 422
Sum frequency, 358
Supercooling, 424
Superfluidity of helium, 191
Supersonic jet, 492
Supersonic velocities, 354, 488
Surface
  formation of, 413, 425
Surface defects, 168
Surface energy, 168, 187, 188
Surface tension, 188
  abnormal, 192
Surface waves on water, 336
Syphon, 200
  for gases, 201
  for liquids, 200

T
Technical work, 381, 387
Temperature, 370
  critical, 386, 414
  on the molecular scale, 429
  thermodynamic definition, 504
Temperature conductivity, 464
Temperature scale, international, 372
Tennis ball, 'cut', 241
Tensile strength
  of liquids, 186, 426
  of solids, 168
  of water, 186, 426
Tensile stress, 154, 426
Thermal conductivity, 463, 470
Thermal expansion coefficient, 370
Thermal motions
  in gases, 195
  in solids, 175
  of molecules, 430
Thermal vibrations, 271
Thermocouple, 372
Thermodiffusion, 473
Thermometer, 371, 372
Thermos bottle, 376
Throttle experiment of GAY-LUSSAC, 396
Throttled release, 396
Tides, 147
Tightrope artist, 125
Time, 10, 16
Time integral, 75
Time measurement, 10
  true, 10
Time measurements
  indirect, 15
Toggle oscillations, 288
Tone, 357
Top, 105, 112
  angular momentum axis, 113
  force-free, 115
  instantaneous axis of rotation, 113
  momentary axis of rotation, 113
  precessional oscillations, 125
  symmetry axis, 113
  with two degrees of freedom, 123
Torque, 97
  definition of, 92
  of an electric motor, 93
  production of, 94
Torsion, 163, 164
Torsion balance, 61
Torsion coefficient, 94, 97, 163
Torsion shaft, 94, 95
Torsional modulus = shear modulus, 157
Torsional oscillation, 97, 100, 264, 271, 273
Torsional pendulum, 95
Torsional waves, 322
Total pressure = dynamic head
  in fluids, 223
Total reflection, 304, 306
Trade winds, 150
Transport processes in gases, 465
Transverse acceleration, 21
Transverse contraction, 155
Transverse force
  applications, 241
  dynamic, 237, 240
Transverse vibrations, 259, 260, 270
Transverse waves, 322
Travelling waves, 291
TREVELYAN's rocker, 288
Triple point, 411, 421, 423
TROUTON's rule, 485
Tsunami, 340
Tuning fork, 282, 352
  auto-control, 249
  damping, 285
  improvement of emission, 353
Tunnel effect, 306
Turbulence, 217, 414
Twisting, 95, 96, 163

U
Ultrasound, 268
Unit of work, 68
Unit system
  SI, 29
Units, 5
Universal gas constant, 393

V
VAN DER WAALS' equation of state, 414, 415
Vapor
  saturation pressure, 377
Vapor pressure, 421, 450
Vapor-pressure curve
  of ice, 424
  of supercooled water, 424
Vector, 20
Vector product, 92
Velocity, 28
  definition, 18, 29
  of a bullet, 137
  of molecules, 198, 430
  vector addition, 20
Velocity distribution
  in laminar flow, 215
  MAXWELL-BOLTZMANN, 436, 450
  of gas molecules, 435
Velocity of light, 88
Velocity of sound, 405
Velocity receiver, 355
Vernier caliper, 5
Vibrating-reed frequency meter, 279, 363
Vibrations, 292
  excitation of, 247
  in pipes, 268
  of a doorbell, 249
  of a gravity pendulum, 248
  of strings, 352
  self-controlling, 247
  undamped, 247
Violin
  as sound source, 353
  frequency spectrum, 354
Viscosity, 212, 213, 470
  dynamic, 218
  kinematic, 218

of gases, 466
Voltage impulse, 75
Volume current, 214, 216
Volume, molar, 393
Vortex field, irrotatonal, 228, 231
Vortex ring, 234
Vortex tube, 474
Vortices and separation surfaces, 232
Vorticity, 231
Vowels, 357

W

Water-wave experiments, 294
Watt = N m/s, 71
Watt second, 68
Wave field on a water surface, 294
Wave maxima, 297
Wave tank, 336
Waves
diffracted, 299
on liquid surfaces, 336
standing, 297
travelling, 291
Weight, 31, 35
of the moon, 59
WESTPHAL, W.H., 434
Wet steam, 411
Wetting, 185, 189
Wobble oscillations, 286
amplitude dependence of the frequency, 286
Work, 67
against a spring, 69
definition, 68
in rotations, 97
of acceleration, 70, 396, 490
of lifting, 68, 69
of rotation, 92
Work of displacement, 381
Work of expansion, 381, 383
Work, technical, 383

Y

YOUNG, THOMAS, 293, 295, 317, 334
YOUNG's modulus = modulus of elasticity, 155

Z

Zone plate, 317
Zones, FRESNEL, 315